Neuro-Fuzzy Nonlinear Dynamical System Approximations using High Order Neural Networks

Manolis A. Christodoulou*
Department of Electronic and Computer Engineering
Technical University of Crete
Chania, Crete, GREECE 73100
manolis@ece.tuc.gr

Dimitris C. Theodoridis**
Department of Electrical and Computer Engineering
Democritus University of Thrace
Xanthi, GREECE, 67100 GR
dtheodo@ee.duth.gr

Yiannis A. Boutalis**
Department of Electrical and Computer Engineering
Democritus University of Thrace
Xanthi, GREECE, 67100 GR
ybout@ee.duth.gr

***Abstract*— A new definition of Adaptive Neuro Fuzzy Systems is presented in this paper for the identification of unknown nonlinear dynamical systems. The proposed scheme uses the concept of Adaptive Fuzzy Systems (AFS) operating in conjunction with High Order Neural Network Functions (F-HONNFs). Since the plant is considered unknown, we first propose its approximation by a special form of an adaptive fuzzy system and in the sequel the fuzzy rules are approximated by appropriate HONNFs. Thus the identification scheme leads up to a Recurrent High Order Neural Network, which however takes into account the fuzzy output partitions of the initial AFS. The proposed scheme does not require a-priori experts' information on the number and type of input variable membership functions making it less vulnerable to initial design assumptions. Weight updating laws for the involved HONNFs are provided, which guarantee that the identification error reaches zero exponentially fast. Simulations illustrate the potency of the method and comparisons with well known benchmarks are given.**

I. INTRODUCTION

It is well known that general nonlinear systems are expressed by nonlinear ordinary differential equations, which may take the following expression

$$\dot{x} = f(x, u) \tag{1}$$

If we know the mathematical model of the system completely, then, provided then exists adequate theory, we are able to control it. In most practical application cases though, the exact mathematical model of the plant, especially when it is highly nonlinear, large scale and complex, it is seldom is known. In those cases we should apply known identification schemes in order to find a suitable model and then control the plant in indirect cases, or, simultaneously identify and control the plant on-line in cases that we apply direct adaptive control algorithms.

It is well known form the scientific literature, that neural and fuzzy systems are universal approximators, [1], [2], [3]. They can approximate a large variety of nonlinear dynamical systems to any desired accuracy provided that sufficient hidden neurons and training data or fuzzy rules are available. Recently, the combination of these two different technologies has given rise to fuzzy neural or neuro fuzzy approaches, that are intended to capture the advantages of both fuzzy logic and neural networks. Numerous works have shown the viability of this approach for system modeling [4] - [12].

The neural and fuzzy approaches are most of the time equivalent, differing between each other mainly in the structure of the approximator chosen. Indeed, in order to bridge the gap between the neural and fuzzy approaches several researchers introduce adaptive schemes using a class of parameterized functions that include both neural networks and fuzzy systems [6] - [12]. Regarding the approximator structure, linear in the parameters approximators are used in [10], [13], and nonlinear in [14], [15], [16].

Adaptive control theory has been an active area of research over the past few years [13]-[30]. The identification procedure is an essential part in any control procedure. In the neuro or neuro fuzzy adaptive control two main approaches are followed. In the indirect adaptive control schemes [13] - [19], first the dynamics of the system are identified and then a control input is generated according to the certainty equivalence principle. In the direct adaptive control schemes [20] - [26] the controller is directly estimated and the control input is generated to guarantee stability without knowledge of the system dynamics. Also, many researchers focus on robust adaptive control that guarantees signal boundness in the presence of modeling errors and bounded disturbances [27]. In [28] both direct and indirect approaches are presented, while in [29],[30] a combined direct and indirect control scheme is used.

Recently [31], [32], high order neural network function approximators (HONNFs) have been proposed for the identification of nonlinear dynamical systems of the form (1), approximated by a Fuzzy Dynamical System. This approximation depends on the fact that fuzzy rules could be identified with the help of HONNFs.

