Proceedings of the IEEE
International Conference on Mechatronics & Automation
Niagara Falls, Canada • July 2005

Backstepping High-Order Differential Neural Network Control of a Type of Nonlinear Systems

Withit Chatlatanagulchai and Peter H. Meckl

School of Mechanical Engineering
Purdue University
West Lafayette, Indiana, USA
{chatlata, meckl}@purdue.edu

Abstract - **This paper presents an output feedback control. The control algorithm does not require plant mathematical model. However, the actual plant is assumed to be affine with respect to local inputs. High-order differential neural networks are used to identify the unknown plant. Luenberger-type observer provides estimated states. Controller is based on backstepping and Lyapunov direct method. Variable structure control handles uncertainties arising from the estimation processes. Closed-loop errors are proved to be bounded. A trajectory tracking example demonstrates the effectiveness of the design.**

Index Terms* – *Differential Neural Network, Backstepping, Variable Structure Control, Observer.

I. Introduction

Technology advance demands better control for systems that are too complicated to be modeled accurately by governing physical laws. Researchers have bypassed this modeling phase by approximating the plant model using intelligent systems such as neural networks and fuzzy logic. These systems have universal approximation property and therefore can approximate any continuous functions to arbitrary accuracy. One of the advantages of doing so is that, once designed, the same intelligent systems can be implemented with physical plants whose dynamics are slightly different without having to redesign.

This so-called model-free scheme is broadly divided into two categories. First, intelligent systems are used to approximate the mapping from input to output of the plant. Then, the control system is designed from the approximated plant to achieve some specific objectives such as tracking or regulating. The approximation is usually performed offline before the actual control process takes place. The offline learning has one major drawback that is poor transient performance during training cycle. The second category, which is more recent, is to represent each part of the unknown system with intelligent system. The approximation process is performed online which eliminates the need of training cycle and the system can adapt itself even when the plant is changed during operation. Early result is presented in [1]. There are several works worth mentioning. In [2], radial basis function neural networks are used with systems in normal form. Variable structure control is used to provide robustness for the system. In [3], three-layer neural network is used to learn unknown SISO plant functions online and a robust state feedback controller is presented for system in strict feedback form. In [4], three-layer neural networks are used to learn an unknown portion of MIMO plant in strict feedback form. In [5], an output feedback controller is presented using three-layer neural networks as function approximators. References [6] and [7] use a type of neuro-fuzzy systems as approximators with direct and indirect control respectively.

In this paper, high-order differential neural network is used as plant functions approximator. Differential neural network differs from feedforward and static neural networks due to the facts that it contains at least one feedback loop and each neuron has its own state variable. High-order network allows higher-order interactions between neurons and has been shown in [8] to have superior storage capacity to the first-order network. In [9], HODNN is used to identify dynamical systems in the form $\dot{x} = f(x,u)$.

Fig. 1 shows the overall system diagram. High-order differential neural network is used as identifier. The estimated plant functions are passed to the observer, as well as actual output and control input. The observer estimates states. Both estimated plant functions and estimated states are fed to controller which generates control input given to the plant. It is shown in the proof to follow that the actual states and the identifier states are close to each other when the neural network weights are ideal. Under an assumption, it will be shown that the observer states are close to the actual states. The controller is designed to reduce the errors between observer states and their desired values. Since uncertainties exist everywhere in the system, backstepping design provides a suitable means to exert the robust control input to each subsystem.

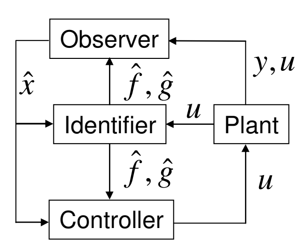

Fig. 1. Overall control system diagram.

The paper is organized as follows. Section II is applicable systems and identifier design. Section III presents observer design. Section IV is controller design. Section V is simulation result. Conclusion is given in Section VI.

II. Applicable Systems and Identifier Design

A. Applicable Systems

Although the closed-form mathematical model representing the plant is not required in the control algorithm, it is assumed that the actual plant is affine with respect to local inputs:

$$\begin{aligned}\dot{x}_i &= f_i\left(\overline{x}_m, t\right) + g_i\left(\overline{x}_m, t\right) x_{i+1},\ 1 \le i \le m-1,\\ \dot{x}_m &= f_m\left(\overline{x}_m, t\right) + g_m\left(\overline{x}_m, t\right) u,\\ y &= x_1,\end{aligned} \tag{1}$$

where $u \in \mathbb{R}$ represents control input, y is system output, $x = \left[x_1, x_2, \ldots, x_m\right]^T \in \mathbb{R}^m$ is vector containing state variables, $\overline{x}_m$ denotes $\left\{x_1, \ldots, x_m\right\}$, and f_j, g_j, $\forall j = 1, \ldots, m$ are unknown functions that depend on all states and may depend on time explicitly. Only output y is measurable.

B. High-Order Differential Neural Network Structure

A HODNN is used to identify the actual plant (1). Each neuron of HODNN represents each subsystem of the actual plant. The identifier system has the following structure

$$\begin{aligned}\dot{\xi}_i &= -a_i \xi_i + b_{hi}\left[\sum_{k=1}^{l} w_{hik}\left(\prod_{j \in I_k} z_j^{d_{jk}}\right)\right]\\ &\quad + b_{gi}\left[\sum_{k=1}^{l} w_{gik}\left(\prod_{j \in I_k} z_j^{d_{jk}}\right)\right]\xi_{i+1}\\ &= -a_i \xi_i + \theta_{hi}^T z + \theta_{gi}^T z \xi_{i+1},\ 1 \le i \le m-1,\\ \dot{\xi}_m &= -a_m \xi_m + \theta_{hm}^T z + \theta_{gm}^T z u,\end{aligned} \tag{2}$$

where ξ_i is the state of the i^{th} neuron, a_i, b_{hi}, b_{gi} are constant scalars, w_{hik}, w_{gik} are the weights connecting the k^{th} node to the i^{th} neuron. z_j is the j^{th} input to the k^{th} node and is defined as

$$\left\{z_1, \ldots, z_m\right\} = \left\{s\left(\xi_1\right), \ldots, s\left(\xi_m\right)\right\},$$

where $s(\bullet)$ is logistic function of the form $s\left(\xi_i\right) = 1/\left(1 + e^{-\xi_i}\right)$. $\left\{I_1, I_2, \ldots, I_l\right\}$ is a set of l not-ordered subsets of $\left\{1, 2, \ldots, m\right\}$. d_{jk} are nonnegative integers. z is a vector given by

$$z = \left[\prod_{j \in I_1} z_j^{d_{j1}}, \prod_{j \in I_2} z_j^{d_{j2}}, \ldots, \prod_{j \in I_l} z_j^{d_{jl}}\right]^T.$$

$\theta_{hi} = b_{hi}\left[w_{hi1}, w_{hi2}, \ldots, w_{hil}\right]^T, \theta_{gi} = b_{gi}\left[w_{gi1}, w_{gi2}, \ldots, w_{gil}\right]^T$ are vectors of adjustable weights. The coefficients $a_i > 0$ are design parameters and are fixed.

