

# Approximation of Dynamical Time-Variant Systems by Continuous-Time Recurrent Neural Networks

Xiao-Dong Li, John K. L. Ho, and Tommy W. S. Chow

**Abstract**—This paper studies the approximation ability of continuous-time recurrent neural networks to dynamical time-variant systems. It proves that any finite time trajectory of a given dynamical time-variant system can be approximated by the internal state of a continuous-time recurrent neural network. Given several special forms of dynamical time-variant systems or trajectories, this paper shows that they can all be approximately realized by the internal state of a simple recurrent neural network.

**Index Terms**—Approximation, dynamical time-variant systems, recurrent neural networks.

## I. INTRODUCTION

In recent years, there have been a lot of research works focusing on the theoretical aspects of neural networks (NNs)-based automatic control system. This is largely due to the increasing demand for applications in system identification, intelligent control system, and signal processing. The performance of NN-based applications relies heavily upon the functional approximation capability of the NN. Most of these applications, particularly in the areas of nonlinear system identification and control, are in fact dealing with the fundamental issue of approximating a nonlinear function or a dynamical system. Feedforward neural networks (FNN) and recurrent neural networks (RNN) are the two major classes of NN widely used in the area of dynamical systems. It is well known that RNN consisting of a large number of feedforward and feedback connections exhibits complex dynamics [1], [2], [13]. When one is dealing with dynamical systems, the dynamical structure of RNN provides significant advantages over the feedforward structure. In addition, an RNN with a relatively smaller network size has been shown to be equivalent in approximation capability to a larger feedforward network architecture [3]. In this paper, the issue of approximating dynamical time-variant systems using an RNN is studied.

Although most current theoretical studies on the NN focus mainly on the stability and convergence of the network trajec-

tory to the equilibrium, there are a considerable amount of results on the approximation capability of the NN reported. For instance, it has been mathematically proved that a given continuous mapping on a compact set could be approximately realized by using a three-layer FNN to any precision [3]–[5]. Li [6] showed that a discrete-time trajectory on a closed finite interval could be represented exactly using a discrete-time RNN. Jin and Nikiforuk [7] also studied the approximation problem of approximating nonlinear discrete-time state-space trajectories with input using discrete-time RNN. In the case of continuous-time RNN, Funahashi and Nakamura [8] studied the approximation of continuous-time dynamic systems using a Hopfield-type RNN. They proved that a continuous-time dynamical system without input, i.e.,  $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x})$ , can be approximated by a class of RNN to an arbitrary degree of accuracy [8]. The dynamical system,  $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x})$ , studied by Funahashi and Nakamura is, in fact, a special type of dynamical systems not usually associated with a control environment. Recently, Chow and Li [9] extended Funahashi and Nakamura's work [8] to hold for more general dynamical continuous-time systems with a control input, i.e.,  $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}, \mathbf{u})$ . Kambhampati, Garces and Warwick [10] studied a special dynamical system of Chow and Li, i.e.  $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x}) \cdot \mathbf{u}$ . All of these works, however, only focus on the approximation problem of continuous-time RNN to dynamical time-invariant system, i.e.,  $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}, \mathbf{u})$ . It is also worth noting that dynamical time-invariant system is a special case of dynamical time-variant systems, i.e.,  $\dot{\mathbf{x}}(t) = \mathbf{F}(\mathbf{x}, \mathbf{u}, t)$ . Dynamical time-variant systems  $\dot{\mathbf{x}}(t) = \mathbf{F}(\mathbf{x}, \mathbf{u}, t)$  apparently have more general appeal, because most industrial dynamical systems are inherently time-variant. For example, the parameters of components in an electrical circuit system often vary with time. On the other hand, in many applications such as aerospace, process control etc., the ever increasing performance demands and more stringent specifications over a wide range of operating conditions diminish the value of time-invariant models as good approximations of the actual plant. Time-variant systems, however, exhibit more complicated dynamics. It is clear that the approximation problem of continuous-time RNN to dynamical time-variant systems is an essential issue.

The main objective of this paper is to extend the approximation capability of continuous-time RNN [9] to the case of dynamical time-variant systems  $\dot{\mathbf{x}}(t) = \mathbf{F}(\mathbf{x}, \mathbf{u}, t)$ . This paper mathematically proves that the finite time trajectory of any dynamical time-variant systems can be approximated by the internal state of the output units of a continuous RNN to any degree of accuracy. The obtained results are very relevant to the approximation theory of NN and RNN.