In this paper HONNFs are also used for the neuro fuzzy identification of unknown nonlinear dynamical systems. In fuzzy or neuro-fuzzy approaches the identification phase usually consists of two categories: structure identification and parameter identification. Structure identification involves finding the main input variables out of all possible, specifying the membership functions, the partition of the input space and determining the number of fuzzy rules which is often based on a substantial amount of heuristic observation to express proper strategy's knowledge. Most of structure identification methods are based on data clustering, such as fuzzy C-means clustering [9], mountain clustering [11], and subtractive clustering [12]. These approaches require that all input-output data are ready before we start to identify the plant. So these structure identification approaches are off-line.

In the proposed approach structure identification is also made off-line based either on human expertise or on gathered data. However, the required a-priori information obtained by linguistic information or data is very limited. The only required information is an estimate of the centers of the output fuzzy membership functions. Information on the input variable membership functions and on the underlying fuzzy rules is not necessary because this is automatically estimated by the HONNFs. This way the proposed method is less vulnerable to initial design assumptions. The parameter identification is then easily addressed by HONNFs, based on the linguistic information regarding the structural identification of the output part and from the numerical data obtained from the actual system to be modelled. So, the parameters of identification model are updated on - line in such a way that the error between the actual system output and the model output reaches zero exponentially fast.

We consider that the nonlinear system is affine in the control and could be approximated with the help of two independent fuzzy subsystems. Every fuzzy subsystem is approximated from a family of HONNFs, each one being related with a group of fuzzy rules. Weight updating laws are given and we prove that when the structural identification is appropriate then the error converges very fast to zero.

The paper is organized as follows. Section II presents preliminaries related to the concept of adaptive fuzzy systems (AFS) and the terminology used in the remaining paper, while Section III reports on the ability of HONNFs to act as fuzzy rule approximators. The new neuro fuzzy representation of affine in the control dynamical systems is introduced in Section IV, where the associated weight adaptation laws are given. Simulation results on the identification of well known benchmark problems are given in Section V and the performance of the proposed scheme is compared to another well known approach of the literature. Finally, Section VI concludes the work.

II. Preliminaries

In this section we briefly present the notion of adaptive fuzzy systems and their conventional representation. We are also introducing the representation of fuzzy systems using the rule firing indicator functions (RFIF), simply called indicator functions (IF), which is used for the development of the proposed method.

A. Adaptive Fuzzy Systems

The performance, complexity, and adaptive law of an adaptive fuzzy system representation can be quite different depending upon the type of the fuzzy system (Mamdani or Takagi-Sugeno). It also depends upon whether the representations is linear or nonlinear in its adjustable parameters. Suppose that the adaptive fuzzy system is intended to approximate the nonlinear function $f(x)$. In the mamdani type, linear in the parameters form, the following fuzzy logic representation is used [2],[3]:

$$f(x) = \sum_{l=1}^{M} \theta_l \xi_l(x) = \theta^T \xi(x) \tag{2}$$

where M is the number of fuzzy rules, $\theta = (\theta_1, ..., \theta_M)^T$, $\xi(x) = (\xi_1(x), ..., \xi_M(x))^T$ and $\xi_l(x)$ is the fuzzy basis function defined by

$$\xi_l(x) = \frac{\prod_{i=1}^{n} \mu_{F_i^l}(x_i)}{\sum_{l=1}^{M} \prod_{i=1}^{n} \mu_{F_i^l}(x_i)} \tag{3}$$

θ_l are adjustable parameters, and $\mu_{F_i^l}$ are given membership functions of the input variables (can be Gaussian, triangular, or any other type of membership functions).

In Tagaki-Sugeno formulation $f(x)$ is given by

$$f(x) = \sum_{l=1}^{M} g_l(x) \xi_l(x) \tag{4}$$

where $g_l(x) = a_{l,0} + a_{l,1}x_1 + \ldots + a_{l,n}x_n$, with $x_i, i = 1 \ldots n$ being the elements of vector x and $\xi_l(x)$ being defined in (3). According to [3], (4) can also be written in the linear to the parameters form, where the adjustable parameters are all $a_{l,i}, l = 1 \ldots M, i = 1 \ldots n$.

From the above definitions it is apparent in both, Mamdani and Tagaki-Sugeno forms that the success of the adaptive fuzzy system representations in approximating the nonlinear function $f(x)$ depends on the careful selection of the fuzzy partitions of input and output variables. Also, the selected type of the membership functions and the proper number of fuzzy rules contribute to the success of the adaptive fuzzy system. This way, any adaptive fuzzy or neuro-fuzzy approach, following a linear in the adjustable parameters formulation becomes vulnerable to initial design assumptions related to the fuzzy partitions and the membership functions chosen. In this paper this drawback is largely overcome by using the concept of rule indicator functions, which are in the sequel approximated by High order Neural Network function approximators (HONNFs). This way there is not any need for initial design assumptions related to the membership values and the fuzzy partitions of the if part.