For convenience in the proofs to follow, the identifier system (2) can be written in the following vector format

$$\dot{\xi} = A\xi + \Theta_h^T z + \overline{\xi}\ \Theta_g^T z, \tag{3}$$

where $\xi = \left[\xi_1, \xi_2, \ldots, \xi_m\right]^T$, $A = diag\left\{-a_1, -a_2, \ldots, -a_m\right\}$, $\Theta_h = \left[\theta_{h1}, \theta_{h2}, \ldots, \theta_{hm}\right]$, $\Theta_g = \left[\theta_{g1}, \theta_{g2}, \ldots, \theta_{gm}\right]$, $\overline{\xi} = diag\left\{\xi_2, \ldots, u\right\}$.

The actual system (1) is also rewritten in the vector format

$$\dot{x} = F\left(x, t\right) + \overline{x}\ G\left(x, t\right) = Ax + H\left(x, t\right) + \overline{x}\ G\left(x, t\right), \tag{4}$$

where $F\left(x, t\right) = \left[f_1, f_2, \ldots, f_m\right]^T$, $G\left(x, t\right) = \left[g_1, g_2, \ldots, g_m\right]^T$, $\overline{x} = diag\left\{x_2, \ldots, u\right\}$, $H\left(x, t\right) = F\left(x, t\right) - Ax.$

C. Approximation Properties of the HODNN

Good approximation requires $\left\|x(t) - \xi(t)\right\|$ to be bounded by arbitrarily small values. This can be shown by exploring the closeness of solutions of the differential equations representing actual plant (4) and identifier system (3).

Proposition 1: Consider the actual plant (4) and identifier system (3), let $F\left(x, t\right)$ and $G\left(x, t\right)$ be continuous in t and locally Lipschitz in x on $D \times \left[t_0, t_1\right]$, where $D \subset \mathbb{R}^m$ is a compact set. Let $x(t)$ be the solution of (4) with $x\left(t_0\right) = x_0 \in D.$ Suppose $x(t)$ is defined and belongs to D for all $t \in \left[t_0, t_1\right]$. Then for any $\varepsilon > 0$ and any finite $t_1 > 0,$ there exist an integer l, matrices Θ_h^* and Θ_g^*, and $\delta > 0$ such that if $\left\|x_0 - \xi_0\right\| \le \delta$ then there is a unique solution $\xi(t)$ of (3) defined on $\left[t_0, t_1\right]$, with l high-order connections and weight values $\Theta_h = \Theta_h^*$ and $\Theta_g = \Theta_g^*$, with $\xi\left(t_0\right) = \xi_0$, and $\xi(t)$ satisfies

$$\sup_{t_0 \le t \le t_1} \left\|x(t) - \xi(t)\right\| \le \varepsilon.$$

Proof:

The right-hand side of the identifier system (3) is continuous differentiable $\forall \xi \in \mathbb{R}^m, \forall t \in \left[t_0, t_1\right]$, using Theorem 3.2 in [11], $\xi(t)$ exists and is unique on $\left[t_0, t_1\right]$.

Let $U = \left\{\left(\chi, t\right) \in \mathbb{R}^m \times \left[t_0, t_1\right] \middle| \left\|\chi(t) - x(t)\right\| \le \varepsilon, x \in D\right\}$. The set U is compact and is ε - larger than $D \times \left[t_0, t_1\right]$, hence $H\left(x, t\right)$ and $G\left(x, t\right)$ are Lipschitz in x on U, that is

$$\left\|H\left(x, t\right) - H\left(\xi, t\right)\right\| \le L_1 \left\|x - \xi\right\|, \left\|G\left(x, t\right) - G\left(\xi, t\right)\right\| \le L_2 \left\|x - \xi\right\|,$$

where L_1 and L_2 are Lipschitz constants.

The identifier system (3) can be written as

$$\begin{aligned}\dot{\xi} &= A\xi + H\left(\xi, t\right) + \overline{\xi} G\left(\xi, t\right)\\ &\quad + \Theta_h^T z - H\left(\xi, t\right) + \overline{\xi}\ \Theta_g^T z - \overline{\xi} G\left(\xi, t\right).\end{aligned} \tag{5}$$

It can be shown that the functions $\Theta_h^T z$ and $\Theta_g^T z$ satisfy the conditions of the Stone-Weierstrass Theorem [10] and, therefore, can approximate any continuous functions over a compact set. Therefore if the number of high order terms in the neural network l is large enough, then there exist ideal weights $\Theta_h = \Theta_h^*$ and $\Theta_g = \Theta_g^*$ such that

$$\sup_{(\xi, t) \in U} \left\|\Theta_h^{*T} z - H\right\| \le \varepsilon_h, \quad \sup_{(\xi, t) \in U} \left\|\Theta_g^{*T} z - G\right\| \le \varepsilon_g, \tag{6}$$

where ε_h and ε_g are constants.