Manuscript received June 19, 2002; revised January 31, 2005. This work was supported by the Research Council of the Hong Kong SAR Government under Project 9040601-570. This paper was recommended by Associate Editor A. Kuh.

X.-D. Li is with the Department of Manufacturing Engineering and Engineering Management, City University of Hong Kong, Hong Kong, and also with the Department of Electronic Engineering, Sun Yat-Sen University, Guangzhou, China (e-mail: 50006160@student.cityu.edu.hk).

J. K. L. Ho is with the Department of Manufacturing Engineering and Engineering Management, City University of Hong Kong, Hong Kong (e-mail: mejohnho@cityu.edu.hk).

T. W. S. Chow is with the Department of Electronic Engineering, City University of Hong Kong, Hong Kong (e-mail: eetchow@cityu.edu.hk).

Digital Object Identifier 10.1109/TCSII.2005.852006

## II. CONTINUOUS-TIME RECURRENT NEURAL NETWORKS

A continuous-time RNN is a complex nonlinear dynamical system described by a set of nonlinear differential equations with extensive connection weights. This can be generally expressed in the following form:

$$\dot{\mathbf{z}}(t) = -\alpha \mathbf{z}(t) + \mathbf{f}(\mathbf{W}_1 \mathbf{z}(t), \mathbf{W}_2 \mathbf{u}(t)) \quad (1)$$

where  $\mathbf{z} \in \mathbb{R}^L$  and  $\mathbf{u} \in \mathbb{R}^m$  are the neural state and the input vector, respectively,  $\mathbf{W}_1 \in \mathbb{R}^{L \times L}$ ,  $\mathbf{W}_2 \in \mathbb{R}^{L \times m}$  are the connection weight matrices associated with the neural state and the input vector, respectively. The parameter  $\alpha$  is a fixed constant and is chosen to lie in the open unit interval, i.e.,  $0 < \alpha < 1$ , and  $\mathbf{f} : \mathbb{R}^L \times \mathbb{R}^m \rightarrow \mathbb{R}^L$  is an appropriately chosen vector-valued nonlinear function. In this paper, we study the following continuous-time RNN:

$$\dot{\mathbf{z}}(t) = -\alpha \mathbf{z}(t) + \mathbf{W}_1 \sigma(\mathbf{z}(t)) + \mathbf{W}_2 \mathbf{u}(t) \quad (2)$$

where  $\sigma(\cdot)$  is a  $C^1$ -sigmoid nonlinear neural activation function (A  $C^k$ -class function is continuously differentiable up to order  $k$ ).

## III. APPROXIMATION THEOREM OF CONTINUOUS RECURRENT NEURAL NETWORKS TO DYNAMICAL TIME-VARIANT SYSTEMS

It has been previously shown that the state of the RNN output units (2) with an arbitrary positive small  $\alpha$  is capable of approximating the dynamics of a nonlinear time-invariant system to any degree [9]. Let  $\mathbf{x} = (x_1, \dots, x_n)'$  denotes a point in  $n$ -dimensional Euclidian space  $\mathbb{R}^n$ , where  $(\cdot)'$  denotes matrix transpose, and the Euclidian norm of  $x$  be defined by  $\|\mathbf{x}\|$ . The aforementioned result can thus be expressed in the following lemma.

*Lemma 1 [9]:* Let  $S \subset \mathbb{R}^n$  and  $U \subset \mathbb{R}^m$  be open sets,  $D_S \subset S$  and  $D_U \subset U$  be compact sets, and  $\mathbf{F} : S \times U \rightarrow \mathbb{R}^n$  be a  $C^1$ -class vector-valued function. For a continuous nonlinear system in the form of

$$\dot{\mathbf{x}}(t) = \mathbf{F}(\mathbf{x}(t), \mathbf{u}(t)), \quad \mathbf{x} \in S, \mathbf{u} \in U, t \in \mathbb{R} \quad (3)$$

with an initial state  $\mathbf{x}(0) \in D_S$ . Then, for an arbitrary  $\varepsilon > 0$ , there exists an integer  $N$  and an RNN in the form (2) with an appropriate initial state  $\mathbf{z}(0)$  and a small enough arbitrary  $\alpha > 0$  such that for any bounded input  $\mathbf{u} : \mathbb{R} \rightarrow D_U$