B. Fuzzy system description using rule indicator functions

Let us consider the system with input space $u \subset R^m$ and state - space $x \subset R^n$, with its i/o relation being governed by the following equation

$$z(k) = f(x(k), u(k)) \tag{5}$$

where $f(\cdot)$ is a continuous function and k denotes the temporal variable. In case the system is dynamic the above equation could be replaced by the following differential equation

$$\dot{x}(k) = f(x(k), u(k)) \tag{6}$$

By setting $y(k) = [x(k), u(k)]$, Eq. (5) may be rewritten as follows

$$z(k) = f\left(y(k)\right) \tag{7}$$

with $y \subset R^{m+n}$

In case f in (7) is unknown we may wish to approximate it by using a fuzzy representation. In this case both $y(k) = [x(k), u(k)]$ and $z(k)$ are initially replaced by fuzzy linguistic variables. Experts or data depended techniques may determine the form of the membership functions of the fuzzy variables and fuzzy rules will determine the fuzzy relations between $y(k)$ and $u(k)$. Sensor input data, possibly noisy and imprecise, enter the fuzzy system, are fuzzified, are processed by the fuzzy rules and the fuzzy implication engine and are in the sequel defuzzified to produce the estimated $z(k)$ [2], [3]. We assume here that a Mamdani type fuzzy system is used.

Let now $\Omega^{l_1,l_2,\ldots,l_n}_{j_1,j_2,\ldots,j_{n+m}}$ be defined as the subset of (x, u) pairs, belonging to the $(j_1, j_2, ..., j_{n+m})^{th}$ input fuzzy patch and pointing - through the vector field $f(\cdot)$ - to the subset of $z(k)$, which belong to the $(j_1, j_2, ..., j_{n+m})^{th}$ output fuzzy patch. In other words, $\Omega^{l_1,l_2,\ldots,l_n}_{j_1,j_2,\ldots,j_{n+m}}$ contains input value pairs that are associated through a fuzzy rule with output values.

In order to present the lemma of Section III, we define the Indicator function (IF) $I^{l_1,l_2,\ldots,l_n}_{j_1,j_2,\ldots,j_{n+m}}$ of the subset $\Omega^{l_1,l_2,\ldots,l_n}_{j_1,j_2,\ldots,j_{n+m}}$, that is,

$$I^{l_1,\ldots,l_n}_{j_1,\ldots,j_{n+m}}(x(k), u(k)) = \begin{cases} \alpha & if\,(x(k), u(k)) \in \Omega^{l_1,\ldots,l_n}_{j_1,\ldots,j_{n+m}} \\ 0 & otherwise \end{cases} \tag{8}$$

where α denotes the firing strength of the rule.

Define now the following system

$$z(k) = \sum \bar{z}^{l_1,\ldots,l_n}_{j_1,\ldots,j_{n+m}} \times I^{l_1,\ldots,l_n}_{j_1,\ldots,j_{n+m}}(x(k), u(k)) \tag{9}$$

Where $\bar{z}^{l_1,\ldots,l_n}_{j_1,\ldots,j_{n+m}} \in R^n$ be any constant vector consisting of the centers of the membership functions of each output variable z_i and $I^{l_1,\ldots,l_n}_{j_1,\ldots,j_{n+m}}(x(k), u(k))$ is the RFIF. Then, according to [31], [32] the system in (9) is a generator for the fuzzy system (FS).

It is obvious that Eq. (9) can be also valid for dynamic systems. In its dynamical form it becomes

$$\dot{x}(k) = \sum \bar{x}^{l_1,\ldots,l_n}_{j_1,\ldots,j_{n+m}} \times I^{l_1,\ldots,l_n}_{j_1,\ldots,j_{n+m}}(x(k), u(k)) \tag{10}$$

Where $\bar{x}^{l_1,\ldots,l_n}_{j_1,\ldots,j_{n+m}} \in R^n$ be again any constant vector consisting of the centers of fuzzy partitions of every variable x_i and $I^{l_1,\ldots,l_n}_{j_1,\ldots,j_{n+m}}(x(k), u(k))$ is the IF.