The solution $x(t)$ of (4) is given by

$$x(t) = x_0 + \int_{t_0}^{t} Ax(s) + H\left(x(s), s\right) + \overline{x}(s) G\left(x(s), s\right) ds.$$

The solution $\xi(t)$ of (5) is given by

$$\begin{aligned}\xi(t) &= \xi_0 + \int_{t_0}^{t} A\xi(s) + H\left(\xi(s), s\right) + \overline{\xi}(s) G\left(\xi(s), s\right) ds\\ &\quad + \int_{t_0}^{t} \Theta_h^{*T} z\left(\xi(s)\right) - H\left(\xi(s), s\right) ds\end{aligned}$$

$$+\int_{t_0}^{t}\bar{\xi}(s)\,\Theta_g^{*T}z(\xi(s))-\bar{\xi}(s)G(\xi(s),s)\,ds.$$

Subtracting the two equations and taking norms yields

$$\begin{aligned}\|x(t)-\xi(t)\| &\le \|x_0-\xi_0\|+\int_{t_0}^{t}(\|A\|+L_1)\|x(s)-\xi(s)\|\\ &+\|\bar{x}G(x,s)-\bar{\xi}G(\xi,s)+\bar{\xi}G(x,s)-\bar{\xi}G(x,s)\|\,ds\\ &+\varepsilon_h(t-t_0)+\varepsilon_g\int_{t_0}^{t}\bar{\xi}(s)\,ds.\end{aligned}$$

By continuity of x and ξ in t and the compactness of $[t_0,t_1]$, x and ξ are bounded on $[t_0,t_1]$. Similarly, since $G(x,t)$ is continuous in x and t, the compactness of $D\times[t_0,t_1]$ implies that $G(x,t)$ is bounded on $D\times[t_0,t_1]$. Later on, we will be able to see that the control input u is also bounded on a compact set. This results in $\|\bar{x}\|\le L_3, \|\bar{\xi}\|\le L_4$, and $\|G(x,t)\|\le L_5$, where L_3, L_4 and L_5 are constants. If $\|x_0-\xi_0\|\le\delta$, the above inequality, then, becomes

$$\begin{aligned}\|x(t)-\xi(t)\| &\le \delta+(\varepsilon_h+L_4\varepsilon_g+(L_3+L_4)L_5)(t-t_0)\\ &+\int_{t_0}^{t}(\|A\|+L_1+L_4L_2)\|x(s)-\xi(s)\|\,ds\\ &=\delta+\mu(t-t_0)+\int_{t_0}^{t}L\|x(s)-\xi(s)\|\,ds,\end{aligned}$$

where μ and L are constants. Using the Bellman-Gronwall Lemma [11] and integration by parts, the inequality becomes

$$\begin{aligned}\|x(t)-\xi(t)\| &\le \delta\exp[L(t-t_0)]+\frac{\mu}{L}\{\exp[L(t-t_0)]-1\}\\ &\le\left(\delta+\frac{\mu}{L}\right)\exp[L(t-t_0)].\end{aligned}$$

Choosing $\delta\le\varepsilon\exp[-L(t_1-t_0)]-\mu/L$ ensures $\|x(t)-\xi(t)\|\le\varepsilon,\ \forall t\in[t_0,t_1]$.

Note that in case (ξ,t) does not belong to U for all $t\in[t_0,t_1]$, δ can be chosen such that $\|x(t)-\xi(t)\|\le\varepsilon_1$, where $\varepsilon_1<\varepsilon$ to ensure that (ξ,t) does not leave $U, \forall t\in[t_0,t_1]$. $[QED]$

The result of Proposition 1 is vital in the design of the robust controller to follow. We are able to claim that the differences between actual states and identifier states are indeed bounded. However, since the proof is quite abstract and the fact that the bound cannot be made arbitrarily small, the identifier states cannot be used as a good estimation of the actual states. In the next section, an observer is designed based on the approximated functions given by the identifier system. We will be able to show that the error between the actual states and the observer states are bounded and the error can be made arbitrarily small by adjusting design parameters.

III. Observer Design

Since actual plant model is unknown, observer is designed from identifier system. The objective is to design an observer such that the difference between observer states and actual states are bounded.

From (2), let

$$\hat{f}_j(\bar{\xi}_m)=-a_j\xi_j+\theta_{hj}^Tz,\ \ \hat{g}_j(\bar{\xi}_m)=\theta_{gj}^Tz,\ \ \forall j=1,\ldots,m, \tag{7}$$

where $\bar{\xi}_m$ denotes $\{\xi_1,\ldots,\xi_m\}$, hence, the identifier system can be rewritten as

$$\begin{aligned}\dot{\xi}_i &= \hat{f}_i(\bar{\xi}_m)+\hat{g}_i(\bar{\xi}_m)\xi_{i+1},\ \ 1\le i\le m-1,\\ \dot{\xi}_m &= \hat{f}_m(\bar{\xi}_m)+\hat{g}_m(\bar{\xi}_m)u,\\ \zeta &= \xi_1,\end{aligned}$$

where ζ is the output of the identifier system. The sequence of output derivatives are

$$\begin{aligned}\zeta &= \xi_1,\\ \dot{\zeta} &= \hat{f}_1(\bar{\xi}_m)+\hat{g}_1(\bar{\xi}_m)\xi_2\\ &=\psi_1(\bar{\xi}_m),\\ \zeta^{(2)} &= \sum_{i=1}^{m-1}\left\{\frac{\partial\psi_1}{\partial\xi_i}\left[\hat{f}_i(\bar{\xi}_m)+\hat{g}_i(\bar{\xi}_m)\xi_{i+1}\right]\right\}\\ &+\frac{\partial\psi_1}{\partial\xi_m}\left[\hat{f}_m(\bar{\xi}_m)+\hat{g}_m(\bar{\xi}_m)u\right]\\ &=\psi_2(\bar{\xi}_m,u),\\ &\vdots\\ \zeta^{(m-1)} &= \psi_{m-1}(\bar{\xi}_m,u,\dot{u},u^{(2)},\ldots,u^{(m-3)}),\end{aligned} \tag{8}$$

We then have the mapping relating the first $m-1$ derivatives of the output ζ to the states of the identifier system and the input derivatives as

$$\begin{aligned}\zeta_e &= \left[\zeta\quad\dot{\zeta}\quad\ldots\quad\zeta^{(m-1)}\right]^T=\hat{H}(\bar{\xi}_m,u,\dot{u},u^{(2)},\ldots,u^{(m-3)})\\ &=\left[\xi_1\quad\psi_1(\bar{\xi}_m)\quad\ldots\quad\psi_{m-1}(\bar{\xi}_m,u,\dot{u},u^{(2)},\ldots,u^{(m-3)})\right]^T.\end{aligned} \tag{9}$$

