15th Triennial World Congress, Barcelona, Spain

DIRECT MNN CONTROL OF A CLASS OF DISCRETE-TIME NON-AFFINE NONLINEAR SYSTEMS

S. S. Ge, [1] **J. Zhang and T. H. Lee**

Department of Electrical and Computer Engineering
National University of Singapore
Singapore 117576

Abstract: In this paper, direct adaptive neural network control is presented for a class of discrete-time SISO non-affine nonlinear systems. Based on the input-output model, multi-layer neural networks are used to emulate the implicit desired feedback control. For the multi-layer neural network control, projection algorithms are used to guarantee the boundness of the neural network weights. The stability of the closed-loop system is proved by using Lyapunov theorem. *Copyright © 2002 IFAC*

Keywords: Non-affine nonlinear systems, implicit function theorem, multi-layer neural networks, projection algorithm, discrete-time systems.

1. INTRODUCTION

While fundamental physical models are almost always developed in continuous-time, computer based control systems function in discrete-time. In addition, the input output data available for model identification is generally only available at discrete time instants. Furthermore, over the past few years, adaptive control for continuous nonlinear system has been studied extensively. These methods cannot be directly applied to discrete-time systems due to some technical difficulties, such as lack of applicability of Lyapunov techniques (Kanellakopoulos, 1994) and loss of linear parameterizability during linearization. These observations motivate us to develop adaptive control scheme for discrete-time nonlinear systems.

Multi-layer neural network is a static feedforward network that consists of a number of layers. It has an important character, that MNNs with one or more hidden layers are capable of approximating any continuous nonlinear function. This make it one of the most widely used neural networks in system modeling and control. In this paper, we use MNN to emulate the implicit desired feedback controller. Because of the residual term of multi-layer neural network approximation, projection algorithms (Goodwin and Mayne, 1987)(Sastry and Bodson, 1989) are used in this paper to guarantee the MNN weights bounded in compact sets.

For nonlinear discrete-time systems, there has been many discussions. In (Chen and Khalil, 1995), a specific class of nonlinear affine systems is investigated. The developed method will lose its effect for non-affine nonlinear systems. In (Goh and Lee, 1994), direct control of a general nonlinear dynamical system is discussed based on implicit function theory. The neural network control method is firstly discussed for first order discrete-time nonlinear system, and then the control scheme is generalized to high order discrete-time nonlinear system without rigorous proof. In (Cabrera and Narendra, 1999), discrete NARMAX (Nonlinear Auto Regressive Moving Average with eXogenous inputs) non-affine systems based on input-output models are discussed without rigorous proof. In this paper, MNN neural net-

[1] Corresponding author. Tel. (65) 8746821; Fax. (65) 7791103; E-mail: elegesz@nus.edu.sg

works are used to construct direct neural network controller for a class of discrete-time non-affine nonlinear systems. The stability analysis method and the weight update laws are different from the literatures listed above.

The main contributions of this paper are that (i) provide an effective neural network control method for non-affine nonlinear discrete-time systems which feedback linearization method is of no use; (ii) propose a different kind of MNN weight update law for discrete-time systems; (iii) propose a modified discrete-time projection algorithm compare to continuous-time projection algorithm used in (Gong and Yao, 2001); (iv) using MNNs to emulate the implicit desired feedback control (IDFC) of non-affine discrete-time systems, which is not only a challenging topic but also of academic interest.

2. PROBLEM

Consider one of the most popular nonlinear representation NARMAX model known as τ-step ahead observer equation

$$\begin{aligned} y(k+\tau) &= f(y(k),\ldots,y(k-n+1),u(k),\ldots,\\ &\quad u(k-n+1),d(k+\tau-1),\ldots,d(k))\\ &= f(\bar{y}_k,u(k),\bar{u}_{k-1},\bar{d}_{k+\tau-1}) \qquad (1)\end{aligned}$$

where $\bar{y}_k = [y(k),\ldots,y(k-n+1)]^T$, $\bar{u}_{k-1} = [u(k-1),\ldots,u(k-n+1)]^T$ and $\bar{d}_{k+\tau-1} = [d(k+\tau-1),\ldots,d(k)]^T$. This model relates an input sequence $\{u(k)\}$ to an output sequence $\{y(k)\}$ by nonlinear difference equation. $\{d(k)\}$ represents a "modeling error" in this relationship.

Assumption 2.1. The unknown nonlinear function $f(\cdot)$ is continuous and differentiable.

Assumption 2.2. The disturbance $d(k)$ is bounded, $|d(k)| \le d$, where d is a little unknown constant and the partial derivative $|\frac{\partial f}{\partial d(k)}| \le g_2$, where g_2 is a positive constant.

Assumption 2.3. Assume that partial derivative $g_1 \ge |\frac{\partial f}{\partial u}| > \epsilon > 0$, where both ϵ and g_1 are positive constants.

This assumption states that $\frac{\partial f}{\partial u}$ is either positive or negative. From now onwards, without lose of generality, we assume that $\frac{\partial f}{\partial u} > 0$.