$$\max_{t \in [0, T]} \|\mathbf{x} - \mathbf{p}\| < \varepsilon, \quad 0 < T < \infty$$

holds, where  $\mathbf{z} = (\mathbf{p}, \mathbf{h})' \in \mathbb{R}^L$ ,  $\mathbf{p} = (p_1, \dots, p_n)' \in \mathbb{R}^n$ ,  $\mathbf{h} = (h_1, \dots, h_N)' \in \mathbb{R}^N$ ,  $\mathbf{p}$  is the internal state of the network output units,  $\mathbf{z}$  is the overall neural state, and  $\mathbf{h}$  is the hidden neural state.

In this lemma, the nonlinear dynamical system to be approximated by a continuous RNN is a time-invariant nonlinear system. For more complicated dynamical time-variant systems, Lemma 1 can be extended to obtain a more general approximation result as shown in the next theorem.

*Theorem 1:* Let  $S \subset \mathbb{R}^n$  and  $U \subset \mathbb{R}^m$  be open sets,  $D_S \subset S$ ,  $D_U \subset U$  be compact sets, and  $\mathbf{F} : S \times U \times \mathbb{R} \rightarrow \mathbb{R}^n$  be a

$C^1$ -class vector-valued function. For a continuous time-variant system of the form

$$\dot{\mathbf{x}}(t) = \mathbf{F}(\mathbf{x}(t), \mathbf{u}(t), t), \quad \mathbf{x}(t) \in S, \mathbf{u}(t) \in U, t \in \mathbb{R} \quad (4)$$

with an initial state  $\mathbf{x}(0) \in D_S$ . Then, for an arbitrary  $\varepsilon > 0$ , there exists an integer  $N$  and an RNN of the form (2) with an appropriate initial state  $\mathbf{z}(0)$  and a small enough arbitrary  $\alpha > 0$  such that for any bounded input  $\mathbf{u} : \mathbb{R} \rightarrow D_U$

$$\max_{t \in [0, T]} \|\mathbf{x} - \mathbf{p}\| < \varepsilon, \quad 0 < T < \infty$$

holds, where  $\mathbf{z} = (\mathbf{p}, \mathbf{h})' \in \mathbb{R}^L$ ,  $\mathbf{p} = (p_1, \dots, p_n)' \in \mathbb{R}^n$ ,  $\mathbf{h} = (h_1, \dots, h_N)' \in \mathbb{R}^N$ ,  $\mathbf{p}$  is the internal state of the network output units,  $\mathbf{z}$  is the overall neural state, and  $\mathbf{h}$  is the hidden neural state.

*Proof:* Let  $\bar{\mathbf{x}}(t) = (\mathbf{x}^{(t)}) \in \mathbb{R}^{n+1}$ , where  $\bar{\mathbf{x}}_{n+1}(t) = t$ . We extend the  $n$ -dimensional vector  $\mathbf{x}(t)$  in dynamical time-variant system (4) to  $(n+1)$ -dimensional vector  $\bar{\mathbf{x}}(t)$ . Then, the continuous dynamical time-variant system (4) may be represented as the following time-invariant form:

$$\dot{\bar{\mathbf{x}}}(t) = \bar{\mathbf{F}}(\bar{\mathbf{x}}(t), \mathbf{u}(t)), \quad \bar{\mathbf{x}} \in (S \times \mathbb{R}), \mathbf{u} \in U \quad (5)$$

where  $\bar{\mathbf{F}}(\bar{\mathbf{x}}(t), \mathbf{u}(t)) = [\mathbf{F}(\mathbf{x}(t), \mathbf{u}(t), \bar{\mathbf{x}}_{n+1}(t))]$ , and  $\bar{\mathbf{F}} : (S \times \mathbb{R}) \times U \rightarrow \mathbb{R}^{n+1}$  is also a  $C^1$ -class vector-valued function. The initial condition of system (5) is  $\bar{\mathbf{x}}(0) = [\mathbf{x}^{(0)}]$ . When  $\mathbf{x}(0) \in D_S$ , we have  $\bar{\mathbf{x}}(0) \in D_S \times [0, T]$ ,  $(0 < T < \infty)$ .  $D_S \times [0, T]$  is a compact subset of  $S \times \mathbb{R}$ , so the time-invariant system (5) satisfies the conditions of Lemma 1. Then, for an arbitrary  $\varepsilon > 0$ , there exists an integer  $N$  and an RNN of the form (2) with an appropriate initial state  $\mathbf{z}(0)$  and a sufficiently small arbitrary  $\alpha > 0$  such that, for any bounded input  $\mathbf{u} : \mathbb{R} \rightarrow D_U$