We obtain equations similar to (8) and (9) for the actual plant (1) as follows:

$$\begin{aligned}y &= x_1,\\ \dot{y} &= f_1(\bar{x}_m,t)+g_1(\bar{x}_m,t)x_2=\varphi_1(\bar{x}_m,t),\\ y^{(2)} &= \frac{\partial\varphi_1}{\partial t}+\sum_{i=1}^{m-1}\left\{\frac{\partial\varphi_1}{\partial x_i}[f_i(\bar{x}_m,t)+g_i(\bar{x}_m,t)x_{i+1}]\right\}\\ &+\frac{\partial\varphi_1}{\partial x_m}[f_m(\bar{x}_m,t)+g_m(\bar{x}_m,t)u]\\ &=\varphi_2(\bar{x}_m,u,t),\\ &\vdots\\ y^{(m-1)} &= \varphi_{m-1}(\bar{x}_m,u,\dot{u},u^{(2)},\ldots,u^{(m-3)},t),\end{aligned}$$

and

$$\begin{aligned}y_e &= \left[y\quad\dot{y}\quad\ldots\quad y^{(m-1)}\right]^T=H(\bar{x}_m,u,\dot{u},u^{(2)},\ldots,u^{(m-3)},t)\\ &=\left[x_1\quad\varphi_1(\bar{x}_m,t)\quad\ldots\quad\varphi_{m-1}(\bar{x}_m,u,\dot{u},u^{(2)},\ldots,u^{(m-3)},t)\right]^T.\end{aligned}$$

Assumption 1: Actual plant (1) and identifier system (2) are uniformly completely observable, i.e., the mapping H and $\hat{H}$ are invertible with respect to $\bar{x}_m$ and $\bar{\xi}_m$ and their inverses, $\bar{x}_m=H^{-1}(y_e,u,\dot{u},u^{(2)},\ldots,u^{(m-3)},t)$ and $\bar{\xi}_m=\hat{H}^{-1}(\zeta_e,u,\dot{u},u^{(2)},\ldots,u^{(m-3)})$, are smooth. Moreover, if $\|y_e-\zeta_e\|\le\varepsilon_1$ for some $\varepsilon_1>0$, then $\|H^{-1}(\cdot)-\hat{H}^{-1}(\cdot)\|\le\varepsilon_2$ for some finite $\varepsilon_2>0$.

The m^{th} derivative of the output is given by

$$\zeta^{(m)} = \sum_{k=1}^{m-2}\left\{\frac{\partial \psi_{m-1}}{\partial u^{(k-1)}}u^{(k)}\right\} + \sum_{l=1}^{m-1}\left\{\frac{\partial \psi_{m-1}}{\partial \xi_l}\left[\hat{f}_l(\bar{\xi}_m) + \hat{g}_l(\bar{\xi}_m)\xi_{l+1}\right\}\right.$$
$$+\frac{\partial \psi_{m-1}}{\partial \xi_m}\left[\hat{f}_m(\bar{\xi}_m) + \hat{g}_m(\bar{\xi}_m)u\right]$$
$$= \hat{\alpha}(\zeta_e) + \hat{\beta}_1(\zeta_e)u + \ldots + \hat{\beta}_{m-1}(\zeta_e)u^{(m-2)}.$$

Similarly, for the actual plant, we have

$$y^{(m)} = \sum_{k=1}^{m-2}\left\{\frac{\partial \varphi_{m-1}}{\partial u^{(k-1)}}u^{(k)}\right\} + \sum_{l=1}^{m-1}\left\{\frac{\partial \varphi_{m-1}}{\partial x_l}\left[f_l(\bar{x}_m,t) + g_l(\bar{x}_m,t)x_{l+1}\right\}\right.$$
$$+\frac{\partial \varphi_{m-1}}{\partial x_m}\left[f_m(\bar{x}_m,t) + g_m(\bar{x}_m,t)u\right] + \frac{\partial \varphi_{m-1}}{\partial t}$$
$$= \alpha(y_e) + \beta_1(y_e)u + \ldots + \beta_{m-1}(y_e)u^{(m-2)},$$

and

$$\dot{y}_e = \bar{A}y_e + \bar{B}\left[\alpha(y_e) + \beta_1(y_e)u + \ldots + \beta_{m-1}(y_e)u^{(m-2)}\right],$$

where

$$\bar{A} = \begin{bmatrix} 0 & I \\ 0 & 0 \end{bmatrix}, \quad \bar{B} = \begin{bmatrix} 0 & \ldots & 0 & 1 \end{bmatrix}^T.$$

Proposition 2: Using nonlinear observer

$$\begin{bmatrix} \dot{\hat{x}}_1 \\ \dot{\hat{x}}_2 \\ \vdots \\ \dot{\hat{x}}_m \end{bmatrix} = \begin{bmatrix} \hat{f}_1(\bar{\hat{x}}_m) + \hat{g}_1(\bar{\hat{x}}_m)\hat{x}_2 \\ \hat{f}_2(\bar{\hat{x}}_m) + \hat{g}_2(\bar{\hat{x}}_m)\hat{x}_3 \\ \vdots \\ \hat{f}_m(\bar{\hat{x}}_m) + \hat{g}_m(\bar{\hat{x}}_m)u \end{bmatrix} + \left[\frac{\partial \hat{H}(\bar{\hat{x}}_m)}{\partial \bar{\hat{x}}_m}\right]^{-1} \varepsilon^{-1}\bar{L}\left[y - \hat{\zeta}\right], \quad (10)$$
$$\hat{\zeta} = \hat{x}_1,$$

where $\hat{x}_j,\ j = 1,\ldots,m$ are observer states, $\hat{\zeta}$ is output of the observer system, $\varepsilon = diag\left[\eta, \eta^2, \ldots, \eta^m\right]$, $0 < \eta \le 1$ and $\left[\partial \hat{H}(\bar{\hat{x}}_m)/\partial \bar{\hat{x}}_m\right]$ is the Jacobian of $\hat{H}$, η is a design parameter and $\bar{L} = [l_1, l_2, \ldots, l_m]^T$ is such that $s^m + l_1 s^{m-1} + \ldots + l_m$ is a Hurwitz polynomial. Then, the estimation error, $\|x - \hat{x}\|$ is globally uniformly ultimately bounded, where $\hat{x} = \left[\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_m\right]^T$.

