16

Adaptive Control for a Class of Non-affine Nonlinear Systems via Neural Networks

Zhao Tong
School of Automatization and Electronic Engineering, Qingdao University of Science and Technology
China

1. Introduction

Adaptive control of highly uncertain nonlinear dynamic systems has been an important research area in the past decades, and in the meantime neural networks control has found extensive application for a wide variety of areas and has attracted the attention of many control researches due to its strong approximation capability. Many significant results on these topics have been published in the literatures (Lewis et al., 1996 ; Yu & Li 2002; Yesidirek & Lewis 1995). It is proved to be successful that neural networks are used in adaptive control. However, most of these works are applicable for a kind of affine systems which can be linearly parameterized. Little has been found for the design of specific controllers for the nonlinear systems, which are implicit functions with respect to control input. We can find in literatures available there are mainly the results of Calise et al. (Calise & Hovakimyan 2001) and Ge et al. (Ge et al. 1997). Calise et al. removed the affine in control restriction by developing a dynamic inversion based control architecture with linearly parameterized neural networks in the feedback path to compensate for the inversion error introduced by an approximate inverse. However, the proposed scheme does not relate to the properties of the functions, therefore, the special properties are not used in design. Ge, S.S. et al., proposed the control schemes for a class of non-affine dynamic systems, using mean value theorem, separate control signals from controlled plant functions, and apply neural networks to approximate the control signal, therefore, obtain an adaptive control scheme.
Furthermore, when controlling large-scale and highly nonlinear systems, the presupposition of centrality is violated due to either due to problems in data gathering when is spread out or due to the lack of accurate mathematical models. To avoid the difficulties, the decentralized control architecture has been tried in controller design. Decentralized control systems often also arise from various complex situations where there exist physical limitations on information exchange among several subsystems for which there is insufficient capability to have a single central controller. Moreover, difficulty and uncertainty in, measuring parameter values within a large-scale system may call for adaptive techniques. Since these restrictions encompass a large group of applications, a variety of decentralized adaptive techniques have been developed (Ioannou 1986).

Earlier literature on the decentralized control methods were focused on control of large-scale linear systems. The pioneer work by Siljak (Siljak 1991) presents stability theorems of interconnected linear systems based on the structure information only. Many works consider subsystems which are linear in a set of unknown parameters (Ioannou 1986 ; Fu 1992 ; Sheikholeslam & Desor 1993 ; Wen 1994 ; Tang et al. 2000), and these results were focused on systems with first order interconnections. When the subsystems has nonlinear dynamics or the interconnected is entered in a nonlinear fashion, the analysis and design problem becomes even challenging.

The use of neural networks' learning ability avoids complex mathematical analysis in solving control problems when plant dynamics are complex and highly nonlinear, which is a distinct advantage over traditional control methods. As an alternative, intensive research has been carried out on neural networks control of unknown nonlinear systems. This motivates some researches on combining neural networks with adaptive control techniques to develop decentralized control approaches for uncertain nonlinear systems with restrictions on interconnections. For example, in (Spooner & Passino 1999), two decentralized adaptive control schemes for uncertain nonlinear systems with radial basis neural networks are proposed, which a direct adaptive approach approximates unknown control laws required to stabilize each subsystem, while an indirect approach is provided which identifies the isolated subsystem dynamics to produce a stabilizing controller. For a class of large scale affine nonlinear systems with strong interconnections, two neural networks are used to approximate the unknown subsystems and strong interconnections, respectively (Huang & Tan 2003), and Huang & Tan (Huang & Tan 2006) introduce a decomposition structure to obtain the solution to the problem of decentralized adaptive tracking control a class of affine nonlinear systems with strong interconnections. Apparently, most of these results are likewise applicable for affine systems described as above. For the decentralized control research of non-affine nonlinear systems, many results can be found from available literatures. Nardi et al. (Nardi & Hovakimyan 2006) extend the results in Calise et al. (Calise & Hovakimyan 2001) to non-affine nonlinear dynamical systems with first order interconnections. Huang (Huang & Tan 2005) apply the results in (Ge & Huang 1999) to a class of non-affine nonlinear systems with strong interconnections.

Inspired by the above researches, in this chapter, we propose a novel adaptive control scheme for non-affine nonlinear dynamic systems. Although the class of nonlinear plant is the same as that of Ge et al. (Ge et al. 1997), utilizing their nice reversibility, and invoking the concept of pseudo-control and inverse function theorem, we find the equitation of error dynamics to design adaptation laws. Using the property of approximation of two-layer neural networks (NN), the control algorithm is gained. Then, the controlled plants are extended to large-scale decentralized nonlinear systems, which the subsystems are composed of the class of non-affine nonlinear functions. Two schemes are proposed, respectively. The first scheme designs a RBFN-based (radial basis function neural networks) adaptive control scheme with the assumption which the interconnections between subsystems in entire system are bounded linearly by the norms of the tracking filtered error. In the scheme, unlike most of other approaches in available literatures, the weight of BBFN and center and width of Gaussian function are tuned adaptively. In another scheme, the interconnection is assumed as stronger nonlinear function. Moreover, in the former, in every subsystem, a RBFN is adopted which is used to approximate unknown function, and in the latter, in every subsystem, two RBFNs are respectively utilized to approximate unknown

function and uncertain strong interconnection function. For those complicated large-scale decentralized dynamic systems, in order to decrease discontinuous factors and make systems run smooth, unlike most of control schemes, the hyperbolic tangent functions are quoted in the design of robust control terms, instead of sign function. Otherwise, the citation of the smooth function is necessary to satisfy the condition of those theorems.
The rest of the paper is organized as follows. Section 2 gives the normal form of a class of non-affine nonlinear systems. Section 3 proposes a novel adaptive control algorithm, which is strictly derived from some mathematical and Lyapunov stability theories, and the effectiveness of the scheme is validated through simulation. Extending the above-mentioned result, Section 4 discusses two schemes of decentralized adaptive neural network control for the class of large-scale nonlinear systems with linear function interconnections and nonlinear function interconnections, respectively. Finally, the Section 5 is concluding remarks.