Assume that $y_m(k+\tau)$ is the reference output at time instant $k+\tau$. Under Assumption 2.3, adding and subtracting $y_m(k+\tau)$ to the right side of equation (1) and using Mean Value Theorem

$$\begin{aligned} y(k+\tau) &= y_m(k+\tau) + f(\bar{y}_k,u(k),\bar{u}_{k-1},\bar{d}_{k+\tau-1})\\ &\quad -y_m(k+\tau)\\ &= y_m(k+\tau) + f(\bar{y}_k,u(k),\bar{u}_{k-1},0)\\ &\quad +\delta_f^T\bar{d}_{k+\tau-1} - y_m(k+\tau)\\ &= y_m(k+\tau) + f(\bar{y}_k,u(k),\bar{u}_{k-1},0)\\ &\quad -y_m(k+\tau) + \delta_{d_k} \qquad (2)\end{aligned}$$

where

$$\begin{aligned}\delta_f &= \Big[\frac{\partial f}{\partial d(k+\tau-1)}|_{d(k+\tau-1)=d_{\xi_{k+\tau-1}}},\cdots,\\ &\quad \frac{\partial f}{\partial d(k)}|_{d(k)=d_{\xi_k}}\Big]^T\\ \bar{d}_\xi &= [d_{\xi_{k+\tau-1}},\cdots,d_{\xi_k}]^T, \quad \delta_{d_k} = \delta_f^T\bar{d}_{k+\tau-1}\end{aligned}$$

and $\bar{d}_\xi \in L(0,\bar{d}_{k+\tau-1})$.

Consider Assumption 2.2, we know that the disturbance item δ_{d_k} in equation (2) is bounded by

$$\begin{aligned}\delta_{d_k} &= \delta_f^T\bar{d}_{k+\tau-1} \le g_2 d + g_2 d + \ldots + g_2 d\\ &= (\tau-1)g_2 d \qquad (3)\end{aligned}$$

Define the tracking error as $e(k) = y(k) - y_m(k)$, then the tracking error dynamic equation is given

$$\begin{aligned} e(k+\tau) &= -y_m(k+\tau)\\ &\quad + f(\bar{y}_k,u(k),\bar{u}_{k-1},0) + \delta_{d_k} \qquad (4)\end{aligned}$$

In the ideal case, there is no disturbance ($\delta_{d_k} = 0$), if the control input $u^*(k)$ satisfying

$$f(\bar{y}_k,u^*(k),\bar{u}_{k-1},0) - y_m(k+\tau) = 0 \qquad (5)$$

system's output tracking error will converge to 0.

Lemma 2.1. According to Assumptions 2.1 and 2.3 if $|\frac{\partial f}{\partial u(k)}| > \epsilon > 0$, then there exists a unique and continuous function $u^*(k) = \alpha^c(\bar{y}_k,\bar{u}_{k-1},y_m(k+\tau))$ such that equation (5) holds.

Proof. See (Ge *et al.*, 2001).

3. PRELIMINARIES

3.1 *Multi-layer Neural Network*

According to the neural network theory, there exists an integer l (the number of hidden neurons) and ideal constant weight W^* and V^*, such that

$$u^*(k) = u^*(z) = {W^*}^T S({V^*}^T\bar{z}) + \varepsilon_u(z) \qquad (6)$$

where $\bar{z} = [z,1]^T$ and z is the input vector. Following assumption is made for this function approximation.

Assumption 3.1. On the compact set Ω_z, the ideal neural network weights W^*, V^* and the NN approximation error are bounded by

$$\|W^*\| \le w_m, \quad \|V^*\|_F \le v_m, \quad |\varepsilon_u(z)| \le \varepsilon_l \qquad (7)$$

with w_m, v_m and ε_l being positive constants.

Choosing $s(x) = \frac{1}{1+e^{-x}}$ as the activation function. Its derivative is

$$s'(x) = \frac{d[s(x)]}{dx} = \frac{e^{-x}}{(1+e^{-x})^2}$$

It is easy to check that

$$0 \le s'(x) \le 0.25, \quad |xs'(x)| \le 0.2239, \quad x \in R \qquad (8)$$

$$\|\hat{S}'\|_F \le \sum_{i=1}^{l} s'(\hat{v}_i^T\bar{z}) \le 0.25l$$

$$\|\hat{S}'\hat{V}^T\bar{z}\| \le \sum_{i=1}^{l} |\hat{v}_i^T\bar{z}s'(\hat{v}_i^T\bar{z})| \le 0.2239l \qquad (9)$$

here $\hat{S}' = diag\{s'(\hat{v}_1^T\bar{z}),\ldots,s'(\hat{v}_i^T\bar{z}),\ldots,s'(\hat{v}_l^T\bar{z})\}$ is a diagonal matrix. Using Taylor series expand

$S(V^{*T}\bar{z})$ about $\hat{V}^T\bar{z}$, noting abbreviation $\hat{S} = S(\hat{V}^T\bar{z})$ and $\tilde{V} = \hat{V} - V^*$, thus

$$S(V^{*T}\bar{z}) = \hat{S} - \hat{S}'\tilde{V}^T\bar{z} + O(\tilde{V}^T\bar{z})^2 \quad (10)$$

Using inequality (9), the high order term $O(\tilde{V}^T\bar{z})^2$ is bounded by