Proof: From (9), replacing identifier states ξ_j with observer states $\hat{x}_j$, we have

$$\hat{\zeta}_e = \left[\hat{x}_1 \quad \psi_1(\bar{\hat{x}}_m) \quad \ldots \quad \psi_{m-1}(\bar{\hat{x}}_m, u, \dot{u}, u^{(2)}, \ldots, u^{(m-3)})\right]^T.$$

All derivatives of $\hat{\zeta}_{e_i}$ can be found as follows:

$$\dot{\hat{\zeta}}_{e_1} = \dot{\hat{x}}_1 = \hat{\zeta}_{e_2} + I_1\left[\partial \hat{H}(\bar{\hat{x}}_m)/\partial \bar{\hat{x}}_m\right]^{-1}\varepsilon^{-1}\bar{L}\left[y - \hat{x}_1\right],$$
$$\dot{\hat{\zeta}}_{e_2} = \dot{\psi}_1(\bar{\hat{x}}_m)$$
$$= \hat{\zeta}_{e_3} + \sum_{l=1}^{m}\left\{\frac{\partial \psi_1}{\partial \hat{x}_l}\left[I_l\left[\partial \hat{H}(\bar{\hat{x}}_m)/\partial \bar{\hat{x}}_m\right]^{-1}\varepsilon^{-1}\bar{L}\left[y - \hat{x}_1\right]\right]\right\},$$
$$\dot{\hat{\zeta}}_{e_m} = \dot{\psi}_{m-1}(\bar{\hat{x}}_m, u, \dot{u}, u^{(2)}, \ldots, u^{(m-3)})$$
$$= \hat{\alpha}(\hat{\zeta}_e) + \hat{\beta}_1(\hat{\zeta}_e)u + \ldots + \hat{\beta}_{m-1}(\hat{\zeta}_e)u^{(m-2)}$$
$$+\sum_{l=1}^{m}\left\{\frac{\partial \psi_{m-1}}{\partial \hat{x}_l}\left[I_l\left[\partial \hat{H}(\bar{\hat{x}}_m)/\partial \bar{\hat{x}}_m\right]^{-1}\varepsilon^{-1}\bar{L}\left[y - \hat{x}_1\right]\right]\right\}.$$

$I_j \in \mathbb{R}^{1\times m}$ is a row vector whose j^{th} element is 1 and 0 otherwise. The above can be put in matrix form as

$$\dot{\hat{\zeta}}_e = \bar{A}\hat{\zeta}_e + \bar{B}\left[\hat{\alpha}(\hat{\zeta}_e) + \hat{\beta}_1(\hat{\zeta}_e)u + \ldots + \hat{\beta}_{m-1}(\hat{\zeta}_e)u^{(m-2)}\right]$$
$$+\varepsilon^{-1}\bar{L}\left[y - \hat{x}_1\right].$$

Define the observer error, $\tilde{\zeta}_e = \hat{\zeta}_e - y_e$. Let $\bar{C} \in \mathbb{R}^{1\times m} = [1, 0, \ldots, 0]$, then the observer error dynamics are given by

$$\dot{\tilde{\zeta}}_e = (\bar{A} - \varepsilon^{-1}\bar{L}\bar{C})\tilde{\zeta}_e + \bar{B}\left[\hat{\alpha}(\hat{\zeta}_e) + \hat{\beta}_1(\hat{\zeta}_e)u + \ldots + \hat{\beta}_{m-1}(\hat{\zeta}_e)\right.$$
$$\left.u^{(m-2)}\right] - \bar{B}\left[\alpha(y_e) + \beta_1(y_e)u + \ldots + \beta_{m-1}(y_e)u^{(m-2)}\right].$$

Define a transformation, $\tilde{v} \triangleq \varepsilon'\tilde{\zeta}_e$, where $\varepsilon' = diag\left[1/\eta^{m-1}, 1/\eta^{m-2}, \ldots, 1\right]$, let P be the solution of the Lyapunov equation $P(\bar{A} - \bar{L}\bar{C}) + (\bar{A} - \bar{L}\bar{C})^T P = -I$, the derivative of a Lyapunov function $V = \tilde{v}^T P\tilde{v} > 0$ is given by

$$\dot{V} = -\tilde{v}^T\tilde{v}/\eta + 2\tilde{v}^T P\bar{B}\left[\hat{\alpha}(\hat{\zeta}_e) + \hat{\beta}_1(\hat{\zeta}_e)u + \ldots + \hat{\beta}_{m-1}(\hat{\zeta}_e)\right.$$
$$\left.u^{(m-2)}\right] - 2\tilde{v}^T P\bar{B}\left[\alpha(y_e) + \beta_1(y_e)u + \ldots + \beta_{m-1}(y_e)u^{(m-2)}\right].$$

From (6) and the fact that $\hat{f}_j, \hat{g}_j$ are smooth functions and later on we will see that the control input u is also smooth, we have

$$\left|\hat{\alpha}(\hat{\zeta}_e) - \alpha(\hat{\zeta}_e)\right| \le k_1, \left|\hat{\beta}_i(\hat{\zeta}_e) - \beta_i(\hat{\zeta}_e)\right| \le k_{2,i},\ i = 1,\ldots,m-1,$$

where $k_1, k_{2,i}$ are bounded constants. The functions α, β_j are Lipschitz and therefore, we have

$$\left|\begin{array}{l}\hat{\alpha}(\hat{\zeta}_e) + \hat{\beta}_1(\hat{\zeta}_e)u + \ldots + \hat{\beta}_{m-1}(\hat{\zeta}_e)u^{(m-2)} - \alpha(y_e) \\ -\beta_1(y_e)u - \ldots - \beta_{m-1}(y_e)u^{(m-2)}\end{array}\right|$$
$$\le k_1 + \left|\alpha(\hat{\zeta}_e) - \alpha(y_e)\right| + k_{2,1}|u| + \left|\beta_1(\hat{\zeta}_e) - \beta_1(y_e)\right||u|$$
$$+ \ldots + k_{2,m-1}\left|u^{(m-2)}\right| + \left|\beta_{m-1}(\hat{\zeta}_e) - \beta_{m-1}(y_e)\right|\left|u^{(m-2)}\right|$$
$$\le k_3 + k_4\left|\tilde{\zeta}_e\right|,$$

where k_3, k_4 are non-negative constants. Thus we have

$$\dot{V} \le -\|\tilde{v}\|^2/\eta + 2k_3\|P\|\|\tilde{v}\| + 2k_4\|P\|\|\tilde{v}\|^2.$$

By using part of $-\|\tilde{v}\|^2/\eta$ to dominate $2k_3\|P\|\|\tilde{v}\|$ for large $\|\tilde{v}\|$, we can conclude that $\tilde{v}$ and hence $\tilde{\zeta}_e$ are globally uniformly ultimately bounded (GUUB). From Assumption 1, $\|x - \hat{x}\|$ is GUUB. [*QED*]

Note that having $\tilde{\zeta}_e$ GUUB is sufficient to guarantee $|x_1 - \hat{x}_1|$ GUUB since $|x_1 - \hat{x}_1| \le \|\tilde{\zeta}_e\|$. Later on, we will see that the controller is designed to make the errors, between observer states $\hat{x}$ (since actual states are not available) and their desired values, small. Therefore, having $|x_1 - \hat{x}_1|$ bounded to an arbitrarily small value results in $|x_1 - x_{1d}|$ small and our tracking objective is achieved.