2. Problem Statement

We consider a general analytic system

$$\begin{cases} \dot{\boldsymbol{\zeta}} = \mathbf{g}(\boldsymbol{\zeta}, u), & \boldsymbol{\zeta} \in R^n, \quad u \in R \\ y = h(\boldsymbol{\zeta}), & y \in R. \end{cases} \tag{1}$$

where $\mathbf{g}(\cdot,\cdot)$ is a smooth vector fields and $h(\cdot)$ is a scalar function. In practice, many physical systems such as chemical reactions, PH neutralization and distillation columns are inherently nonlinear, whose input variables may enter in the systems nonlinearly as described by the above general form (Ge et al. 1998). Then, the Lie derivative (Tsinias & Kalouptsidis 1983) of $h(\boldsymbol{\zeta})$ with respect to $\mathbf{g}(\boldsymbol{\zeta}, u)$ is a scalar function defined by $L_{\mathbf{g}} h = [\partial h(\boldsymbol{\zeta}) / \partial \boldsymbol{\zeta}] \mathbf{g}(\boldsymbol{\zeta}, u)$. Repeated Lie derivatives can be defined recursively as $L_{\mathbf{g}}^i h = L_{\mathbf{g}}(L_{\mathbf{g}}^{i-1} h), \quad for \quad i = 1, 2 \cdots$. The system (1) is said to have relative degree α at $(\boldsymbol{\zeta}_0, u_0)$, if there exists a smallest positive integer α such that $\partial L_{\mathbf{g}}^i h / \partial u = 0$, $\partial L_{\mathbf{g}}^{\alpha} h / \partial u \neq 0, \quad i = 1, \cdots, \alpha - 1$.

Let $\Omega_\zeta \subset R^n$ and $\Omega_u \subset R$ be compact subsets containing $\boldsymbol{\zeta}_0$ and u_0, respectively. System (1) is said to have a strong relative degree α in a compact set $D = \Omega_\zeta \times \Omega_u$, if it has relative degree α at every point $(\boldsymbol{\zeta}_0, u_0) \in D$. Therefore, system (1) is feedback linearizable and the mapping $\Phi(\boldsymbol{\zeta}) = [\phi_1(\boldsymbol{\zeta}), \phi_2(\boldsymbol{\zeta}), \cdots \phi_n(\boldsymbol{\zeta})]$, with $\phi_j(\boldsymbol{\zeta}) = L_{\mathbf{g}}^{j-1} h, \quad j = 1, 2, \cdots \alpha$ has a Jacobian matrix which is nonsingular for all $\mathbf{x} \in \Phi(\boldsymbol{\zeta})$, system (1) can be transformed into a normal form

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = x_3 \\ \quad \vdots \\ \dot{x}_n = f(x,u) \\ y = x_1 \end{cases} \tag{2}$$

where $f(x,u) = L_{\mathbf{g}}^n h$ and $x = \Phi^{-1}(\zeta)$ with $x = [x_1, x_2, \cdots, x_n]^T$. Define the domain of normal system (2) as $\bar{D} \triangleq \{(x,u) | x \in \Phi(\Omega_\zeta); u \in \Omega_u\}$.

3. Adaptive Control for a Class of Non-affine Nonlinear Systems via Two-Layer Neural Networks

Now we consider the $n-th$ order nonlinear systems of the described form as (2). For the considered systems in the chapter, we may make the following assumptions.

Assumption 1. $\partial f(x,u)/\partial u \neq 0$ for all $(x,u) \in \Omega \times R$.

Assumption 2. $f(\cdot): R^{n+1} \to R$, is an unknown continuous function and $f(x,u)$ a smooth function with respect to control input u.

The control objective is: determine a control law, force the output, y , to follow a given desired output, x_d with an acceptable accuracy, while all signals involved must be bounded.

Assumption 3. The desired signals $x_d(t) = [y_d, y_d^{(1)}, \cdots, y_d^{(n-1)}]$, and $X_d = [x_d^T, y_d^{(n)}]^T$ are bounded, with $\|X_{\mathbf{d}}\| \leq \bar{X}_d$, $\bar{X}_d$ a known positive constant.

Define the tracking error vector as

$$e = x - x_d, \tag{3}$$

and a filtered tracking error as

$$\tau = [\Lambda^T \quad 1]e, \tag{4}$$

with Λ a gain parameter vector selected so that $e(t) \to 0$ as $\tau \to 0$. Differentiating (4), the filtered tracking error can be written as

$$\dot{\tau} = \dot{x}_n - x_d^{(n)} + [0 \quad \Lambda^T]\mathbf{e}. \tag{5}$$

Define a continuous function

$$\delta = -k\tau + x_d^{(n)} - [0 \quad \Lambda^T]\mathbf{e}. \tag{6}$$

where k is a positive constant. We know $\partial f(x,u)/\partial u \neq 0$ (Assumption 1), thus, $\partial[f(x,u)-\delta]/\partial u \neq 0$. Considering the fact that $\partial\delta/\partial u = 0$, we invoke the implicit function theorem (Lang 1983), there exists a continuous ideal control input u^* in a neighborhood of $(x,u) \in \Omega \times R$, such that $f(x,u^*) - \delta = 0$, i.e. $\delta = f(x,u^*)$ holds. $\delta = f(x,u^*)$ may represent ideal control inverse.

Adding and subtracting δ to the right-hand side of $\dot{x}_n = f(x,u)$ of (2), one obtains

$$\dot{x}_n = f(x,u) - \delta - k\tau + x_d^{(n)} - [0 \quad \Lambda^T]e, \tag{7}$$

and yields

$$\dot{\tau} = -k\tau + f(x,u) - \delta. \tag{8}$$

Considering the following state dependent transformation $\psi = \dot{x}_n$, where ψ is commonly referred to as the pseudo-control (Calise & Hovakimyan 2001). Apparently, the pseudo-control is not a function of the control u but rather a state dependent operator. Then, $\partial\psi/\partial u = 0$, from Assumption 1, $\partial f(x,u)/\partial u \neq 0$ thus $[\psi - f(x,u)]/\partial u \neq 0$. With the implicit function theorem, for every $(x,u) \in \Omega \times R$, there exists a implicit function such that $\psi - f(x,u) = 0$ holds, i.e. $\psi = f(x,u)$. Therefore, we have

$$\psi = f(x,u). \tag{9}$$

Furthermore, using inverse function theorem, with the fact that $[\psi - f(x,u)]/\partial u \neq 0$ and $f(x,u)$ is a smooth with respect to control input, u , then, $f(x,u)$ defines a local diffeomorphism (Slotine & Li 1991), such that, for a neighborhood of u, there exists a smooth inverse function and $u = f^{-1}(x,\psi)$ holds. If the inverse is available, the control problem is easy. But this inverse is not known, we can generally use some techniques, such as neural networks, to approximate it. Hence, we can obtain an estimated function, $\hat{u} = f^{-1}(x,\hat{\psi})$. This result in the following equation holding:

$$\hat{\psi} = f(x,\hat{u}), \tag{10}$$

where $\hat{\psi}$ may be referred to as approximation pseudo-control input which represents actual dynamic approximation inverse.