$$\max_{t \in [0, T]} \|\bar{\mathbf{x}}(t) - \bar{\mathbf{p}}(t)\| < \varepsilon \quad (6)$$

holds, where  $\mathbf{z} = (\bar{\mathbf{p}}, \bar{\mathbf{h}})' \in \mathbb{R}^L$ ,  $\bar{\mathbf{p}} \in \mathbb{R}^{n+1}$ ,  $\bar{\mathbf{h}} \in \mathbb{R}^{N-1}$ , and  $\bar{\mathbf{p}}$  is the internal state of the first  $(n+1)$  output units of the network. Let  $\bar{\mathbf{p}} = (\mathbf{p}, \bar{\mathbf{p}}_{n+1})'$ , then  $\mathbf{p} \in \mathbb{R}^n$  is the internal state of the first  $n$  output units of the network. From the definition of the Euclidian norm  $\|\cdot\|$ , we have

$$\|\mathbf{x}(t) - \mathbf{p}(t)\|^2 + (\bar{\mathbf{x}}_{n+1}(t) - \bar{\mathbf{p}}_{n+1}(t))^2 = \|\bar{\mathbf{x}}(t) - \bar{\mathbf{p}}(t)\|^2.$$

Therefore

$$\max_{t \in [0, T]} \|\mathbf{x}(t) - \mathbf{p}(t)\| \leq \max_{t \in [0, T]} \|\bar{\mathbf{x}}(t) - \bar{\mathbf{p}}(t)\| < \varepsilon. \quad (7)$$

Theorem 1 is proved.

Theorem 1 illustrates that the finite-time trajectories of a given time-variant continuous system can be approximated by the first  $n$  internal state out of the  $n+N$  neural units in an RNN (2). It clearly has more general applications compared to the results previously obtained in [9]. Fig. 1 shows a diagram of an RNN (2) with 4 neural units and 1 input. According to Theorem 1, we set the first two units as output units of the RNN and the other as hidden units if a continuous time-variant system (4) with one input and two outputs is approximated by the RNN

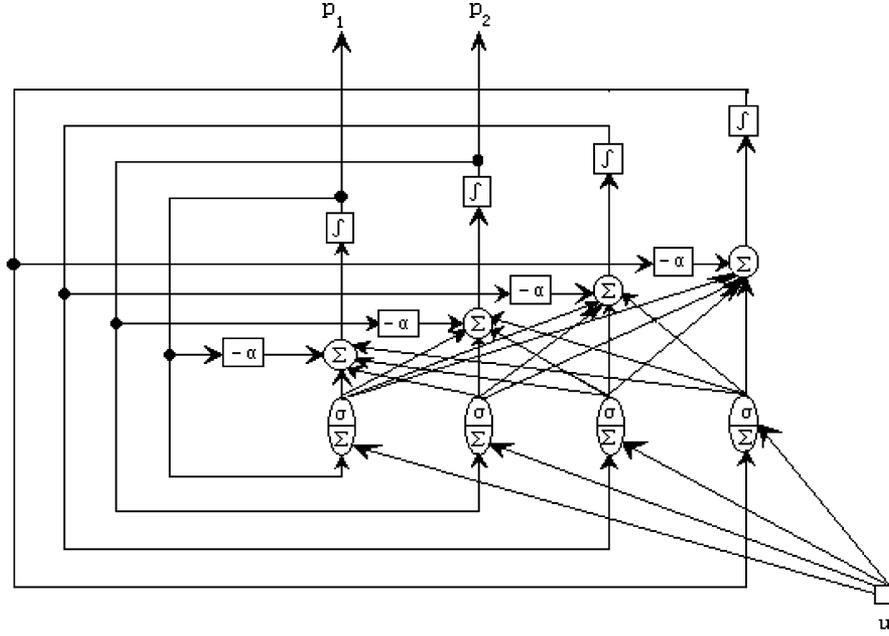


Fig. 1. Diagram of RNN (2) with 4 neural units and 1 external input, where  $\mathbf{p}_1, \mathbf{p}_2$  represent the output units' internal states.

of Fig. 1. The internal states  $\mathbf{p}_1$  and  $\mathbf{p}_2$  of the RNN output units are used to approximate the trajectories of a continuous time-variant system.