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

Considering (9), $\|V^*\|_F \le v_m$ and the fact that $\|S(V^{*T}\bar{z}) - S(\hat{V}^T\bar{z})\| \le l$, we have

$$\|O(\tilde{V}^T\bar{z})^2\| \le 1.2239l + 0.25v_m l\|\bar{z}\| \quad (11)$$

3.2 *Projection Algorithm*

It is assumed that each element of W^* and V^* is bounded, i.e., $\rho_{w_i,\min} \le w_i \le \rho_{w_i,\max}$ for $i = 1,\dots,l$ and $\rho_{v_{ij},\min} \le v_{ij} \le \rho_{v_{ij},\max}$ for $i = 1,\dots,n, j = 1,\dots,l$, where the lower and upper bounds $\rho_{w,\min}$, $\rho_{w,\max}$, $\rho_{v,\min}$, $\rho_{v,\max}$ maybe unknown with n the input dimension of neural networks and l the neuron numbers. It is natural to require that the estimates of the weights be within the corresponding bounds. However, due to the fact that these bounds may not be known a prior, certain fictitious bounds have to be used.

In this paper, the following projection mapping is used. Let $\hat{\rho}_{\Theta_{ij},\min}$ and $\hat{\rho}_{\Theta_{ij},\max}$ be the fictitious lower and upper bound for Θ_{ij}, where Θ could be any of the unknown weight vector or matrix. Based on these fictitious lower and upper bounds, similar to the projection algorithm in continuous time(Goodwin and Mayne, 1987), a discontinuous projection mapping $\mathrm{Proj}(*)$ can be defined as $\mathrm{Proj}_{\hat{\Theta}}(*) = \left[\mathrm{Proj}_{\hat{\Theta}}(*_{ij})\right]$ with its ijth element being

$$\mathrm{Proj}_{\hat{\Theta}}(*_{ij}) = \begin{cases} -*_{ij} & \text{if} \begin{cases} \hat{\Theta}_{ij} = \hat{\rho}_{\Theta_{ij},\max} \\ \quad \text{and } *_{ij} < 0 \\ \hat{\Theta}_{ij} = \hat{\rho}_{\Theta_{ij},\min} \\ \quad \text{and } *_{ij} > 0 \end{cases} \\ *_{ij} & \quad \text{otherwise} \end{cases} \quad (12)$$

where $*$ denotes a vector or a matrix, then $*_{ij}$ denotes its element.

Choosing weight update law as

$$\hat{\Theta}(k+\tau) = \hat{\Theta}(k) - \mathrm{Proj}_{\hat{\Theta}}(\Gamma\eta) \quad (13)$$

where $\Gamma = \Gamma^T > 0$ is diagonal positive-definite adaptation matrix, and η is adaptation function. For simplicity, assume $\Gamma = \lambda I$ with $\lambda > 0$.

Lemma 3.1. For the projection algorithm (12) and parameter adaptation laws (13) used in this paper, the following properties hold:

(i) Parameter estimates are within the known prescribed range, $\hat{\rho}_{\Theta_{ij},\min} \le \hat{\Theta}_{ij} \le \hat{\rho}_{\Theta_{ij},\max}$.

(ii) In addition, noting $\tilde{\Theta} = \hat{\Theta} - \Theta$, then

$$\tilde{\Theta}^T(\Gamma^{-1}\mathrm{Proj}_{\hat{\Theta}}(\Gamma\eta) - \eta) \ge 0, \Theta \in R^n$$
$$\mathrm{tr}\{\tilde{\Theta}^T(\Gamma^{-1}\mathrm{Proj}_{\hat{\Theta}}(\Gamma\eta) - \eta)\} \ge 0, \Theta \in R^{m\times n}$$

Proof. According to (12) and (13), it is obvious that the first property always holds. Now we prove the second property.

If Θ is a vector, consider the diagonal positive-definite adaptation matrix Γ, noticing that the operation of projection operator $\mathrm{Proj}_{\hat{\Theta}}(*_{ij})$ is to change the sign of $*_{ij}$, thus

$$\begin{aligned}\tilde{\Theta}^T(\Gamma^{-1}\mathrm{Proj}_{\hat{\Theta}}(\Gamma\eta) - \eta) &= \tilde{\Theta}^T(\Gamma^{-1}\Gamma\mathrm{Proj}_{\hat{\Theta}}(\eta) - \eta) \\ &= \tilde{\Theta}^T(\mathrm{Proj}_{\hat{\Theta}}(\eta) - \eta) \\ &= \sum_{i=1}^{l}\tilde{\Theta}_i(\mathrm{Proj}_{\hat{\Theta}}(\eta_i) - \eta_i)\end{aligned}$$

Then considering

$$\begin{aligned}&\tilde{\Theta}_i(\mathrm{Proj}_{\hat{\Theta}}(\eta_i) - \eta_i) \\ &= (\hat{\Theta}_i - \Theta_i)(\mathrm{Proj}_{\hat{\Theta}}(\eta_i) - \eta_i) \\ &= \begin{cases} (a) \quad (\hat{\Theta}_i - \Theta_i)(-\eta_i - \eta_i) > 0 \\ \quad \text{if} \begin{cases} \hat{\Theta}_i = \hat{\rho}_{\Theta_i,\max} \ \ (\text{i.e. } \hat{\Theta}_i > \Theta_i) \\ \quad \text{and } \eta_i < 0 \\ \hat{\Theta}_i = \hat{\rho}_{\Theta_i,\min} \ \ (\text{i.e. } \hat{\Theta}_i < \Theta_i) \\ \quad \text{and } \eta_i > 0 \end{cases} \\ (b) \quad (\hat{\Theta}_i - \Theta_i)(\eta_i - \eta_i) = 0 \\ \quad \text{otherwise} \end{cases}\end{aligned}$$

we have $\tilde{\Theta}^T(\Gamma^{-1}\mathrm{Proj}_{\hat{\Theta}}(\Gamma\eta) - \eta) \ge 0$ holds.