Remark 1. According to the above-mentioned conditions, when one designs the pseudo-control signal, $\hat{\psi}$, must be a smooth function. Therefore, in order to satisfy the condition, we adopt hyperbolic tangent function, instead of sign function in design of input. This also makes control signal tend smooth and system run easier. The hyperbolic tangent function has a good property as follows (Polycarpou 1996) :

$$0 < |\eta| - \eta \tanh(\frac{\eta}{\alpha}) \le \varsigma\alpha, \tag{11}$$

with $\varsigma = 0.2785$, α any positive constant. Moreover, theoretically, $\hat{\psi}$ is approximation inverse, generally a nonlinear function, but it must be bounded and play a dynamic approximation role and make system stable. Hence, it represents actual dynamic approximation inverse.
Based on the above conditions, in order to control the system and make it be stable, we design the approximation pseudo-control input $\hat{\psi}$ as follows:

$$\hat{\psi} = f(x, u^*) + u_{ad} + v_r, \tag{12}$$

where u_{ad} is output of a neural network controller, which adopts a two-layer neural network, v_r is robustifying control term designed in stability analysis.
Adding and subtracting $\hat{\psi}$ to the right-hand side of (8), with $\delta = f(x, u^*)$, we have

$$\begin{aligned} \dot{\tau} &= -k\tau + f(x,u) + \hat{\psi} - f(x,u^*) - u_{ad} - v_r - \delta \\ &= -k\tau + \tilde{\Delta}(x,u,u^*) + \hat{\psi} - \delta - u_{ad} - v_r, \end{aligned} \tag{13}$$

where $\tilde{\Delta}(x,u,u^*) = f(x,u) - f(x,u^*)$ is error between nonlinear function and its ideal control function, we can use the neural network to approximate it.

3.1 Neural network-based approximation

A two-layer NN consists of two layers of tunable weights, a hidden layer and an output layer. Given a $\varepsilon > 0$, there exists a set of bounded weights M and N such that the nonlinear error $\tilde{\Delta} \in C(\Omega)$, with Ω compact subset of R^n, can be approximated by a two-layer neural network, i.e.

$$\tilde{\Delta} = M^T \sigma(N^T x_{nn}) + \varepsilon(x_{nn}), \tag{14}$$

with $x_{nn} = [1, x_d^T, e^T, \hat{\psi}]$ input vector of NN.

Assumption 4. The approximation error ε is bounded as follows:

$$\left|\varepsilon\right| \leq \varepsilon_N , \tag{15}$$

where $\varepsilon_N > 0$ is an unknown constant.

Let $\hat{M}$ and $\hat{N}$ be the estimates respectively of M and N . Based on these estimates, let u_{ad} be the output of the NN

$$u_{ad} = \hat{M}^T \sigma(\hat{N}^T x_{nn}). \tag{16}$$

Define $\tilde{M} = M - \hat{M}$ and $\tilde{N} = N - \hat{N}$, where we use notations: $Z = diag[M, N]$, $\tilde{Z} = diag[\tilde{M}, \tilde{N}]$, $\hat{Z} = diag[\hat{M}, \hat{N}]$ for convenience. Then, the following inequality holds:

$$tr(\tilde{Z}^T \hat{Z}) \leq \left\|\tilde{Z}\right\|_F \left\|Z\right\|_F - \left\|\tilde{Z}\right\|_F^2 . \tag{17}$$

The Taylor series expansion of $\sigma(N^T x_{nn})$ for a given x_{nn} can be written as:

$$\sigma(N^T x_{nn}) = \sigma(\hat{N}^T x_{nn}) + \sigma'(\hat{N}^T x_{nn})\tilde{N}^T x_{nn} + O(\tilde{N}^T x_{nn})^2 , \tag{18}$$

with $\hat{\sigma} := \sigma(\hat{N}^T x_{nn})$ and $\hat{\sigma}'$ denoting its Jacobian, $O(\tilde{N}^T x_{nn})^2$ the term of order two. In the following, we use notations: $\sigma := \sigma(N^T x_{nn})$, $\tilde{\sigma} := \sigma(\tilde{N}^T x_{nn})$.

With the procedure as Appendix A, the approximation error of function can be written as

$$M^T \sigma(N^T x_{nn}) - \hat{M}^T \sigma(\hat{N}^T x_{nn}) = \tilde{M}^T (\hat{\sigma} - \hat{\sigma}' \hat{N}^T x_{nn}) + \hat{M}^T \hat{\sigma}' \tilde{N}^T x_{nn} + \omega , \tag{19}$$

and the disturbance term ω can be bounded as

$$\left|\omega\right| \leq \left\|N\right\|_F \left\|x_{nn} \hat{M}^T \hat{\sigma}'\right\|_F + \left\|M\right\| \left\|\hat{\sigma}' \hat{N}^T x_{nn}\right\| + \left\|M\right\|_1 , \tag{20}$$

where the subscript "F" denotes Frobenius norm, and the subscript "1" the 1-norm. Redefine this bound as

$$\left|\omega\right| \leq \rho_\omega \vartheta_\omega(\hat{M}, \hat{N}, x_{nn}) , \tag{21}$$

where $\rho_\omega = \max\{\|M\|, \|N\|_F, \|M\|_1\}$ and $\vartheta_\omega = \left\|x_{nn}\hat{M}^T\hat{\sigma}'\right\|_F + \left\|\hat{\sigma}'\hat{N}^T x_{nn}\right\| + 1$. Notice that ρ_ω is an unknown coefficient, whereas ϑ_ω is a known function.