III. The HONNF's as fuzzy rule approximators

The main idea in presenting the main result of this section lies on the fact that functions of high order neurons are capable of approximating discontinuous functions; thus, we use high order neural network functions in order to approximate the indicator functions $I^{l_1,\ldots,l_n}_{j_1,\ldots,j_{n+m}}$. However, in order the approximation problem to make sense the space $\mathcal{Y} := \mathcal{X} \times \mathcal{U}$ must be compact. Thus, our first assumption is the following:

(A.1) $\mathcal{Y} := \mathcal{X} \times \mathcal{U}$ is a compact set.

Notice that since $\mathcal{Y} \subset \Re^{n+m}$ the above assumption is identical to the assumption that it is closed and bounded. Also, it is noted that even if $\mathcal{Y}$ is not compact we may assume that there is a time instant T such that $(x(k), u(k))$ remain in a compact subset of $\mathcal{Y}$ for all $t < T$; i.e. if $\mathcal{Y}_T := \{(x(k), u(k)) \in \mathcal{Y}, k < T\}$ We may replace assumption (A.1) by the following assumption

(A.2) $\mathcal{Y}_T$ is a compact set.

It is worth noticing, that while assumption (A.1) requires the system in Eq. (6) solutions to be bounded for all $u^t \in U$ and $x^0 \in X$, assumption (A.2) requires the system in Eq. (6) solutions to be bounded for a finite time period; thus, assumption (A.1) requires the system in Eq. (6) to be BIBS stable while assumption (A.2) is valid for systems that are not BIBS stable and, even more, for unstable systems and systems with finite escape times.

We are now ready to show that high order neural network functions are capable of approximating the indicator functions $I^{l_1,\ldots,l_n}_{j_1,\ldots,j_{n+m}}$ Let us define the following high order neural network functions (HONNFs).

$$N(x, u; w, L) = \sum_{k=1}^{L} w_k \prod_{j \in I_k} \Phi_j^{d_j(k)} \tag{11}$$

Where $\{I_1, I_2, ..., I_L\}$ is a collection of L not-ordered subsets of $\{1, 2, ..., m+n\}$, $d_j(k)$ are non-negative integers, Φ_j are sigmoid functions of the state or the input and $w := [w_1 \cdots w_L]^T$ are the HONNF weights. Eq. (11) can also be written

$$N(x, u; w, L) = \sum_{k=1}^{L} w_k s_k(x, u) \tag{12}$$

Where $s_k(x, u)$ are high order terms of sigmoid functions of the state and/or input.

The next lemma [31] states that a HONNF of the form in Eq. (12) can approximate the indicator function $I^{l_1,\ldots,l_n}_{j_1,\ldots,j_{n+m}}$.

Lemma 1: Consider the indicator function $I^{l_1,\ldots,l_n}_{j_1,\ldots,j_{n+m}}$ and the family of the HONNFs $N(x, u; w, L)$. Then for any $\epsilon > 0$ there is a vector of weights $w^{j_1,\ldots,j_{n+m};l_1,\ldots,l_n}$ and a number of $L^{j_1,\ldots,j_{n+m};l_1,\ldots,l_n}$ high order connections such that

$$\sup_{(x,u)\in\bar{\mathcal{Y}}} \{I^{l_1,\ldots,l_n}_{j_1,\ldots,j_{n+m}}(x, u) - \\ -N(x, u; w^{j_1,\ldots,j_{n+m};l_1,\ldots,l_n}, L^{j_1,\ldots,j_{n+m};l_1,\ldots,l_n})\} < \varepsilon$$

where $\bar{\mathcal{Y}} \equiv \mathcal{Y}$ if assumption (A.1) is valid and $\bar{\mathcal{Y}}_T \equiv \mathcal{Y}$ if assumption (A.2) is valid.