If Θ is a matrix, following the same procedure

$$\mathrm{tr}\{\tilde{\Theta}^T(\Gamma^{-1}\mathrm{Proj}_{\hat{\Theta}}(\Gamma\eta) - \eta)\} \ge 0$$

Its proof is omitted here. **Q.E.D.**

4. CONTROLLER DESIGN

Assumption 4.1. Considering the projection algorithms we used, on the compact set Ω_z, the estimate neural network weights $\hat{W}$, $\hat{V}$ and the weight approximation error $\tilde{W}$, $\tilde{V}$ are bounded by

$$\|\hat{W}\| \le \hat{w}_m, \quad \|\hat{V}\|_F \le \hat{v}_m,$$
$$\|\tilde{W}\| \le \tilde{w}_m, \quad \|\tilde{V}\|_F \le \tilde{v}_m \quad (14)$$

where $\tilde{W} = \hat{W} - W^*$, $\tilde{V} = \hat{V} - V^*$ and $\tilde{w}_m$, $\tilde{v}_m$, $\hat{w}_m$, $\hat{v}_m$ are positive constants.

Choosing auxiliary variables as follows

$$\eta_w = \hat{S}(k)e(k+\tau) \quad (15)$$
$$\eta_v = (\bar{z}\hat{W}^T(k)\hat{S}'(k))e(k+\tau) \quad (16)$$

where $\bar{z} = [z^T, 1]^T$ with $z = [\bar{y}_k, \bar{u}_{k-1}, y_m(k+\tau)]^T$, $\bar{y}_k$, $\bar{u}_{k-1}$ and $y_m(k+\tau)$ are defined in Section 2. $\hat{S}'(k) = diag[\hat{s}'_1(k), \dots, \hat{s}'_l(k)]$ is a diagonal matrix and $\hat{s}'_i(k) = s'(\hat{v}_i^T\bar{z}(k))$.

Define the multi-layer neural networks update law

$$\hat{W}(k+\tau) = \hat{W}(k) - \mathrm{Proj}_{\hat{W}}[\Gamma_w\eta_w] \quad (17)$$
$$\hat{V}(k+\tau) = \hat{V}(k) - \mathrm{Proj}_{\hat{V}}[\Gamma_v\eta_v] \quad (18)$$

Subtract W^* and V^* to the both sides of the equation (17) and (18), we obtain

$$\tilde{W}(k+\tau) = \tilde{W}(k) - \mathrm{Proj}_{\hat{W}}(\Gamma_w \eta_w) \quad (19)$$
$$= \tilde{W}(k) - \mathrm{Proj}_{\hat{W}}[\Gamma_w \hat{S}(k) e(k+\tau)]$$
$$\tilde{V}(k+\tau) = \tilde{V}(k) - \mathrm{Proj}_{\hat{V}}(\Gamma_v \eta_v) \quad (20)$$
$$= \tilde{V}(k) - \mathrm{Proj}_{\hat{V}}[\Gamma_v(\bar{z}\hat{W}^T(k)\hat{S}'(k))e(k+\tau)]$$

where $\Gamma_w = \Gamma_w^T = \lambda_w I$ and $\Gamma_v = \Gamma_v^T = \lambda_v I$.

Lemma 4.1. Consider Lemma 3.1, we have the following inequalities

$$\tilde{W}(\Gamma_w^{-1}\mathrm{Proj}_{\hat{W}}(\Gamma_w \eta_w) - \eta_w) \geq 0 \quad (21)$$
$$\mathrm{tr}\{\tilde{V}(\Gamma_v^{-1}\mathrm{Proj}_{\hat{V}}(\Gamma_v \eta_v) - \eta_v)\} \geq 0 \quad (22)$$

Furthermore we have

$$\mathrm{Proj}_{\hat{W}}^T \Gamma_w^{-1} \mathrm{Proj}_{\hat{W}} - \eta_w^T \Gamma_w^T \eta_w = 0 \quad (23)$$
$$\mathrm{tr}\{\mathrm{Proj}_{\hat{V}}^T \Gamma_v^{-1} \mathrm{Proj}_{\hat{V}}\} - \mathrm{tr}\{\eta_v^T \Gamma_v^T \eta_v\} = 0 \quad (24)$$

where $\mathrm{Proj}_{\hat{W}} = \mathrm{Proj}_{\hat{W}}(\Gamma_w \eta_w)$ and $\mathrm{Proj}_{\hat{V}} = \mathrm{Proj}_{\hat{V}}(\Gamma_v \eta_v)$.

Proof. It is obvious that following Lemma 3.1, inequalities (21) and (22) hold.