3.2 Parameters update law and stability analysis

Substituting (14) and (16) into (13), we have

$$\dot{\tau} = -k\tau + M^T\sigma(N^T x_{nn}) - \hat{M}^T\sigma(\hat{N}^T x_{nn}) + \hat{\psi} - v_r - \delta + \varepsilon(x_{nn}). \tag{22}$$

Using(19), the above equation can become

$$\dot{\tau} = -k\tau + \tilde{M}^T(\hat{\sigma} - \hat{\sigma}'\hat{N}^T x_{nn}) + \hat{M}^T\hat{\sigma}'\tilde{N}^T x_{nn} + \hat{\psi} - \delta - v_r + \omega + \varepsilon. \tag{23}$$

Theorem 1. Consider the nonlinear system represented by Eq. (2) and let Assumption 1-4 hold. If choose the approximation pseudo-control input $\hat{\psi}$ as Eq.(12), use the following adaptation laws and robust control law

$$\begin{aligned}
\dot{\hat{M}} &= F\left[(\hat{\sigma} - \hat{\sigma}'Nx_{nn})\tau - k_1\hat{M}|\tau|\right], \\
\dot{\hat{N}} &= R\left[x_{nn}\hat{M}^T\hat{\sigma}'\tau - k_1\hat{N}|\tau|\right], \\
\dot{\hat{\phi}} &= \gamma\left\{\tau(\vartheta_\omega + 1)\tanh\left[\frac{\tau(\vartheta_\omega + 1)}{\alpha}\right] - \lambda\hat{\phi}\right\} \\
v_r &= -\hat{\phi}(\vartheta_\omega + 1)\tanh\left[\frac{\tau(\vartheta_\omega + 1)}{\alpha}\right]
\end{aligned} \tag{24}$$

where $F = F^T > 0, R = R^T > 0$ are any constant matrices, $k_1 > 0$ and $\gamma > 0$ are scalar design parameters, $\hat{\phi}$ is the estimated value of the uncertain disturbance term $\phi = \max(\rho_\omega, \varepsilon_N)$, defining $\tilde{\phi} = \phi - \hat{\phi}$ with $\tilde{\phi}$ error of ϕ, then, guarantee that all signals in the system are uniformly bounded and that the tracking error converges to a neighborhood of the origin.

Proof. Consider the following positive define Lyapunov function candidate as

$$L = \frac{1}{2}\tau^2 + \frac{1}{2}tr(\tilde{M}^T F^{-1}\tilde{M}) + \frac{1}{2}tr(\tilde{N}^T R^{-1}\tilde{N}) + \frac{1}{2}\gamma^{-1}\tilde{\phi}^2 \tag{25}$$

The time derivative of the above equation is given by

$$\dot{L} = \dot{\tau}\tau + tr(\tilde{M}^T F^{-1}\dot{\tilde{M}}) + tr(\tilde{N}^T R^{-1}\dot{\tilde{N}}) + \gamma^{-1}\tilde{\phi}\dot{\tilde{\phi}} \tag{26}$$

Substituting (23) and the anterior two terms of (24) into (26), after some straightforward manipulations, we obtain

$$
\begin{aligned}
\dot{L} &= -k\tau^2 + \tau[\tilde{M}^T(\hat{\sigma} - \hat{\sigma}'\hat{N}^T x_{nn}) + \hat{M}^T\hat{\sigma}'\tilde{N}^T x_{nn} + (\hat{\psi} - \delta) - v_r + \omega + \varepsilon] \\
&\quad + tr(\tilde{M}^T F^{-1}\dot{\hat{M}}) + tr(\tilde{N}^T R^{-1}\dot{\hat{N}}) + \gamma^{-1}\tilde{\phi}\dot{\hat{\phi}} \\
&= -k\tau^2 + \tau(\hat{\psi} - \delta) - \tau v_r + \tau(\omega + \varepsilon) + \gamma^{-1}\tilde{\phi}\dot{\hat{\phi}} + k_1|\tau| tr(\tilde{Z}^T\hat{Z}). \\
&\le -k\tau^2 + \tau(\hat{\psi} - \delta) - \tau v_r + |\tau|\phi(\vartheta_\omega + 1) + \gamma^{-1}\tilde{\phi}\dot{\hat{\phi}} + k_1|\tau| tr(\tilde{Z}^T\hat{Z}).
\end{aligned}
\tag{27}
$$

With (4),(6),(12),(16) and the last two equations of (24), the approximation error between actual approximation inverse and ideal control inverse is bounded by

$$
|\hat{\psi} - \delta| \le c_1 + c_2|\tau| + c_3\|\tilde{Z}\|_F, \tag{28}
$$

where c_1, c_2, c_3 are positive constants.

Using (11) and the last two terms of (24), we obtain

$$
\begin{aligned}
\dot{L} &\le -k\tau^2 + \tau(\hat{\psi} - \delta) - \tau\hat{\phi}(\vartheta_\omega + 1)\tanh\left[\frac{\tau(\vartheta_\omega + 1)}{\alpha}\right] \\
&\quad + |\tau|\phi(\vartheta_\omega + 1) - \tilde{\phi}\left\{\tau(\vartheta_\omega + 1)\tanh\left[\frac{\tau(\vartheta_\omega + 1)}{\alpha}\right] - \lambda\hat{\phi}\right\} + k_1|\tau| tr(\tilde{Z}^T\hat{Z}) \\
&\le -k\tau^2 + \tau(\hat{\psi} - \delta) + \varsigma\phi\alpha + \lambda\tilde{\phi}\hat{\phi} + k_1|\tau| tr(\tilde{Z}^T\hat{Z})
\end{aligned}
\tag{29}
$$

Applying (17),(28) , and $\tilde{\phi}\hat{\phi} \le |\tilde{\phi}||\phi| - |\tilde{\phi}|^2$, after completing square, we have the following inequality

$$
\dot{L} \le -(k - c_2)|\tau|^2 + D_1|\tau| + D_2 \tag{30}
$$

where $D_1 = c_1 + \frac{k_1}{4}(Z_M + \frac{c_3}{k_1})^2, D_2 = \frac{1}{4}\lambda\phi^2 + \varsigma\phi\alpha$.

Let $D_3 = \sqrt{D_1^2 + 4D_2(k - c_2)} + D_1$, thus, as long as $|\tau| \ge D_3/[2(k - c_2)]$, and $k > c_2$, then $\dot{L} \le 0$ holds.

Now define

$$\Omega_\phi = \left\{\tilde{\phi} \,\middle|\, \left|\tilde{\phi}\right| \le \phi\right\},\ \ \Omega_Z = \left\{\tilde{Z} \,\middle|\, \left\|\tilde{Z}\right\|_F \le \frac{1}{k_1}(k_1 Z_M + c_3)\right\},\ \ \Omega_\tau = \left\{\tau \,\middle|\, \left|\tau\right| \le \frac{1}{2(k-c_2)} D_3\right\}. \tag{31}$$

Since $Z_M, k_1, k, D_1, D_2, D_3, c_2, c_3$ are positive constants, as long as k is chosen to be big enough, such that $k > c_2$ holds, we conclude that Ω_ϕ, Ω_Z and Ω_τ are compact sets. Hence $\dot{L}$ is negative outside these compacts set. According to a standard Lyapunov theorem, this demonstrates that $\tilde{\phi}, \tilde{Z}$ and τ are bounded and will converge to Ω_ϕ, Ω_Z and Ω_τ, respectively. Furthermore, this implies e is bounded and will converge to a neighborhood of the origin and all signals in the system are uniformly bounded.