IV. THE PROPOSED IDENTIFICATION SCHEME

We consider affine in the control, nonlinear dynamical systems of the form

$$\dot{x} = f(x) + G(x) \cdot u \tag{13}$$

where the state $x \in R^n$ is assumed to be completely measured, the control u is in R^n, f is an unknown smooth vector field called the drift term and G is a matrix with columns the unknown smooth controlled vector fields g_i, $i = 1, 2, ..., n$ and $G = [g_1, g_2, \ldots, g_n]$. The above class of continuous-time nonlinear systems are called affine, because in (13) the control input appears linear with respect to G. The main reason for considering this class of nonlinear systems is that most of the systems encountered in engineering, are by nature or design, affine. Furthermore, we note that non affine systems of the form given in (1) can be converted into affine, by passing the input through integrators, a procedure known as dynamic extension. The following mild assumptions are also imposed on (13), to guarantee the existence and uniqueness of solution for any finite initial condition and $u \in U$.

Proposition 1: Given a class U of admissible inputs, then for any $u \in U$ and any finite initial condition, the state trajectories are uniformly bounded for any finite $T > 0$. Hence, $|x(T)| < \infty$.

Proposition 2: The vector fields f, g_i, $i = 1, 2, ..., n$ are continuous with respect to their arguments and satisfy a local Lipchitz condition so that the solution $x(t)$ of (13) is unique for any finite initial condition and $u \in U$.

We are using an affine in the control fuzzy dynamical system, which approximates the system in (13) and uses two fuzzy subsystem blocks for the description of $f(x)$ and $G(x)$ as follows

$$f(\chi) = A\chi + \sum \bar{f}^{l_1,\ldots,l_n}_{j_1,\ldots,j_n} \times I^{l_1,\ldots,l_n}_{j_1,\ldots,j_n}(\chi) \tag{14}$$

$$g_i(\chi) = \sum (\bar{g}_i)^{l_1,\ldots,l_n}_{j_1,\ldots,j_n} \times {I_1}^{l_1,\ldots,l_n}_{j_1,\ldots,j_n}(\chi) \tag{15}$$

where the summation is carried out over the number of all available fuzzy rules, I, I_1 are appropriate fuzzy rule indicator functions and the meaning of indices $\bullet^{l_1,\ldots,l_n}_{j_1,\ldots,j_n}$ has already been described in Section II-B.

According to Lemma 1, every indicator function can be approximated with the help of a suitable HONNF. Therefore, every I, I_1 can be replaced with a corresponding HONNF as follows

$$f(\chi) = A\chi + \sum \bar{f}^{l_1,\ldots,l_n}_{j_1,\ldots,j_n} \times N^{l_1,\ldots,l_n}_{j_1,\ldots,j_n}(\chi) \tag{16}$$

$$\bar{g}_i(\chi) = \sum (\bar{g}_i)^{l_1,\ldots,l_n}_{j_1,\ldots,j_n} \times {N_1}^{l_1,\ldots,l_n}_{j_1,\ldots,j_n}(\chi) \tag{17}$$

where N, N_1 are appropriate HONNFs.

In order to simplify the model structure, since some rules result to the same output partition, we could replace the NNs associated to the rules having the same output with one NN and therefore the summations in (16),(17) are carried out over the number of the corresponding output partitions. Therefore, the affine in the control fuzzy dynamical system in (14), (15) is replaced by the following equivalent affine Recurrent High Order Neural Network (RHONN), which depends on the centers of the fuzzy output partitions $\bar{f}_l$ and $\bar{g}_{i,l}$

$$\dot{\hat{\chi}} = A\hat{\chi} + \sum_{l=1}^{Npf} \bar{f}_l \times N_l(\chi) + \sum_{i=1}^{n} \left(\sum_{l=1}^{Npg_i} (\bar{g}_i)_l \times N_{1l}(\chi) \right) u_i \tag{18}$$

Or in a more compact form

$$\dot{\hat{\chi}} = A\hat{\chi} + XWS(\chi) + X_1 W_1 S_1(\chi) u \tag{19}$$

Where A is a $n \times n$ stable matrix which for simplicity can be taken to be diagonal as $A = diag[a_1, a_2, ..., a_n]$, X, X_1 are matrices containing the centres of the partitions of every fuzzy output variable of $f(x)$ and $g(x)$ respectively, $S(\chi), S_1(\chi)$ are matrices containing high order combinations of sigmoid functions of the state χ and W, W_1 are matrices containing respective neural weights according to (18) and (12). The dimensions and the contents of all the above matrices are chosen so that $XWS(\chi)$ is a $n \times 1$ vector and $X_1 W_1 S_1(\chi)$ is a $n \times n$ matrix. Without compromising the generality of the model we assume that the vector fields in (15) are such that the matrix G is diagonal. For notational simplicity we assume that all output fuzzy variables are partitioned to the same number, m, of partitions. Under these specifications X is a $n \times n \cdot m$ block diagonal matrix of the form $X = diag(X^1, X^2, \ldots, X^n)$ with each X^i being an m-dimensional raw vector of the form