Considering equation (23), because $\Gamma_w = \lambda_w I$, we have

$$\mathrm{Proj}_{\hat{W}}^T \Gamma_w^{-1} \mathrm{Proj}_{\hat{W}} - \eta_w^T \Gamma_w^T \eta_w$$
$$= \mathrm{Proj}_{\hat{W}}^T(\eta_w) \Gamma_w^T \Gamma_w^{-1} \Gamma_w \mathrm{Proj}_{\hat{W}}(\eta_w) - \eta_w^T \Gamma_w^T \eta_w$$
$$= \lambda_w [\mathrm{Proj}_{\hat{W}}^T(\eta_w)\mathrm{Proj}_{\hat{W}}(\eta_w) - \eta_w^T \eta_w] = 0$$

then equation (23) holds.

Considering equation (24), because $\Gamma_v = \lambda_v I$, we have

$$\mathrm{tr}\{\mathrm{Proj}_{\hat{V}}^T \Gamma_v^{-1} \mathrm{Proj}_{\hat{V}}\} - \mathrm{tr}\{\eta_v^T \Gamma_v^T \eta_v\}$$
$$= \mathrm{tr}\{\mathrm{Proj}_{\hat{V}}^T(\eta_v) \Gamma_v^T \Gamma_v^{-1} \Gamma_v \mathrm{Proj}_{\hat{V}}(\eta_v)\}$$
$$- \mathrm{tr}\{\eta_v^T \Gamma_v^T \eta_v\}$$
$$= \lambda_v \mathrm{tr}\{\mathrm{Proj}_{\hat{V}}^T(\eta_v)\mathrm{Proj}_{\hat{V}}(\eta_v) - \eta_v^T \eta_v\}$$
$$= 0$$

then equation (24) holds. **Q.E.D.**

Choose the practical control input as

$$u(k) = \hat{W}^T(k) S(\hat{V}^T(k)\bar{z}) \quad (25)$$

where $\bar{z} = [z^T, 1]^T$ with $z = [\bar{y}_k, \bar{u}_{k-1}, y_m(k+\tau)]^T$, $\bar{y}_k$, $\bar{u}_{k-1}$ and $y_m(k+\tau)$ are defined in Section 2. Noticing equation (6), then we have

$$u(k) - u^*(k)$$
$$= \hat{W}^T(k) S(\hat{V}^T(k)\bar{z})$$
$$- W^{*T}(k) S(V^{*T}(k)\bar{z}) - \varepsilon_u(z)$$
$$= \tilde{W}^T(k) S(\hat{V}^T(k)\bar{z}) + W^{*T}(k)[S(\hat{V}^T(k)\bar{z})$$
$$- S(V^{*T}(k)\bar{z})] - \varepsilon_u(z)$$
$$= \tilde{W}^T \hat{S} + W^{*T}(\hat{S} - S^*) - \varepsilon_u(z)$$

Substitute $u(k)$ into the error equation (4), using Mean Value Theorem, noticing equation (5), then we have

$$e(k+\tau)$$
$$= -y_m(k+\tau) + \delta_{d_k} + f(\bar{y}_k, u^*(k), \bar{u}_{k-1}, 0)$$
$$+ \frac{\partial f}{\partial u}|_{u=\xi}(\tilde{W}^T \hat{S} + W^{*T}(\hat{S} - S^*) - \varepsilon_u(z))$$
$$= [\tilde{W}^T \hat{S} + W^{*T}(\hat{S} - S^*) - \varepsilon_u(z)] f_u + \delta_{d_k} \quad (26)$$

where

$$f_u = \frac{\partial f}{\partial u}|_{u=\xi} \quad \xi \in [u^*(k), u(k)]$$

Theorem 4.1. For the non-affine discrete-time system (1), neural network controller (25) and neural network weight update laws (17) and (18). There exist compact sets Ω_y, Ω_w, Ω_v and positive constants l^*, λ_w^* and λ_v^* such that if

(i) the initial parameter set $\Omega_{y_0} \in \Omega_y$, $\Omega_{w_0} \in \Omega_w$, $\Omega_{v_0} \in \Omega_v$;
(ii) the neural number $l > l^*$, adaptive gain $\lambda_w < \lambda_w^*$, with λ_w^* being the largest eigenvalue of Γ_w, $\lambda_v < \lambda_v^*$, with λ_v^* being the largest eigenvalue of Γ_v;
(iii) the initial future output $y(k_0), \ldots, y(k_0+\tau-1)$ are kept in the compact set Ω_y, initial input sequence $u(k_0)$ are kept in the compact set Ω_u;

then the output of system (1) will track the desired trajectory and the tracking error is bounded. The closed-loop system is stable and all the signals are bounded.

Proof. Choose the Lyapunov function as follows

$$J(k) = \frac{1}{g_1}\sum_{j=0}^{\tau-1} e^2(k+j)$$
$$+ \sum_{j=0}^{\tau-1} \tilde{W}^T(k+j)\Gamma_w^{-1}\tilde{W}(k+j)$$
$$+ \sum_{j=0}^{\tau-1} tr\{\tilde{V}^T(k+\tau)\Gamma_v^{-1}\tilde{V}(k+\tau)\} \quad (27)$$

The first difference of (27) is given

$$\triangle J(k) = J(k+1) - J(k)$$
$$= \frac{1}{g_1}[e^2(k+\tau) - e^2(k)]$$
$$+ \tilde{W}^T(k+\tau)\Gamma_w^{-1}\tilde{W}(k+\tau)$$
$$- \tilde{W}^T(k)\Gamma_w^{-1}\tilde{W}(k)$$
$$+ tr\{\tilde{V}^T(k+\tau)\Gamma_v^{-1}\tilde{V}(k+\tau)$$
$$- \tilde{V}^T(k)\Gamma_v^{-1}\tilde{V}(k)\}$$