Theorem 1 can be extended to the approximation of continuous time-variant input-output systems by RNN in the following corollary 1.

*Corollary 1:* Let  $S \subset \mathfrak{R}^n, U \subset \mathfrak{R}^m$ , be open sets,  $D_S \subset S, D_U \subset U$  be compact sets, and  $\mathbf{F} : S \times U \times \mathfrak{R} \rightarrow \mathfrak{R}^n$  be a  $C^1$ -class vector-valued function,  $\mathbf{G} : S \times \mathfrak{R} \rightarrow \mathfrak{R}^r$  be a  $C^2$ -class reversible vector-valued function for each  $t \in \mathfrak{R}$ . For a continuous time-variant input-output system of the form

$$\dot{\mathbf{x}}(t) = \mathbf{F}(\mathbf{x}(t), \mathbf{u}(t), t), \quad \mathbf{x}(t) \in S, \mathbf{u}(t) \in U, t \in \mathfrak{R} \quad (8a)$$

$$\mathbf{y}(t) = \mathbf{G}(\mathbf{x}(t), t) \quad (8b)$$

with an initial state  $\mathbf{x}(0) \in D_S$ , then, for an arbitrary  $\varepsilon > 0$ , there exists an integer  $N$  and an RNN of the form (2) with an appropriate initial state  $\mathbf{z}(0)$  and a small enough arbitrary  $\alpha > 0$  such that for any bounded input  $\mathbf{u} : \mathfrak{R} \rightarrow D_U$

$$\max_{t \in [0, T]} \|\mathbf{y} - \mathbf{p}\| < \varepsilon, \quad 0 < T < \infty$$

holds, where  $\mathbf{z} = (\mathbf{p}, \mathbf{h})' \in \mathfrak{R}^L, \mathbf{p} \in \mathfrak{R}^r, \mathbf{h} \in \mathfrak{R}^N$ , and  $\mathbf{p}$  is the internal state of the network output units,  $\mathbf{z}$  is the overall neural state,  $\mathbf{h}$  is the hidden neural state.

*Proof:* In the nonlinear system (8), as the vector-valued function  $\mathbf{F}$  is a  $C^1$ -class mapping  $S \times U \times \mathfrak{R} \rightarrow \mathfrak{R}^n$ , and the vector-valued function  $\mathbf{G}$  is a  $C^2$ -class mapping  $S \times \mathfrak{R} \rightarrow \mathfrak{R}^r$ , the set of output trajectories of the nonlinear dynamical system (8) defined as

$$Y = \{\mathbf{y}(t) \in \mathfrak{R}^r : \mathbf{x}(t) \in S, \mathbf{u}(t) \in U, t \in \mathfrak{R}\} \quad (9)$$

forms an open subset of  $\mathfrak{R}^r$ , and the set of initial output

$$D_Y = \{\mathbf{y}(0) \in \mathfrak{R}^r : \mathbf{x}(0) \in D_S\} \quad (10)$$

forms a compact subset of  $\mathfrak{R}^r$ .

From the nonlinear system (8), we have

$$\dot{\mathbf{y}}(t) = \frac{\partial \mathbf{G}}{\partial \mathbf{x}} \mathbf{F}(\mathbf{x}, \mathbf{u}, t) + \frac{\partial \mathbf{G}}{\partial t}. \quad (11)$$

Let  $\mathbf{H}(\mathbf{x}, \mathbf{u}, t) = (\partial \mathbf{G} / \partial \mathbf{x}) \cdot \mathbf{F}(\mathbf{x}, \mathbf{u}, t) + (\partial \mathbf{G} / \partial t)$ , then  $\mathbf{H} : S \times U \times \mathfrak{R} \rightarrow \mathfrak{R}^r$  is a  $C^1$ -class vector-valued function, and the nonlinear system (8) can be represented as

$$\dot{\mathbf{y}}(t) = \mathbf{H}(\mathbf{x}(t), \mathbf{u}(t), t), \mathbf{x}(t) \in S, \mathbf{u}(t) \in U, t \in \mathfrak{R} \quad (12)$$