$$X^i = [\, \bar{f}^i_1 \quad \bar{f}^i_2 \quad \cdots \quad \bar{f}^i_m \,]$$

where $\bar{f}^i_p$ denotes the centre of the p-th partition of f_i. Also, $S(\chi) = [\, s_1(\chi) \quad \ldots \quad s_k(\chi) \,]^T$, where each $s_i(\chi)$ is a high order combination of sigmoid functions of the state variables and W is a $n \cdot m \times k$ matrix with neural weights. W assumes the form $W = [\, W^1 \quad \cdots \quad W^n \,]^T$, where each W^i is a matrix $\left[w^i_{j\,l} \right]_{m \times k}$. X_1 is a $n \times n \cdot m$ block diagonal matrix $X_1 = diag({}^1X^1, {}^1X^2, \ldots, {}^1X^n)$ with each ${}^1X^i$ being an m-dimensional raw vector of the form

$${}^1X^i = [\, \bar{g}^{i,i}_1 \quad \bar{g}^{i,i}_2 \quad \cdots \quad \bar{g}^{i,i}_m \,],$$

where $\bar{g}^{i\,i}_k$ denotes the center of the k-th partition of g_{ii}. W_1 is a $m \cdot n \times n$ block diagonal matrix $W_1 = diag({}^1W^1, {}^1W^2, \ldots, {}^1W^n)$, where each ${}^1W^i$ is a column vector $\left[{}^1w^i_{j\,l} \right]_{m \times 1}$ of neural weights. Finally, $S_1(\chi)$ is a $n \times n$ diagonal matrix with each diagonal element $s_i(\chi)$ being a high order combination of sigmoid functions of the state variables.

We assume the existence of only parameter uncertainty, so, we can take into account that the actual system (13) can be modeled by the following neural form

$$\dot{\chi} = A\chi + XW^* S(\chi) + X_1 W_1^* S_1(\chi) u \tag{20}$$

Define now, the error between the identifier states and the real states as

$$e = \hat{\chi} - \chi \tag{21}$$

Then from (19) and (21) we obtain the error equation

$$\dot{e} = Ae + X\tilde{W}S(\chi) + X_1\tilde{W}_1S_1(\chi)u \tag{22}$$

Where $\tilde{W} = W - W^*$ and $\tilde{W}_1 = W_1 - W_1^*$. Regarding the identification of W and W_1 in (19) we are now able to state the following theorem.

Theorem 1: Consider the identification scheme given by (22). The learning law

a) For the elements of W^i

$$\dot{w}^i_{j\,l} = -\bar{f}^i_j p_i e_i s_l(\chi) \tag{23}$$

b) For the elements of ${}^1W^i$

$${}^1\dot{w}^i_{j\,1} = -\bar{g}^{i\,i}_j p_i e_i u_i s_i(\chi) \tag{24}$$

or equivalently ${}^1\dot{W}^i = -({}^1X^i)^T p_i e_i u_i s_i(\chi)$ with $i = 1, \ldots, n$, $j = 1, \ldots, m$, $l = 1, \ldots, k$ guarantees the following properties.

- $e, \hat{\chi}, \tilde{W}, \tilde{W}_1 \in L_\infty, \quad e \in L_2$
- $\lim_{t\to\infty} e(t) = 0$, $\lim_{t\to\infty} \dot{\tilde{W}}(t) = 0$, $\lim_{t\to\infty} \dot{\tilde{W}}_1(t) = 0$

Proof: Consider the Lyapunov function candidate,

$$V(e, \tilde{W}, \tilde{W}_1) = \frac{1}{2}e^TPe + \frac{1}{2\gamma_1}tr\{\tilde{W}^T\tilde{W}\} + \frac{1}{2\gamma_2}tr\{\tilde{W}_1^T\tilde{W}_1\}$$