Considering the neural network weight update laws (19) and (20), we have

$$\triangle J(k) = \frac{1}{g_1}[e^2(k+\tau) - e^2(k)]$$
$$- \tilde{W}^T(k)\Gamma_w^{-1}\mathrm{Proj}_{\hat{W}} - \mathrm{Proj}_{\hat{W}}^T \Gamma_w^{-1}\tilde{W}(k)$$
$$+ \mathrm{Proj}_{\hat{W}}^T \Gamma_w^{-1} \mathrm{Proj}_{\hat{W}}$$
$$+ tr\{-\tilde{V}^T(k)\Gamma_v^{-1}\mathrm{Proj}_{\hat{V}} - \mathrm{Proj}_{\hat{V}}^T \Gamma_v^{-1}\tilde{V}(k)$$
$$+ \mathrm{Proj}_{\hat{V}}^T \Gamma_v^{-1} \mathrm{Proj}_{\hat{V}}\} \quad (28)$$

Considering the projection algorithms used, there are four possible Conditions:

(1) All the elements of $\hat{W}(k)$ and $\hat{V}(k)$ are within the known prescribed fictitious bounds;
(2) Only some elements of $\hat{W}(k)$ reach the fictitious bounds, equation (17) is applied;
(3) Only some elements of $\hat{V}(k)$ reach the fictitious bounds, equation (18) is applied;
(4) Some elements of both $\hat{W}(k)$ and $\hat{V}(k)$ reach the fictitious bound, equations (17) and (18) are applied.

We will discuss them one by one in details below.

Condition 1. When all the elements of weight $\hat{W}(k)$ and $\hat{V}(k)$ are within the known prescribed bounds, equation (28) becomes

$$\begin{aligned}\triangle J(k) = & \frac{1}{g_1}[e^2(k+\tau)-e^2(k)] - 2\tilde{W}^T(k)\eta_w \\ & +\eta_w^T\Gamma_w^T\eta_w - 2tr\{\tilde{V}^T(k)\eta_v\} + tr\{\eta_v^T\Gamma_v^T\eta_v\}\end{aligned}$$

From equation (26), we obtain

$$\tilde{W}^T\hat{S} = \frac{e(k+\tau)-\delta_{d_k}}{f_u} - W^{*T}(\hat{S}-S^*) + \varepsilon_u(z)$$

Furthermore, considering the adaptive function (15) and (16), noticing that $tr\{\tilde{V}^T\bar{z}\hat{W}^T\hat{S}'\} = \hat{W}^T\hat{S}'\tilde{V}^T\bar{z}$, then

$$\begin{aligned}\triangle J(k) = & [\frac{1}{g_1} - \frac{2}{f_u}]e^2(k+\tau) - \frac{1}{g_1}e^2(k) \\ & +2W^{*T}(\hat{S}-S^*)e(k+\tau) \\ & -2[\varepsilon_u(z) - \frac{\delta_{d_k}}{f_u}]e(k+\tau) \\ & +\hat{S}^T\Gamma_w^T\hat{S}e^2(k+\tau) - 2\hat{W}^T\hat{S}'\tilde{V}^T\bar{z}e(k+\tau) \\ & +tr\{(\bar{z}\hat{W}^T\hat{S}')^T\Gamma_v^T(\bar{z}\hat{W}^T\hat{S}')\}e^2(k+\tau)\end{aligned}$$

Noticing equation (10), Assumption 3.1 and

$$\begin{aligned}-2[\varepsilon_u(z) & - \frac{\delta_{d_k}}{f_u}]e(k+\tau) \\ & \le k_0[|\varepsilon_u(z)| + |\frac{\delta_{d_k}}{f_u}|]^2 + \frac{1}{k_0}e^2(k+\tau) \\ & \le k_0[\varepsilon_l + \frac{(\tau-1)g_2 d}{\epsilon}]^2 + \frac{1}{k_0}e^2(k+\tau)\end{aligned}$$

$$\hat{S}^T\Gamma_w^T\hat{S}e^2(k+\tau) \le \lambda_w^* l e^2(k+\tau)$$

$$\begin{aligned}tr\{(\bar{z}\hat{W}^T\hat{S}')^T & \Gamma_v^T(\bar{z}\hat{W}^T\hat{S}')\}e^2(k+\tau) \\ & \le \lambda_v^*\|\bar{z}\|^2\sum_{i=1}^{l}[\hat{s}_i'\hat{w}_i]^2 e^2(k+\tau)\end{aligned}$$

where k_0 is a positive number. Therefore

$$\begin{aligned}\triangle J(k) \le & [\frac{1}{g_1} - \frac{2}{f_u} + \frac{1}{k_0} + \lambda_w^* l]e^2(k+\tau) \\ & -\frac{1}{g_1}e^2(k) + \lambda_v^*\|\bar{z}\|^2\sum_{i=1}^{l}[\hat{s}_i'\hat{w}_i]^2 e^2(k+\tau) \\ & +2\hat{W}^T O(\tilde{V}^T\bar{z})^2 e(k+\tau) \\ & -2\tilde{W}^T(\hat{S}-S^*)e(k+\tau) \\ & +k_0[\varepsilon_l + \frac{(\tau-1)g_2 d}{\epsilon}]^2\end{aligned}$$