Also,  $\mathbf{G} : S \times \mathfrak{R} \rightarrow \mathfrak{R}^r$  is a reversible vector-valued function for each  $t \in \mathfrak{R}$ , so  $\mathbf{x}(t) = \mathbf{G}^{-1}(\mathbf{y}(t), t)$ , and the nonlinear continuous system (12) can be further written as

$$\dot{\mathbf{y}}(t) = \tilde{\mathbf{H}}(\mathbf{y}(t), \mathbf{u}(t), t), \mathbf{y}(t) \in Y, \mathbf{u}(t) \in U, t \in \mathfrak{R} \quad (13)$$

with an initial state  $\mathbf{y}(0) \in D_Y$ , where  $\tilde{\mathbf{H}} : Y \times U \times \mathfrak{R} \rightarrow \mathfrak{R}^r$  is a  $C^1$ -class vector-valued function.

Applying Theorem 1 to the nonlinear system (13), for an arbitrary  $\varepsilon > 0$ , there exists an integer  $N$  and an RNN of the form (2) with an appropriate initial state  $\mathbf{z}(0)$  and a small enough arbitrary  $\alpha > 0$  such that for any bounded input  $\mathbf{u} : \mathfrak{R} \rightarrow D_U$

$$\max_{t \in [0, T]} \|\mathbf{y} - \mathbf{p}\| < \varepsilon, \quad 0 < T < \infty$$

holds, where  $\mathbf{z} = (\mathbf{p}, \mathbf{h})' \in \mathfrak{R}^L, \mathbf{p} \in \mathfrak{R}^r, \mathbf{h} \in \mathfrak{R}^N$ , and  $\mathbf{p}$  is the internal state of the network output units.

Corollary 1 is proved.

Based on Theorem 1, we next show that several special forms of dynamical time-variant systems or trajectories can be approximately realized by the internal state of the output units of an RNN or a simple RNN.

*Corollary 2:* Let  $S \subset \mathbb{R}^n$  and  $U \subset \mathbb{R}^m$  be open sets,  $D_S \subset S$ ,  $D_U \subset U$  be compact sets, and  $\mathbf{F}_1 : \mathbb{R} \times U \rightarrow \mathbb{R}^n$ ,  $\mathbf{F}_2 : S \times \mathbb{R} \rightarrow \mathbb{R}^n$  and  $\mathbf{F}_3 : \mathbb{R} \rightarrow \mathbb{R}^n$  be  $C^1$ -class vector-valued functions. For each dynamical time-variant system of the following forms

$$\dot{\mathbf{x}}(t) = \mathbf{F}_1(t, \mathbf{u}(t)), \quad t \in \mathbb{R}, \mathbf{u} \in U \quad (14)$$

$$\dot{\mathbf{x}}(t) = \mathbf{F}_2(\mathbf{x}(t), t), \quad \mathbf{x}(t) \in S, t \in \mathbb{R} \quad (15)$$

$$\dot{\mathbf{x}}(t) = \mathbf{F}_3(t), \quad t \in \mathbb{R} \quad (16)$$

with an initial state  $\mathbf{x}(0) \in D_S$ , then, for an arbitrary  $\varepsilon > 0$ , there exists an integer  $N$  and an RNN of the form (2) with an appropriate initial state  $\mathbf{z}(0)$  and a small enough arbitrary  $\alpha > 0$  such that for any bounded input  $\mathbf{u} : \mathbb{R} \rightarrow D_U$

$$\max_{t \in [0, T]} \|\mathbf{x} - \mathbf{p}\| < \varepsilon, \quad 0 < T < \infty$$

holds, where  $\mathbf{z} = (\mathbf{p}, \mathbf{h})' \in \mathbb{R}^L$ ,  $\mathbf{p} \in \mathbb{R}^n$ ,  $\mathbf{h} \in \mathbb{R}^N$ , and  $\mathbf{p}$  is the internal state of the network's output units,  $\mathbf{z}$  is the overall neural state,  $\mathbf{h}$  is the hidden neural state.