Where $P > 0$ is chosen to satisfy the Lyapunov equation

$$PA + A^TP = -I$$

Taking the derivative of the Lyapunov function candidate we get

$$\dot{V}(e, \tilde{W}, \tilde{W}_1) = \frac{1}{2}\dot{e}^TPe + \frac{1}{2}e^TP\dot{e} + \frac{1}{\gamma_1}tr\{\dot{\tilde{W}}^T\tilde{W}\} + \frac{1}{\gamma_2}tr\{\dot{\tilde{W}}_1^T\tilde{W}_1\} \Rightarrow$$

$$\dot{V} = \frac{1}{2}e^TA^TPe + \frac{1}{2}S^T\tilde{W}^TX^TPe + \frac{1}{2}U^TS_1^T\tilde{W}_1^TX^TPe + \frac{1}{2}e^TPAe + \frac{1}{2}e^TPX\tilde{W}S + \frac{1}{2}e^TPX\tilde{W}_1S_1U + \frac{1}{\gamma_1}tr\{\dot{\tilde{W}}^T\tilde{W}\} + \frac{1}{\gamma_2}tr\{\dot{\tilde{W}}_1^T\tilde{W}_1\} \Rightarrow$$

$$\dot{V} = \frac{1}{2}e^T\left(A^TP + PA\right)e + \left(\frac{1}{2}e^TPX\tilde{W}S + \frac{1}{2}e^TPX\tilde{W}S\right) + \left(\frac{1}{2}e^TPX\tilde{W}_1S_1U + \frac{1}{2}e^TPX\tilde{W}_1S_1U\right) + \frac{1}{\gamma_1}tr\{\dot{\tilde{W}}^T\tilde{W}\} + \frac{1}{\gamma_2}tr\{\dot{\tilde{W}}_1^T\tilde{W}_1\} \Rightarrow$$

$$\dot{V} = -\frac{1}{2}e^Te + e^TPX\tilde{W}S + e^TPX\tilde{W}_1S_1U + \frac{1}{\gamma_1}tr\{\dot{\tilde{W}}^T\tilde{W}\} + \frac{1}{\gamma_2}tr\{\dot{\tilde{W}}_1^T\tilde{W}_1\} \Rightarrow$$

$$\dot{V} = -\frac{1}{2}e^Te \leq 0$$

when

$$\frac{1}{\gamma_1}tr\{\dot{\tilde{W}}^T\tilde{W}\} = -e^TPX\tilde{W}S \tag{25}$$

$$\frac{1}{\gamma_2}tr\{\dot{\tilde{W}}_1^T\tilde{W}_1\} = -e^TPX_1\tilde{W}_1S_1u \tag{26}$$

Then, taking into account the form of W and W_1 the above equations result in the element wise learning laws given in (23), (24). These laws can also be written in the following compact form

$$\dot{W} = -\gamma_1X^TPeS^T \tag{27}$$

$$\dot{W}_1 = -\gamma_2X_1^TPEUS_1^T \tag{28}$$

Where E and U are diagonal matrices such that $E = diag(e_1, \ldots e_n)$ and $U = diag(u_1, \ldots u_n)$.

Using the above Lyapunov function candidate V and proving that $\dot{V} \leq 0$ all properties of the theorem are assured [23]. ■

V. Experimental results

To demonstrate the potency of the proposed scheme we present two simulation results. One of them is the well known benchmark "Inverted Pendulum" and the other "Van der pol" oscillator. Both of them present comparisons of the proposed method with two well established approaches on adaptive fuzzy system identification [2] and on the use of RHONN approximators [35] respectively. The comparison shows off the minimal parameter requirements of the proposed method with respect to the traditional method of [2]. It also demonstrates its functional approximation superiority against both approaches.