Noticing inequalities (8) and (11), we have

$$\begin{aligned}\lambda_v^*\|\bar{z}\|^2 & \sum_{i=1}^{l}[\hat{s}_i'\hat{w}_i]^2 e^2(k+\tau) \\ & \le 0.0625\lambda_v^*\|\bar{z}\|^2\|\hat{W}\|^2 e^2(k+\tau) \\ & \le 0.0625\lambda_v^*\|\bar{z}\|^2\hat{w}_m^2 e^2(k+\tau)\end{aligned}$$

$$\begin{aligned}2\hat{W}^T & O(\tilde{V}^T\bar{z})^2 e(k+\tau) \\ & \le 2\|\hat{W}\|(1.2239l + 0.25v_m l\|\bar{z}\|)|e(k+\tau)| \\ & \le 2\hat{w}_m(1.2239l + 0.25v_m l\|\bar{z}\|)|e(k+\tau)| \\ & \le 2.4478\hat{w}_m l|e(k+\tau)| + 0.5\hat{w}_m v_m l\|\bar{z}\||e(k+\tau)| \\ & \le 1.2239k_1\hat{w}_m^2 l^2 + \frac{1.2239}{k_1}e^2(k+\tau) \\ & \quad +\frac{0.25}{k_1}\|\bar{z}\|^2 e^2(k+\tau) + 0.25k_1\hat{w}_m^2 v_m^2 \\ & = \frac{1.2239}{k_1}e^2(k+\tau) + \frac{0.25}{k_1}\|\bar{z}\|^2 e^2(k+\tau) + \beta_1\end{aligned}$$

$$\begin{aligned}-2\tilde{W}^T(\hat{S}-S^*)e(k+\tau) & \le 2\|\tilde{W}\|l|e(k+\tau)| \\ & \le 2\tilde{w}_m l|e(k+\tau)| \\ & \le \frac{1}{k_2}e^2(k+\tau) + k_2\tilde{w}_m^2 l^2 \\ & = \frac{1}{k_2}e^2(k+\tau) + \beta_2\end{aligned}$$

where k_1, k_2, β_1 and β_2 are positive constants and

$$\begin{aligned}\beta_1 & = 1.2239k_1\hat{w}_m^2 l^2 + 0.25k_1\hat{w}_m^2 v_m^2 \\ \beta_2 & = k_2\tilde{w}_m^2 l^2\end{aligned}$$

then we have

$$\begin{aligned}\triangle J(k) \le & [\frac{1}{g_1} - \frac{2}{f_u} + \lambda_w^* l + 0.0625\lambda_v^*\|\bar{z}\|^2\hat{w}_m^2 \\ & +\frac{1.2239}{k_1} + \frac{0.25}{k_1}\|\bar{z}\|^2 + \frac{1}{k_2} + \frac{1}{k_0}]e^2(k+\tau) \\ & -\frac{1}{g_1}e^2(k) + \beta\end{aligned}$$

where

$$\beta = k_0[\varepsilon_l + \frac{(\tau-1)g_2 d}{\epsilon}]^2 + \beta_1 + \beta_2 \qquad (29)$$

is a positive constant. Noticing that k_1 and k_2 are not design parameters, they can be assumed sufficient large. Further the initial value of vector z are in a compact set Ω_z, which will make $\|\bar{z}\|$ bounded, then by choosing appropriate parameters λ_w, λ_v and neuron number l to make the coefficient of $e^2(k+\tau)$ negative, that is to say there exist a positive α satisfying

$$\begin{aligned}-\alpha = & \frac{1}{g_1} - \frac{2}{f_u} + \lambda_w l + 0.0625\lambda_v\|\bar{z}\|^2\hat{w}_m^2 \\ & +\frac{1.2239}{k_1} + \frac{0.25}{k_1}\|\bar{z}\|^2 + \frac{1}{k_2} + \frac{1}{k_0} \qquad (30)\end{aligned}$$

then

$$\triangle J(k) \leq -\alpha e^2(k+\tau) - \frac{1}{g_1} e^2(k) + \beta$$
$$\leq -\alpha e^2(k+\tau) + \beta$$
$$= -\alpha[e^2(k+\tau) - \frac{\beta}{\alpha}]$$

Define compact set

$$\Omega_e := \{e(k) \big| |e(k)| \leq \sqrt{\frac{\beta}{\alpha}}\}$$

then we can see that if the initial parameter $z_0 = [\bar{y}_{k_0}, \bar{u}_{k_0-1}, y_m(k_0+\tau)] \in \Omega_z$ which means that $\|z_0\|$ is bounded, then we can always choose λ_w, λ_v and l to guarantee the existence of positive α. Then the tracking error $e(k)$ will converge to Ω_e as soon as $e(k)$ out of compact Ω_e which lead to the following sequence $\{z(k)\}$ resident in the compact set Ω_z. This will guarantee that for the sequence $\bar{z}(k)$, there is always $\|\bar{z}(k)\|^2 \leq C$ (C is a positive number) holds. Then the positive number α always exist by choose appropriate λ_w, λ_v and l.