IV. Controller Design

Since actual plant states are not available, the objective is to reduce the errors between observer state and desired state of each subsystem in the backstepping process. Let $e_j = \hat{x}_j - x_{jd},\ j = 1,\ldots,m$ be those errors. Let $\tilde{x}_j = \hat{x}_j - x_j,\ j = 1,\ldots,m$ be the error between observer state and actual state and $\tilde{\theta}_j = \theta_j - \theta_j^*,\ j = 1,\ldots,m$ be the error between the weight used in the identifier system and the ideal weight.

From (6), $\left|h_1 - \theta_{h1}^{*T} z\right| \le \varepsilon_{h1}, \left|g_1 - \theta_{g1}^{*T} z\right| \le \varepsilon_{g1}$. From Proposition 1, $\left|\xi_1 - x_1\right| \le \varepsilon_1$. From (1), (10) and the fact that control input u is smooth, we have $\left|\dot{\hat{x}}_1\right| \le \varepsilon_{\dot{\hat{x}}1}$. Therefore, the following inequality holds

$$\left|h_1 - \theta_{h1}^{*T} z\right| + \left|a_1(\xi_1 - x_1)\right| + \left|\dot{\hat{x}}_1\right| + \left|\left(g_1 - \theta_{g1}^{*T} z\right)x_{2d}\right| \le K_1^{*T}\varphi_1,$$

where $K_1^* = \left[\varepsilon_{h1}, \varepsilon_1, \varepsilon_{\dot{\hat{x}}1}, \varepsilon_{g1}\right]^T$, $\varphi_1 = \left[1, |a_1|, 1, |x_{2d}|\right]^T$.

Let the virtual control of the first subsystem be

$$\begin{aligned} x_{2d} &= -\hat{g}_1^{-1}\left(c_1 e_1 + \hat{f}_1 - \dot{x}_{1d} + u_{2dvsc}\right) \\ &= -\left(\theta_{g1}^T z\right)^{-1}\left(c_1 e_1 - a_1\xi_1 + \theta_{h1}^T z - \dot{x}_{1d} + u_{2dvsc}\right), \end{aligned} \tag{11}$$

where the smooth variable structure control is given by

$$u_{2dvsc} = K_1^T \bar{\varphi}_1, \tag{12}$$

where

$$\begin{aligned} \bar{\varphi}_1 = \Bigg[&\frac{2}{\pi}\arctan\left(\frac{e_1}{\mu_1}\right), |a_1|\frac{2}{\pi}\arctan\left(\frac{e_1}{\mu_1}|a_1|\right) \\ &, \frac{2}{\pi}\arctan\left(\frac{e_1}{\mu_1}\right), |x_{2d}|\frac{2}{\pi}\arctan\left(\frac{e_1}{\mu_1}|x_{2d}|\right)\Bigg]^T, \end{aligned} \tag{13}$$

μ_1 is a small positive number. K_1 approximates K_1^*. The $\dot{e}_1$ equation becomes

$$\begin{aligned} \dot{e}_1 &= \dot{\hat{x}}_1 - \dot{x}_{1d} \\ &= \left(f_1 - \hat{f}_1\right) + g_1 e_2 + \dot{\hat{x}}_1 + \left(g_1 - \hat{g}_1\right)x_{2d} - c_1 e_1 - u_{2dvsc} \\ &= \left(h_1 - \theta_{h1}^{*T} z\right) + a_1(\xi_1 - x_1) + \dot{\hat{x}}_1 + \left(g_1 - \theta_{g1}^{*T} z\right)x_{2d} \\ &\quad -\tilde{\theta}_{h1}^T z - \tilde{\theta}_{g1}^T z x_{2d} + g_1 e_2 - c_1 e_1 - K_1^T \bar{\varphi}_1. \end{aligned}$$

Similarly, the virtual control of the i^{th} subsystem, $2 \le i \le m-1$, is given by

$$\begin{aligned} x_{(i+1)d} &= -\hat{g}_i^{-1}\left(g_{(i-1)U} e_{i-1} + c_i e_i + \hat{f}_i - \dot{x}_{id} + u_{(i+1)dvsc}\right) \\ &= -\left(\theta_{gi}^T z\right)^{-1}\Big(g_{(i-1)U} e_{i-1} + c_i e_i - a_i \xi_i + \theta_{hi}^T z \\ &\quad -\dot{x}_{id} + u_{(i+1)dvsc}\Big), \end{aligned}$$

where $\left|g_j(\bar{x}_m, t)\right| \le g_{jU}, \forall j = 1,\ldots,m-1$. The fact that $g_j(\bar{x}_m, t)$ is bounded by a constant can be seen from g_i being smooth and $x_1,\ldots,x_m$ and t are in compact sets. The variable structure control $u_{(j+1)dvsc}, j = 2,\ldots,m$ is the same as those in (12) and (13), only the index is changed. The $\dot{e}_i$ equation becomes

$$\begin{aligned} \dot{e}_i &= \left(h_i - \theta_{hi}^{*T} z\right) + a_i(\xi_i - x_i) + \dot{\hat{x}}_i + \left(g_i - \theta_{gi}^{*T} z\right)x_{(i+1)d} \\ &\quad -\tilde{\theta}_{hi}^T z - \tilde{\theta}_{gi}^T z x_{(i+1)d} + g_i e_{(i+1)} - c_i e_i - K_i^T \bar{\varphi}_i - g_{(i-1)U} e_{i-1}. \end{aligned}$$

The actual control input is given by

$$\begin{aligned} u = -\left(\theta_{gm}^T z\right)^{-1}\Big(&g_{(m-1)U} e_{m-1} + c_m e_m - a_m \xi_m + \theta_{hm}^T z \\ &-\dot{x}_{md} + u_{(m+1)dvsc}\Big). \end{aligned} \tag{14}$$