*Remark:* Corollary 2 is a direct result of Theorem 1. For approximating the special forms of dynamical time-variant systems or trajectories (15) and (16) with  $\mathbf{u}(t) \equiv 0$ , the RNN (2) is practically simplified as the following form:

$$\dot{\mathbf{z}}(t) = -\alpha \mathbf{z}(t) + W\sigma(\mathbf{z}(t)) \quad (17)$$

which resembles a simple form of Hopfield-type continuous RNN.

Correspondingly, if (8a) has certain special forms, we can also derive useful results on continuous time-variant input-output systems. From the above described theorem and corollaries, the approximation takes place only on a finite closed interval. There is no information regarding the exact number of hidden neural units required to achieve such an approximation. However, from [11], the storage capacity of a fully connected neural network must be at least the same size as the net. When the approximation interval becomes longer, the RNN memory requirements for tracking the trajectory of the given dynamical system increases accordingly. Therefore, generally speaking, the longer the approximation interval, the larger the number of neural units needed.

#### IV. SIMULATION RESULTS

*Example 1:* In order to further verify our approximation theorem, a dynamical time-variant second-order system

$$\ddot{x}(t) - t \cdot \dot{x}(t) - \sin(x(t)) - u(t) = 0 \quad (18)$$

with  $x(0) = \dot{x}(0) = 0$  was approximated in a finite time interval  $[0, 1]$  by the internal state of the RNN output units (2). The training of the RNN (2) used for the approximation of the dynamical time-variant system (18) was carried out by means of a kind of hybrid method that combines Genetic Algorithm (GA) and another gradient-descent based method called recurrent backpropagation algorithm by Pearlmutter [12]. The option

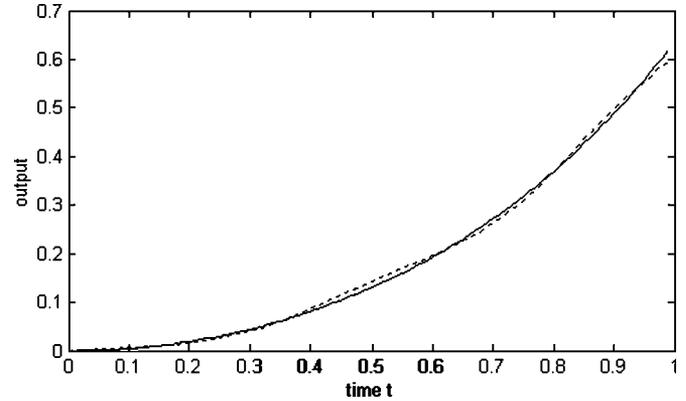


Fig. 2. Outputs of the time-variant system (18) and RNN (2) when excited by a step input. [Solid line: output of the time-variant system (18); dashed line: modeling output of the RNN (2).]

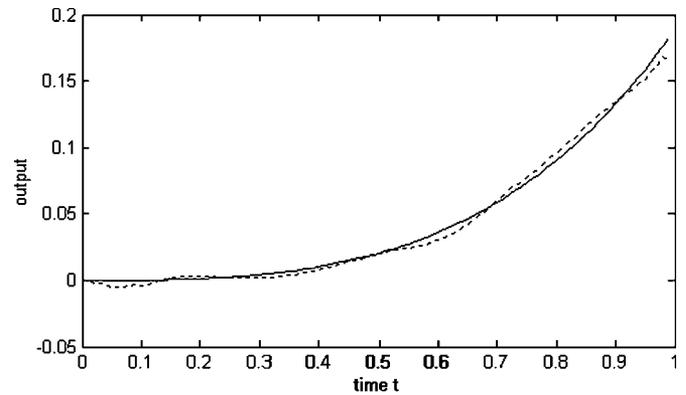


Fig. 3. Outputs of the time-variant system (18) and RNN (2) when excited by a sinusoidal input. [Solid line: output of the time-variant system (18); dashed line: modeling output of the RNN (2).]

of using only recurrent backpropagation algorithm was deemed not viable. The implementation of gradient-descent training algorithms is computationally expensive and does not provide an exhaustive search of the error surface. In our training scheme of RNN (2), after selecting a set of suitable training data, the GA was first used to obtain a rough initial point, then the recurrent backpropagation algorithm was used to bring the total square training error over a finite time interval to minimum.