Finally, for all $k \geq 0$, $J(k)$ is bounded because

$$J(k) = J(0) + \sum_{j=0}^{k} \triangle J(j) < \infty$$

Condition 2. When only some elements of $\hat{W}(k)$ reach the fictitious bounds, equation (28) becomes

$$\begin{aligned}\triangle J(k) = & \frac{1}{g_1}[e^2(k+\tau) - e^2(k)] \\ & -\tilde{W}^T(k)\Gamma_w^{-1}\text{Proj}_{\hat{W}} - \text{Proj}_{\hat{W}}^T \Gamma_w^{-1}\tilde{W}(k) \\ & +\text{Proj}_{\hat{W}}^T\Gamma_w^{-1}\text{Proj}_{\hat{W}} \\ & +tr\{-\tilde{V}^T(k)\Gamma_v^{-1}\Gamma_v\eta_v - (\Gamma_v\eta_v)^T\Gamma_v^{-1}\tilde{V}(k) \\ & +(\Gamma_v\eta_v)^T\Gamma_v^{-1}\Gamma_v\eta_v\}\end{aligned}$$

Adding and subtracting $-\tilde{W}^T(k)\eta_w - \eta_w^T\tilde{W}(k) + \eta_w^T\Gamma_w^T\eta_w$ to the right side of the above equation, we obtain

$$\begin{aligned}\triangle J(k) = & \frac{1}{g_1}[e^2(k+\tau) - e^2(k)] \\ & -\tilde{W}^T(k)\eta_w - \eta_w^T\tilde{W}(k) + \eta_w^T\Gamma_w^T\eta_w \\ & +tr\{-\tilde{V}^T(k)\Gamma_v^{-1}\Gamma_v\eta_v - (\Gamma_v\eta_v)^T\Gamma_v^{-1}\tilde{V}(k) \\ & +(\Gamma_v\eta_v)^T\Gamma_v^{-1}\Gamma_v\eta_v\} \\ & -\tilde{W}^T(k)(\Gamma_w^{-1}\text{Proj}_{\hat{W}} - \eta_w) \\ & -(\text{Proj}_{\hat{W}}^T\Gamma_w^{-1} - \eta_w^T)\tilde{W}(k) \\ & +\text{Proj}_{\hat{W}}^T\Gamma_w^{-1}\text{Proj}_{\hat{W}} - \eta_w^T\Gamma_w^T\eta_w\end{aligned}$$

Noticing Lemma 4.1, using equation (21) and (23), we have

$$\begin{aligned}\triangle J(k) \leq & \frac{1}{g_1}[e^2(k+\tau) - e^2(k)] \\ & -\tilde{W}^T(k)\eta_w - \eta_w^T\tilde{W}(k) + \eta_w^T\Gamma_w^T\eta_w \\ & +tr\{-\tilde{V}^T(k)\Gamma_v^{-1}\Gamma_v\eta_v - (\Gamma_v\eta_v)^T\Gamma_v^{-1}\tilde{V}(k) \\ & +(\Gamma_v\eta_v)^T\Gamma_v^{-1}\Gamma_v\eta_v\}\end{aligned}$$

which is the same as we discussed in Condition 1. Thus, we have the same result

$$\triangle J(k) \leq -\alpha[e^2(k+\tau) - \frac{\beta}{\alpha}]$$

where α and β are defined in (29) and (30).

Following the same procedure in Condition 2, by adding and subtracting appropriate items, the proof of Conditions 3 and 4 also can be transformed to condition 1, which we have proved that the tracking error to be bounded in a compact set. Its proof is omitted here for clarity. Then Theorem 4.1 holds. **Q.E.D**.

5. CONCLUSION

In this paper, MNN control of a class of non-affine nonlinear discrete-time systems have been investigated. Based on implicit function theorem, multilayer neural networks are used as the emulators to approximate the IDFC controller. Projection algorithms are used to guarantee the boundness of the multi-layer neural network weights. The stability of the closed-loop system is proved rigorously using Lyapunov theorem. The simulation results show the effectiveness of the developed control method. Notice that because of the space limit, simulation results are not given here.

6. REFERENCES

Cabrera, João. B. D. and Kumpati S. Narendra (1999). Issues in the application of neural networks for tracking based on inverse control. *IEEE Transaction On Automatic Control.*

Chen, F. C. and H. K. Khalil (1995). Adaptive control of a class of nonlinear discrete-time systems using neural networks. *IEEE Trans. Automat. Control* **40**, 791–801.

Ge, S. S., C. C. Hang, T. H. Lee and T. Zhang (2001). *Stable Adaptive Neural Network Control.* Kluwer Academic. Norwell, USA.

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

Gong, J. Q. and B. Yao (2001). Neural network adaptive robust control of nonlinear systems in semi-strict feedback form. *Automatica* pp. 1149–1160.

Goodwin, G. C. and D. Q. Mayne (1987). A parameter estimation perspective of continuous time model reference adaptive control. *Automatica* **23**, 57–70.

Kanellakopoulos, I. (1994). A discrete-time adaptive nonlinear system. *IEEE Trans. Automat. Control* **39**, 2362–2365.

Sastry, S. S. and M. Bodson (1989). *Adaptive Control: Stability, Convergence, and Robustness.* Englewood Cliff, NJ: Prentice-Hall.