3.3 Simulation Study

In order to validate the performance of the proposed neural network-based adaptive control scheme, we consider a nonlinear plant, which described by the differential equation

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -\omega^2 x_1 - 0.02(\omega + x_1^2)x_2 + u^3 + (x_1^2 + x_2^2)\sigma(u) + \tanh(0.2u) + d \end{aligned} \tag{32}$$

where $\omega = 0.4\pi$, $\sigma(u) = (1-e^{-u})/(1+e^{-u})$ and $d = 0.2$. The desired trajectory $x_d = 0.1\pi[\sin(2t) - \cos(t)]$.

To show the effectiveness of the proposed method, two controllers are studied for comparison. A fixed-gain PD control law is first used as Polycarpou, (Polycarpou 1996). Then, the adaptive controller based on NN proposed is applied to the system.

Input vector of neural network is $x_{nn} = [1, x_d^T, e^T, \hat{\psi}]$, and number of hidden layer nodes 25. The initial weight of neural network is $\hat{M}(0) = (0), \hat{N}(0) = (0)$. The initial condition of controlled plant is $x(0) = [0.1, 0.2]^T$. The other parameters are chosen as follows: $k_1 = 0.01, \gamma = 0.1, \lambda = 0.01, \alpha = 10$, $\Lambda = 2, F = 8I_M$, $R = 5I_N$, with I_M, I_N corresponding identity matrices.

Fig.1, 2, and 3 show the results of comparisons, the PD controller and the adaptive controller based on NN proposed, of tracking errors, output tracking and control input, respectively. These results indicate that the adaptive controller based on NN proposed presents better control performance than that of the PD controller. Fig.4 depicts the results of output of NN, norm values of $\hat{M}, \hat{N}$, respectively, to illustrate the boundedness of the estimates of $\hat{M}, \hat{N}$ and the control role of NN. From the results as figures, it can be seen that the learning rate of neural network is rapid, and tracks objective in less than 2 seconds. Moreover, as desired, all signals in system, including control signal, tend to be smooth.

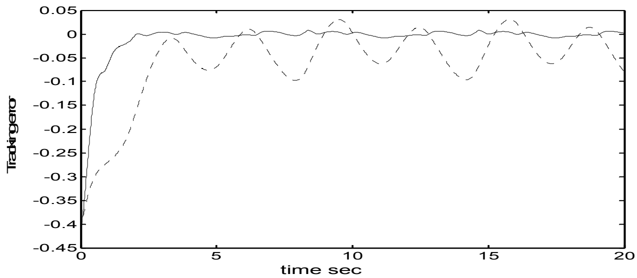

Fig. 1. Tracking errors: PD(dot) and NN(solid).

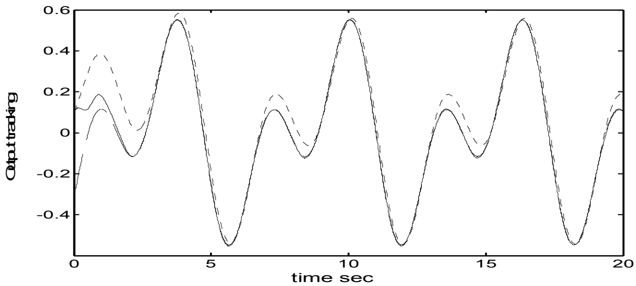

Fig. 2. Output tracking: desired (dash), NN(solid) and PD(dot).

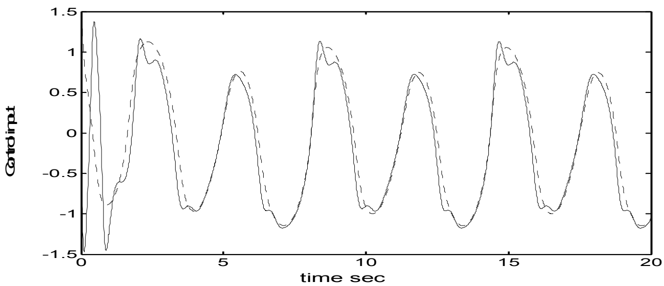

Fig. 3. Control input: PD (dash), NN(solid)

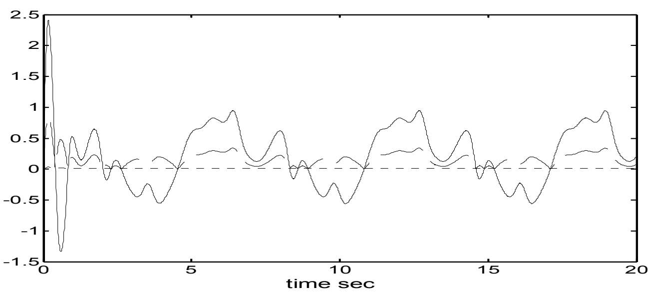

Fig. 4. $\|\hat{M}\|$ (dash), $\|\hat{N}\|$ (dot), output of NN(solid)

4. Decentralized Adaptive Neural Network Control of a Class of Large-Scale Nonlinear Systems with linear function interconnections

In the section, the above proposed scheme is extended to large-scale decentralized nonlinear systems, which the subsystems are composed of the class of the above-mentioned non-affine nonlinear functions. Two schemes are proposed, respectively. The first scheme designs a RBFN-based adaptive control scheme with the assumption which the interconnections between subsystems in entire system are bounded linearly by the norms of the tracking filtered error. In another scheme, the interconnection is assumed as stronger nonlinear function.

We consider the differential equations in the following form described, and assume the large-scale system is composed of the nonlinear subsystems:

$$\begin{cases} \dot{x}_{i1} = x_{i2} \\ \dot{x}_{i2} = x_{i3} \\ \quad \vdots \\ \dot{x}_{il_i} = f_i(x_{i1}, x_{i2}, \cdots, x_{ili}, u_i) + g_i(x_1, x_2, \cdots, x_n) \\ y_i = x_{i1} \end{cases} \tag{33}$$
$$i = 1, 2, \cdots n,$$

where $x_i \in R^{l_i}$ is the state vector, $x_i = [x_{i1}, x_{i2}, \cdots, x_{il_i}]^T$, $u_i \in R$ is the input and $y_i \in R$ is the output of the $i-th$ subsystem.