The $\dot{e}_m$ equation is given by

$$\begin{aligned} \dot{e}_m &= \left(h_m - \theta_{hm}^{*T} z\right) + a_m(\xi_m - x_m) + \dot{\hat{x}}_m + \left(g_m - \theta_{gm}^{*T} z\right)u \\ &\quad -\tilde{\theta}_{hm}^T z - \tilde{\theta}_{gm}^T z u - c_m e_m - K_m^T \bar{\varphi}_m - g_{(m-1)U} e_{m-1}. \end{aligned}$$

Choose the Lyapunov function of the whole system as

$$V = \sum_{i=1}^{m}\left\{\frac{1}{2}e_i^2 + \frac{1}{2}\tilde{K}_i^T \Gamma_{ki}^{-1}\tilde{K}_i + \frac{1}{2}\tilde{\theta}_{hi}^T \Gamma_{hi}^{-1}\tilde{\theta}_{hi} + \frac{1}{2}\tilde{\theta}_{gi}^T \Gamma_{gi}^{-1}\tilde{\theta}_{gi}\right\},$$

where $\tilde{K}_j = K_j - K_j^*,\ j = 1,\ldots,m,$ and $\Gamma_{kj}, \Gamma_{hj}, \Gamma_{gj},\ j = 1,\ldots,m$ are positive constants. For all $j = 1,\ldots,m-1,$ let the weight adaptation laws be

$$\begin{aligned} \dot{K}_j &= \dot{\tilde{K}}_j = \Gamma_{kj}\left[\bar{\varphi}_j e_j - \sigma_{kj} K_j\right], \\ \dot{\theta}_{hj} &= \dot{\tilde{\theta}}_{hj} = \Gamma_{hj}\left[z e_j - \sigma_{hj}\theta_{hj}\right], \\ \dot{\theta}_{gj} &= \dot{\tilde{\theta}}_{gj} = \Gamma_{gj}\left[z e_j x_{(j+1)d} - \sigma_{gj}\theta_{gj}\right]. \end{aligned} \tag{15}$$

For the last step $(j = m)$, replace $x_{(j+1)d}$ by u. $\sigma_{kj}, \sigma_{hj}, \sigma_{gj},\ j = 1,\ldots,m$ are positive constants. Using the following facts

$$0 \le |\alpha| - \alpha\frac{2}{\pi}\arctan\left(\frac{\alpha}{\eta}\right) \le 0.2785\eta, \quad \forall \alpha \in \mathbb{R},$$

$$2\tilde{K}_j^T K_j = \left\|\tilde{K}_j\right\|^2 + \left\|K_j\right\|^2 - \left\|K_j^*\right\|^2 \ge \left\|\tilde{K}_j\right\|^2 - \left\|K_j^*\right\|^2,$$

$$2\tilde{\theta}_{hj}^T \theta_{hj} = \left\|\tilde{\theta}_{hj}\right\|^2 + \left\|\theta_{hj}\right\|^2 - \left\|\theta_{hj}^*\right\|^2 \ge \left\|\tilde{\theta}_{hj}\right\|^2 - \left\|\theta_{hj}^*\right\|^2,$$

$$2\tilde{\theta}_{gj}^T \theta_{gj} = \left\|\tilde{\theta}_{gj}\right\|^2 + \left\|\theta_{gj}\right\|^2 - \left\|\theta_{gj}^*\right\|^2 \ge \left\|\tilde{\theta}_{gj}\right\|^2 - \left\|\theta_{gj}^*\right\|^2,$$

the derivative $\dot{V}$ can be found as follows:

$$\begin{aligned} \dot{V} &= \sum_{i=1}^{m}\left\{e_i\dot{e}_i + \tilde{K}_i^T \Gamma_{ki}^{-1}\dot{\tilde{K}}_i + \tilde{\theta}_{hi}^T \Gamma_{hi}^{-1}\dot{\tilde{\theta}}_{hi} + \tilde{\theta}_{gi}^T \Gamma_{gi}^{-1}\dot{\tilde{\theta}}_{gi}\right\} \\ &\le \sum_{i=1}^{m}\left\{-c_i e_i^2 - \frac{\sigma_{ki}}{2}\left\|\tilde{K}_i\right\|^2 - \frac{\sigma_{hi}}{2}\left\|\tilde{\theta}_{hi}\right\|^2 - \frac{\sigma_{gi}}{2}\left\|\tilde{\theta}_{gi}\right\|^2 + \phi_i\right\}, \end{aligned}$$

where

$$\begin{aligned} \phi_i &= 0.2785\mu_i\left(\varepsilon_{hi} + \varepsilon_i + \varepsilon_{\dot{\hat{x}}i} + \varepsilon_{gi}\right) \\ &\quad + \frac{\sigma_{ki}}{2}\left\|\tilde{K}_i^*\right\|^2 + \frac{\sigma_{hi}}{2}\left\|\tilde{\theta}_{hi}^*\right\|^2 + \frac{\sigma_{gi}}{2}\left\|\tilde{\theta}_{gi}^*\right\|^2. \end{aligned}$$

Let $\varsigma = \min_{1\le i\le m}\{2c_i\} \ge 0$, $\delta = \sum_{i=1}^{m}\phi_i \ge 0$ and choose

$$\sigma_{hi} \ge \varsigma\lambda_{\max}\left\{\Gamma_{hi}^{-1}\right\}, \sigma_{gi} \ge \varsigma\lambda_{\max}\left\{\Gamma_{gi}^{-1}\right\},$$

$$\sigma_{ki} \ge \varsigma\lambda_{\max}\left\{\Gamma_{ki}^{-1}\right\},\ i = 1,\ldots,m,$$

we have

$$\dot{V} \le -\varsigma V + \delta.$$

From this point on, using standard stability analysis, it can be shown that the trajectory $e_i, \tilde{K}_i, \tilde{\theta}_{hi}, \tilde{\theta}_{gi}$ are GUUB.

V. Simulation Result

The following system is used as the actual plant

$$\begin{aligned} \dot{x}_1 &= \left[0.5x_1^2 x_2 \sin t - \left(1 + 0.1x_1^2\right)\left(2 + \sin x_1\right)\right] \\ &\quad + \left[1 + 0.1x_1^2 + 0.1x_2^2\right]x_2, \end{aligned}$$

$$\dot{x}_2 = \left[x_1 x_2 - (2+\cos x_1)\left(0.3\left(e^{x_1}+e^{-x_1}\right)\right)\right] + \left[2+\cos(x_1 x_2)+0.1\cos t\right]u,$$

$$y = x_1.$$

It is assumed that all functions in the actual plant are unknown. Only output y is available from measurement. The objective is to make the output track a signal, obtained from passing a square wave of amplitude 7, zero mean, and 20-sec period into the filter $1/(s+2)^3$, as close as possible while all the closed-loop signals remain bounded.