In the results presented, the parameter  $\alpha$  of RNN (2) was set to 0.01. The number of neurons was set to 5, using stepwise forward selection with a satisfactory index. The *sigmoid* non-linear function  $\tanh(\cdot)$  was chosen for  $\sigma(\cdot)$ . For simulation convenience, both the time-variant system (18) and RNN (2) were discretized with a sampling time interval 0.01. Once the network structure was fixed, the GA-based hybrid method was used to determine the network weights. The initial weights of RNN (2) were random gaussian values. The first 2 RNN output neurons (2) were initialized to the initial values of the time-variant system (18) while the rest of the RNN neuron states (2) were initialized to 0. The number of the initial population involved in the GA-based algorithm was fixed at 140. This is a reasonable number that provides ample variety of individuals and low computational cost for each generation. The mutation and crossover probabilities were fixed at 0.001 and 0.8, respectively. When the GA-based hybrid algorithm converged, the trained weights of RNN (2) were determined. Figs. 2 and 3 present the modeling

outputs when system (18) and RNN (2) are excited by a step input and a sine input, respectively. Their modeling error indices (total square error) are  $5.4 \times 10^{-3}$  and  $2.3 \times 10^{-3}$ , respectively. This example shows the nonlinear time-variant system (18) can be approximated by an RNN (2) with 5 neural units to a very high degree of accuracy.

## V. CONCLUSION

This paper proves that dynamical time-variant systems, which have considerable general applications, can be approximated in a finite time interval to any accuracy level by using an RNN. As discussed, the special forms of continuous dynamical time-variant systems or trajectories can be approximately realized in a finite time interval by the internal state of the output units of a simple RNN. In implementing the above approximation, the connection weight matrices of the RNN must be determined through an adaptive learning process.

## ACKNOWLEDGMENT

The authors would like to thank Associate Editor and anonymous reviewers for their valuable suggestions to this paper.

## REFERENCES

[1] J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," *Proc. Nat. Acad. Sci. USA*, vol. 79, pp. 2554–2558, 1982.

- [2] F. J. Pineda, "Dynamics and architecture for neural computation," *J. Complexity*, vol. 4, pp. 216–245, 1988.
- [3] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Netw.*, vol. 2, no. 5, pp. 359–366, 1989.
- [4] K. Funahashi, "On the approximation realization of continuous mappings by neural networks," *Neural Netw.*, vol. 2, no. 3, pp. 183–191, 1989.
- [5] G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Math. Control, Signal, Syst.*, vol. 2, pp. 303–314, 1989.
- [6] K. Li, "Approximation theory and recurrent networks," in *Proc. IJCNN'91*, vol. 2, 1991, pp. 266–271.
- [7] L. Jin, P. N. Nikiforuk, and M. M. Gupta, "Approximation of discrete-time state-space trajectories using dynamic recurrent neural networks," *IEEE Trans. Autom. Control*, vol. 40, no. 7, pp. 1266–1270, Jul. 1995.
- [8] K. Funahashi and Y. Nakamura, "Approximation of dynamical systems by continuous time recurrent neural networks," *Neural Netw.*, vol. 6, pp. 801–806, 1993.
- [9] T. W. S. Chow and X.-D. Li, "Modeling of continuous time dynamical systems with input by recurrent neural networks," *IEEE Trans. Circuit Syst. I, Fundam. Theory Appl.*, vol. 47, no. 4, pp. 575–578, Apr. 2000.
- [10] C. Kambhampati, F. Garces, and K. Warwick, "Approximation of nonautonomous dynamical systems by continuous time recurrent neural networks," in *Proc. IEEE-INNS-ENNS Int. Joint Conf.*, vol. 1, 2000, pp. 64–69.
- [11] L. K. Li, "Capacity of recurrent networks," in *Proc. IJCNN'91*, vol. 3, Singapore, 1991, pp. 1330–1335.
- [12] B. A. Pearlmutter, "Gradient calculations for dynamic recurrent neural networks: a survey," *IEEE Trans. Neural Netw.*, vol. 6, no. 5, pp. 1212–1228, Sep. 1995.
- [13] T. W. S. Chow and Y. Fang, "A recurrent neural network based real-time learning control strategy applying to nonlinear systems with unknown dynamics," *IEEE Trans. Ind. Electron.*, vol. 45, no. 1, pp. 151–161, Feb. 1998.