$f_i(x_i, u_i): R^{li+1} \to R$ is an unknown continuous function and implicit and smooth function with respect to control input u_i .

Assumption 5. $\partial f_i(x_i, u_i)/\partial u_i \neq 0$ for all $(x_i, u_i) \in \Omega_i \times R$.

$g_i(x_1, x_2, \cdots, x_n)$ is the interconnection term. In according to the distinctness of the interconnection term, two schemes are respectively designed in the following.

4.1 RBFN-based decentralized adaptive control for the class of large-scale nonlinear systems with linear function interconnections

Assumption 6. The interconnection effect is bounded by the following function:

$$\left| g_i(x_1, x_2, \cdots, x_n) \right| \le \sum_{j=1}^{n} \gamma_{ij} \left| \tau_j \right|, \tag{34}$$

where γ_{ij} are unknown coefficients, τ_j is a filtered tracking error to be defined shortly .

The control objective is: determine a control law, force the output, y_i , to follow a given desired output, x_{di} , with an acceptable accuracy, while all signals involved must be bounded.

Define the desired trajectory vector $x_{di} = [y_{di}, \dot{y}_{di}, \cdots, y_{di}^{l_i - 1}]^T$ and $X_{di} = \left[y_{di}, \dot{y}_{di}, \cdots, y_{di}^{(l_i)} \right]^T$, tracking error $e_i = x_i - x_{di} = [e_{i1}, e_{i2}, \cdots, e_{il_i}]^T$, thus, the filter tracking error can be written as

$$\tau_i = [\Lambda_i^T \quad 1]\mathbf{e}_i = k_{i,1}e_i + k_{i,2}\dot{e}_i + \cdots + k_{i,l_i-1}e_i^{(l_i-2)} + e_i^{(l_i-1)}, \tag{35}$$

where the coefficients are chosen such that the polynomial $k_{i,1} + k_{i,2}s + \cdots + k_{i,l_i-1}s^{(l_i-2)}$ $+s^{(l_i-1)}$ is Hurwitz.

Assumption 7. The desired signal $x_{di}(t)$ is bounded, so that $\|X_{di}\| \le \bar{X}_{di}$, where $\bar{X}_{di}$ is a known constant.

For an isolated subsystem, without interconnection function, by differentiating (35), the filtered tracking error can be rewritten as

$$\dot{\tau}_i = \dot{x}_{il_l} - x_{di}^{(l_i)} + [0 \quad \Lambda_i^T]e_i = f_i(x_i, u_i) + Y_{di} \tag{36}$$

with $Y_{di} = -x_{di}^{(l_i)} + [0 \quad \Lambda_i^T]e_i$.

Define a continuous function

$$\delta_i = -k_i\tau_i - Y_{di} \tag{37}$$

where k_i is a positive constant. With Assumption 5, we know $\partial f(x_i, u_i)/\partial u_i \ne 0$, thus, $\partial[f(x_i, u_i) - \delta_i]/\partial u_i \ne 0$. Considering the fact that $\partial\delta_i/\partial u_i = 0$, we invoke the implicit function theorem, there exists a continuous ideal control input u_i^* in a neighborhood of $(x_i, u_i) \in \Omega_i \times R$, such that $f(x_i, u_i^*) - \delta_i = 0$, i.e. $\delta_i = f_i(x_i, u_i^*)$ holds. $\delta_i = f_i(x_i, u_i^*)$ represents ideal control inverse.

Adding and subtracting δ_i to the right-hand side of $\dot{x}_{il_i} = f_i(x_i, u_i) + g_i$ of (33), one obtains

$$\dot{x}_{il_i} = f_i(x_i, u_i) + g_i - \delta_i - k_i\tau_i - Y_{di}, \tag{38}$$

and yields

$$\dot{\tau}_i = -k_i\tau_i + f_i(x_i, u_i) + g_i - \delta_i \; . \tag{39}$$

In the same the above-discussed manner as equations (9)-(10) , we can obtain the following equation:

$$\hat{\psi}_i = f_i(x_i, \hat{u}_i) \, . \tag{40}$$

Based on the above conditions, in order to control the system and make it be stable, we design the approximation pseudo-control input $\hat{\psi}_i$ as follows:

$$\hat{\psi}_i = -k_i\tau_i - Y_{di} + u_{ci} + v_{ri} \, , \tag{41}$$

where u_{ci} is output of a neural network controller, which adopts a RBFN, v_{ri} is robustifying control term designed in stability analysis.
Adding and subtracting $\hat{\psi}_i$ to the right-hand side of (39), with $\delta_i = -k_i\tau_i - Y_{di} = f_i(x_i, u_i^*)$, we have

$$\dot{\tau}_i = -k_i\tau_i + \tilde{\Delta}_i(x_i, u_i, u_i^*) - u_{ci} + \hat{\psi}_i - \delta_i - v_{ri} + g_i \, , \tag{42}$$

where $\tilde{\Delta}_i(x_i, u_i, u_i^*) = f_i(x_i, u_i) - f_i(x_i, u_i^*)$ is error between nonlinear function and its ideal control function, we can use the RBFN to approximate it.

4.1.1 Neural network-based approximation

Given a multi-input-single-output RBFN, let n_{1i} and m_{1i} be node number of input layer and hidden layer, respectively. The active function used in the RBFN is Gaussian function, $S_l(\mathbf{x}) = \exp[-0.5(\|z_i - \mu_{lk}\|^2)/\sigma_k^2]$, $l = 1, \cdots, n_{1i}$, $k = 1, \cdots, m_{1i}$ where $z_i \in R^{n_{1i}\times 1}$ is input vector of the RBFN, $\mu_i \in R^{n_{1i}\times m_{1i}}$ and $\sigma_i \in R^{m_{1i}\times 1}$ are the center matrix and the width vector.
Based on the approximation property of RBFN, $\tilde{\Delta}_i(x_i, u_i, u_i^*)$ can be written as

$$\tilde{\Delta}_i(x_i, u_i, u_i^*) = W_i^T S_i(z_i, \mu_i, \sigma_i) + \varepsilon_i(z_i) \, , \tag{43}$$

where $\varepsilon_i(z_i)$ is approximation error of RBFN, $W_i \in R^{m_{1i}\times 1}$.
Assumption 8. The approximation error $\varepsilon(x_{nn})$ is bounded by $|\varepsilon_i| \le \varepsilon_{Ni}$, with $\varepsilon_{Ni} > 0$ is an unknown constant.
The input of RBFN is chosen as $z_i = [x_i^T, \tau_i, \hat{\psi}_i]^T$. Moreover, output of RBFN is designed as

$$u_{ci} = \hat{W}_i^T S_i(z_i, \hat{\mu}_i, \hat{\sigma}_i). \tag{44}$$

Define $\hat{W}_i, \hat{\mu}_i, \hat{\sigma}_i$ as estimates of ideal W_i, μ_i, σ_i, which are given by the RBFN tuning algorithms.