The identifier system (2) has the following design parameters, $a_1 = a_2 = 10, l = 4, z = \left[z_1, z_2, z_1 z_2, z_1^2\right]^T$. The observer system (10) has the following design parameters, $\eta = 0.1, \bar{L} = [5, 7]^T$. The virtual control input x_{2d} in (11) and the actual control input u in (14) has the following design parameters, $c_1 = c_2 = 50, g_{1U} = 5$. The variable structure control u_{2dvsc} and u_{3dvsc} in (12) has $\mu_1 = \mu_2 = 0.1$.
The weight adaptation laws (15) has the following design parameters, $\Gamma_{k1} = \Gamma_{k2} = I^{4\times4}, \Gamma_{h1} = \Gamma_{h2} = 5I^{4\times4}$, $\Gamma_{g1} = \Gamma_{g2} = 0.005I^{4\times4}, \sigma_{k1} = \sigma_{k2} = 0, \sigma_{h1} = \sigma_{h2} = 1$, $\sigma_{g1} = \sigma_{g2} = 0$. The control saturation limits are set at ± 20. The sampling period is 0.001 sec. The initial conditions are set as follows: $x_1(0) = x_2(0) = 0$,
$\hat{x}_1(0) = \hat{x}_2(0) = 0.1, \xi_1(0) = \xi_2(0) = 0$,
$\theta_{h1}(0) = \theta_{h2}(0) = [0.5, 0.5, 0.5, 0.5]^T$,
$\theta_{g1}(0) = [0.7, 0.7, 0.7, 0.7]^T, \theta_{g2}(0) = [1.9, 1.9, 1.9, 1.9]^T$,
$K_1(0) = K_2(0) = [0,0,0,0]^T$.

Fig. 2a and b show good overall tracking performance. The controller is designed to reduce the errors between observer states and their desired values. In Fig. 2c, the controller demonstrates good performance by obtaining small errors between observer state $\hat{x}_2$ and desired state x_{2d}.

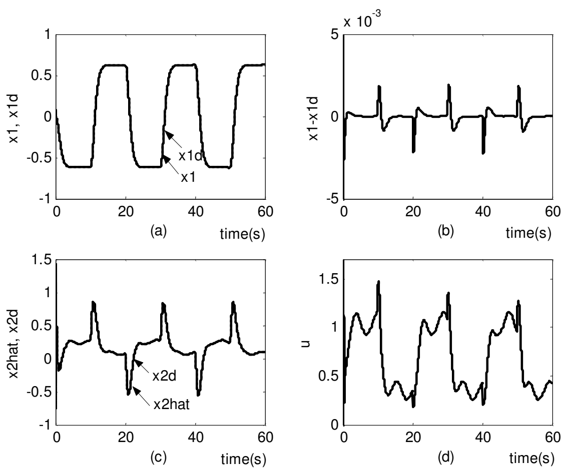

Fig. 2. Overall system performance. (a) Actual output x_1 versus desired output x_{1d}. (b) Tracking error $x_1 - x_{1d}$. Controller performance. (c) Observer state $\hat{x}_2$ versus desired value x_{2d}. (d) Control input u.

VI. Conclusion

The main contribution of this paper is that it incorporates the use of high-order differential neural network, primarily used in system identification, to the controller design. The structure of the differential neural network allows us to show that the identifier state trajectories stay close to the actual states when the neural network weights are ideal. We are able to show that, using the Luenberger-type observer, the differences between actual states and observer states are bounded by an adjustable bound. Since both function and state estimation processes cause uncertainties to arise in each subsystem, backstepping framework provides possible way to exert the control force directly into each subsystem. The variable structure control, used to provide robustness to the system, does not require the exact bounds of the uncertainties since the parameter in the controller is adapted based on the size of the tracking errors. With all things put together, we are able to control a rather complicated plant, e.g., the one given in the simulation, quite well using only the output of the system.

References

[1] F. L. Lewis, S. Jagannathan, and A. Yesildirek, *Neural Network Control of Robot Manipulators and Nonlinear Systems.* UK: Taylor & Francis, 1999.

[2] S. S. Ge, C. C. Hang, T. H. Lee, and T. Zhang, *Stable Adaptive Neural Network Control.* The Netherlands: Kluwer, 2002.

[3] J. Q. Gong and B. Yao, "Neural network adaptive robust control of nonlinear systems in semi-strict feedback form," *Automatica,* vol. 37, no. 8, pp. 1149-1160, 2001.

[4] C. Kwan, and F. L. Lewis, "Robust backstepping control of nonlinear systems using neural networks," *IEEE Trans. on Systems, Man, and Cybernetics,* vol. 30, no. 6, pp. 753-766, 2000.

[5] W. Chatlatanagulchai and P. H. Meckl, "Robust observer backstepping neural network control of nonlinear systems in strict feedback form," in *Proc. American Control Conf.,* Boston, 2004, pp. 3035-3040.

[6] W. Chatlatanagulchai and P. H. Meckl, "Observer backstepping neuro-fuzzy control design for a type of nonlinear system," in *Proc. ASME Int. Mech. Eng. Cong. and RD&D Expo.,* Anaheim, 2004.

[7] W. Chatlatanagulchai and P. H. Meckl, "Model-free observer backstepping control design for nonlinear systems in strict feedback form," in *Proc. of IEEE Conf. on Decision and Control.,* Nassau, 2004, pp. 5439-5444.

[8] P. Paretto and J. J. Niez, "Long term memory storage capacity of multiconnected neural networks," *Biol. Cybern.,* vol. 54, pp. 53-63, 1986.

[9] E. B. Kosmatopoulos, M. M. Polycarpou, M. A. Christodoulou, and P. A. Ioannou, "High-order neural network structures for identification of dynamical systems," *IEEE Trans. Neural Networks*, vol. 6, no. 2, pp. 422-430, March 1995.

[10] N. E. Cotter, "The Stone-Weierstrass theorem and its application to neural networks," *IEEE Trans. on Neural Networks*, vol. 1, no. 4, Dec 1990.

[11] H. K. Khalil, *Nonlinear Systems.* Englewood Cliffs, NJ: Prentice Hall, 2002.