Assumption 9. The ideal values of W_i, μ_i, σ_i satisfy

$$\|W_i\| \le W_{iM}, \quad \|\mu_i\|_F \le \mu_{iM}, \quad \|\sigma_i\| \le \sigma_{iM}, \tag{45}$$

where $W_{iM}, \mu_{iM}, \sigma_{iM}$ are positive constants. $\|\cdot\|_F$ and $\|\cdot\|$ denote Frobenius norm and 2-norm, respectively. Define their estimation errors as

$$\tilde{W}_i = W_i - \hat{W}_i, \quad \tilde{\mu}_i = \mu_i - \hat{\mu}_i, \quad \tilde{\sigma}_i = \sigma_i - \hat{\sigma}_i. \tag{46}$$

Using the notations: $Z_i = diag[W_i, \mu_i, \sigma_i], \tilde{Z}_i = diag[\tilde{W}_i, \tilde{\mu}_i, \tilde{\sigma}_i], \hat{Z}_i = diag[\hat{W}_i, \hat{\mu}_i, \hat{\sigma}_i]$ for convenience.

The Taylor series expansion for a given μ_i and σ_i is

$$S_i(z_i, \mu_i, \sigma_i) = S_i(z_i, \hat{\mu}_i, \hat{\sigma}_i) + \hat{S}'_{\mu i}\tilde{\mu}_i + \hat{S}'_{\sigma i}\tilde{\sigma}_i + O(\tilde{\mu}_i, \tilde{\sigma}_i)^2 \tag{47}$$

where $\hat{S}'_{\mu i} \triangleq \partial S_k(z_i, \hat{\mu}_i, \hat{\sigma}_i)/\partial \mu_i$, $\hat{S}'_{\sigma i} \triangleq \partial S_k(z_i, \hat{\mu}_i, \hat{\sigma}_i)/\partial \sigma_i$ evaluated at $\mu_i = \hat{\mu}_i$, $\sigma_i = \hat{\sigma}_i$, $O(\tilde{\mu}_i, \tilde{\sigma}_i)^2$ denotes the terms of order two. We use notations: $\hat{S}_i := S_i(z_i, \hat{\mu}_i, \hat{\sigma}_i)$, $\tilde{S}_i := S_i(z_i, \tilde{\mu}_i, \tilde{\sigma}_i)$, $S_i := S_i(z_i, \mu_i, \sigma_i)$.

Following the procedure in Appendix B, it can be shown that the following operation. The function approximation error can be written as

$$W_i^T S_i - \hat{W}_i^T \hat{S}_i = \tilde{W}_i^T(\hat{S}_i - \hat{S}'_{\mu i}\hat{\mu}_i - \hat{S}'_{\sigma i}\hat{\sigma}_i) + \hat{W}_i^T(\hat{S}'_{\mu i}\tilde{\mu}_i + \hat{S}'_{\sigma i}\tilde{\sigma}_i) + \omega_i(t), \tag{48}$$

The disturbance term $\omega_i(t)$ is given by

$$\omega_i(t) = W_i^T(S_i - \hat{S}_i) + W_i^T(\hat{S}'_{\mu i}\hat{\mu}_i + \hat{S}'_{\sigma i}\hat{\sigma}_i) - \hat{W}_i^T(\hat{S}'_{\mu i}\mu_i + \hat{S}'_{\sigma i}\sigma_i) \tag{49}$$

Then, the upper bound of $\omega_i(t)$ can be written as

$$|\omega_i(t)| \le \|W_i\|(\|\hat{S}'_{\mu i}\hat{\mu}_i\|_F + \|\hat{S}'_{\sigma i}\hat{\sigma}_i\|_F) + \|\hat{W}_i^T\hat{S}'_{\mu i}\|_F\|\mu_i\|_F + \|\hat{W}_i^T\hat{S}'_{\sigma i}\|_F\|\sigma_i\| + 2\|W_i\|_1 \le \rho_{\omega i}\vartheta_{\omega i} \tag{50}$$

where $\rho_{\omega i} = \max(\|W_i\|, \|\mu_i\|_F, \|\sigma_i\|, 2\|W_i\|_1)$, $\vartheta_{\omega i} = \|\hat{S}'_{\mu i}\hat{\mu}_i\|_F + \|\hat{S}'_{\sigma i}\hat{\sigma}_i\|_F + \|\hat{W}_i^T\hat{S}'_{\mu i}\|_F + \|\hat{W}_i^T\hat{S}'_{\sigma i}\|_F + 1$, with $\|\cdot\|_1$ 1 norm. Notice that $\rho_{\omega i}$ is an unknown coefficient, whereas $\vartheta_{\omega i}$ is a known function.

4.1.2 Controller design and stability analysis

Substituting (43) and (44) into (42), we have

$$\dot{\tau}_i = -k_i\tau_i + W_i^T S_i - \hat{W}_i^T\hat{S}_i + \hat{\psi}_i - \delta_i - v_{ri} + g_i + \varepsilon_i(z_i), \tag{51}$$

using (48), the above equation can become

$$\begin{aligned}\dot{\tau}_i = -k_i\tau_i &+ \tilde{W}_i^T(\hat{S}_i - \hat{S}'_{\mu i}\hat{\mu}_i - \hat{S}'_{\sigma i}\hat{\sigma}_i) + \hat{W}_i^T(\hat{S}'_{\mu i}\tilde{\mu}_i + \hat{S}'_{\sigma i}\tilde{\sigma}_i)\\ &+ \hat{\psi}_i - \delta_i - v_{ri} + g_i + \varepsilon_i(z_i) + \omega_i(t).\end{aligned} \tag{52}$$

Theorem 2. Consider the nonlinear subsystems represented by Eq. (33) and let assumptions hold. If choose the pseudo-control input $\hat{\psi}_i$ as Eq.(41), and use the following adaptation laws and robust control law

$$\dot{\hat{W}}_i = F_i\left[(\hat{S}_i - \hat{S}'_{\mu i}\hat{\mu}_i - \hat{S}'_{\sigma i}\hat{\sigma}_i)\tau_i - \gamma_{Wi}\hat{W}_i|\tau_i|\right], \tag{53}$$

$$\dot{\hat{\mu}}_i = G_i\left[\hat{S}'^{\,T}_{\mu i}\hat{W}_i\tau_i - \gamma_{Wi}\hat{\mu}_i|\tau_i|\right], \tag{54}$$

$$\dot{\hat{\sigma}}_i = H_i\left[\hat{S}'^{\,T}_{\sigma i}\hat{W}_i\tau_i - \gamma_{Wi}\hat{\sigma}_i|\tau_i|\right], \tag{55}$$

$$\dot{\hat{\phi}}_i = \gamma_{\phi i}\left[\tau_i\vartheta^*_{\omega i}\tanh\left(\frac{\tau_i\vartheta^*_{\omega i}}{\alpha_i}\right) - \lambda_{\phi i}\hat{\phi}_i|\tau_i|\right], \tag{56}$$

$$\dot{\hat{d}}_i = \gamma_{di}(\tau_i^2 - \lambda_{di}\hat{d}_i|\tau_i|), \tag{57}$$

$$v_{ri} = \hat{\phi}_i\vartheta^*_{\omega i}\tanh\left(\frac{\tau_i\vartheta^*_{\omega i}}{\alpha_i}\right) + \hat{d}_i\tau_i, \tag{58}$$

where $\vartheta_{\omega i}^{*} = \vartheta_{\omega i} + 1$, $F_i = F_i^T > 0, G_i = G_i^T > 0, H_i = H_i^T > 0$ are any constant matrices, $\gamma_{Wi}, \gamma_{\phi i}, \gamma_{di}, \lambda_{\phi i}, \lambda_{di}$ and α_i are positive design parameters, $\hat{\phi}_i$ is the estimated value of the uncertain disturbance term $\phi_i = \max(\rho_{\omega i}, \varepsilon_{Ni})$, defining $\tilde{\phi}_i = \phi_i - \hat{\phi}_i$ with $\tilde{\phi}_i$ error, $d_i > 0$ is used to estimate unknown positive number to shield interconnection effect, $\hat{d}_i$ is its estimated value, with $\tilde{d}_i = d_i - \hat{d}_i$ estimated error, then, guarantee that all signals in the system are bounded and the tracking error e_i will converge to a neighborhood of the origin.

Proof. Consider the following positive define Lyapunov function candidate as

$$L_i = \frac{1}{2}\tau_i^2 + \frac{1}{2}\left[tr(\tilde{W}_i^T F_i^{-1} \tilde{W}_i) + tr(\tilde{\mu}_i^T G_i^{-1} \tilde{\mu}_i) + tr(\tilde{\sigma}_i^T H_i^{-1} \tilde{\sigma}_i) + \gamma_{\phi i}^{-1} \tilde{\phi}_i^2 + \gamma_{di}^{-1} \tilde{d}_i^2 \right] \quad (59)$$

The time derivative of the above equation is given by

$$\dot{L}_i = \tau_i \dot{\tau}_i + tr(\tilde{W}_i^T F_i^{-1} \dot{\tilde{W}}_i) + tr(\tilde{\mu}_i^T G_i^{-1} \dot{\tilde{\mu}}_i) + tr(\tilde{\sigma}_i^T H_i^{-1} \dot{\tilde{\sigma}}_i) + \gamma_{\phi i}^{-1} \tilde{\phi}_i \dot{\tilde{\phi}}_i + \gamma_{di}^{-1} \tilde{d}_i \dot{\tilde{d}}_i \quad (60)$$

Applying(52) to (60), we have

$$\begin{aligned} \dot{L}_i = \tau_i &\begin{bmatrix} -k_i \tau_i + \tilde{W}_i^T (\hat{S}_i - \hat{S}'_{\mu i} \hat{\mu}_i - \hat{S}'_{\sigma i} \hat{\sigma}_i) + \hat{W}_i^T (\hat{S}'_{\mu i} \tilde{\mu}_i + \hat{S}'_{\sigma i} \tilde{\sigma}_i) \\ + \hat{\psi}_i - \delta_i - v_{ri} + g_i + \varepsilon_i + \omega_i \end{bmatrix} \\ &+ tr(\tilde{W}_i^T F_i^{-1} \dot{\tilde{W}}_i) + tr(\tilde{\mu}_i^T G_i^{-1} \dot{\tilde{\mu}}_i) + tr(\tilde{\sigma}_i^T H_i^{-1} \dot{\tilde{\sigma}}_i) + \gamma_{\phi i}^{-1} \tilde{\phi}_i \dot{\tilde{\phi}}_i + \gamma_{di}^{-1} \tilde{d}_i \dot{\tilde{d}}_i \end{aligned} \quad (61)$$

Substituting the adaptive laws (53), (54) and (55) into (61), and $(\dot{\tilde{\cdot}}) = -(\dot{\hat{\cdot}})$,yields

$$\begin{aligned} \dot{L}_i &= \tau_i \left[-k_i \tau_i + \hat{\psi}_i - \delta_i - v_{ri} + g_i + \varepsilon_i + \omega_i \right] + \gamma_{Wi} \left| \tau_i \right| tr(\tilde{Z}_i^T \hat{Z}_i) + \gamma_{\phi i}^{-1} \tilde{\phi}_i \dot{\tilde{\phi}}_i + \gamma_{di}^{-1} \tilde{d}_i \dot{\tilde{d}}_i \\ &\le -k_i \tau_i^2 + \tau_i (\hat{\psi}_i - \delta_i) - v_{ri} \tau_i + \tau_i g_i + \left| \tau_i \right| (\rho_{\omega i} \vartheta_{\omega i} + \varepsilon_{Ni}) \\ &+ \gamma_{Wi} \left| \tau_i \right| tr(\tilde{Z}_i^T \hat{Z}_i) + \gamma_{\phi i}^{-1} \tilde{\phi}_i \dot{\tilde{\phi}}_i + \gamma_{di}^{-1} \tilde{d}_i \dot{\tilde{d}}_i \\ &\le -k_i \tau_i^2 + \tau_i (\hat{\psi}_i - \delta_i) - v_{ri} \tau_i + \tau_i g_i + \left| \tau_i \right| \phi_i \vartheta_{\omega i}^{*} \\ &+ \gamma_{Wi} \left| \tau_i \right| tr(\tilde{Z}_i^T \hat{Z}_i) + \gamma_{\phi i}^{-1} \tilde{\phi}_i \dot{\tilde{\phi}}_i + \gamma_{di}^{-1} \tilde{d}_i \dot{\tilde{d}}_i \end{aligned} \quad (62)$$