A. Comparison of function approximation abilities on the well known benchmarks of inverted pendulum and Van der Pol oscillator

1) Inverted Pendulum: Let the well known system of an inverted pendulum. Its dynamical equations can assume the following Brunovsky canonical form [34]

$$\dot{x}_1 = x_2$$

$$\dot{x}_2 = \frac{g\sin x_1 - \frac{mlx_2^2\cos x_1 \sin x_1}{m_C+m}}{l\left(\frac{4}{3} - \frac{m\cos^2 x_1}{m_C+m}\right)} + \frac{\frac{cos x_1}{m_C+m}}{l\left(\frac{4}{3} - \frac{m\cos^2 x_1}{m_C+m}\right)}u \tag{29}$$

where $x_1 = \theta$ and $x_2 = \dot{\theta}$ are the angle from the vertical position and the angular velocity respectively. Also, $g = 9.8\ m/s^2$ is the acceleration due to gravity, m_c is the mass of the cart, m is the mass of the pole, and l is the half-length of the pole. We choose $m_c = 1\ kg$, $m = 0.1\ kg$, and $l = 0.5\ m$ in the following simulation. In this case we also have that $|x_1| \leq \pi/6$ and $|x_2| \leq \pi/6$.

It is our intention to compare the approximation abilities of the proposed Neuro-Fuzzy approach with Wang [2] adaptive Fuzzy approach and RHONN [35]. Eq. (29) is similar with Eq. (13), so we assume that $f(x)$ and $g(x)$ can be approximated using Wang's approach and Eq. (2) or alternatively by the *XWS* and $X_1W_1S_1$ term of Eq. (19) in the proposed approach, or *WS* and W_1S_1 for RHONN approach [35] respectively, . The weight updating laws are chosen to be: For the Wang approach ([2], page 115)

$$\dot{\theta}_f = -\gamma_1e^TPb_c\xi(x) \tag{30}$$

TABLE I

COMPARISON OF WANG, RHONN AND FHONNF APPROACHES FOR THE INVERTED PENDULUM WITH 2 HOST.

	Wang	RHONN	FHONNF
$MSEx_1$	0.0456	0.0189	0.0064
$MSEx_2$	0.0530	0.0261	0.0068

$$\dot{\theta}_g = -\gamma_2 e^T P b_c \xi(x) u_c \tag{31}$$

where only the simplified approach, without parameter projection case was necessary to be used.

For the RHONN approach we use tha adaptive laws, which are described in [35], page 37.

For the proposed F-HONNF approach we use the adaptive laws which are described by Eqs. (27) and (28). Numerical training data were obtained by using Eq. (29) with initial conditions $[\, x_1(0) \quad x_2(0)\,] = [\, \frac{\pi}{6} \quad -\frac{\pi}{6}\,]$, and a persistently exciting input $u = 1 + 0.8\sin(0.001t)$.

The approximation of the dynamical equations using conventional fuzzy system approach requires a very large number of fuzzy rules for the approximation of the unknown functions. Choosing 40 or more membership functions for each variable x_i results in very accurate fuzzy representation. This representation requires 1600 rules, which in turn leads to a parameter explosion when using an adaptive scheme like that of Eq. (2) and consequently, it takes plenty of time for the simulations.

We are using the proposed approach with Eq. (19) to approximate Inverted Pendulum dynamics. The proposed Neuro-Fuzzy model was chosen to use 5 output partitions of f and 5 output partitions of g. The number of high order sigmoidal terms (HOST) used in HONNF's were chosen to be first 2 $(s(x_1), s(x_2))$ and secondly 5 $(s(x_1), s(x_2), s(x_1) \cdot s(x_2), s^2(x_1), s^2(x_2))$ for two different simulations with the same benchmark. Therefore, the number of adjustable weights is 20 or 50 respectively, which is a much smaller number to that used in the conventional fuzzy approach.

In order our model to be equivalent with regard to other parameters except the adjustable weights we have chosen terms $\gamma_1 P b_c$ in Eq. (30) and $\gamma_1 P_1$ (the updating learning rates) in Eq. (31) to have the same values. Also, the RHONN model given from [35] is constructed with the same learning parameters and number of high order terms with these of F-HONNF approach. The parameters of the sigmoidal terms were chosen to be $a_1 = 0.1$, $a_2 = 6$, $b_1 = b_2 = 1$ and $c_1 = c_2 = 0$. Fig. (1) and (2) shows the approximation of states x_1 and x_2 respectively while fig. (3) and (4) gives the evolution of errors x_1 and x_2.

The mean squared error (MSE) for Wang's, RHONN and F-HONNF approaches were measured and are shown in Tables I and II, demonstrating a significant (order of magnitude) increase in the approximation performance with the maximum diferrence appearing between Wang's and F-HONNF, although in the F-HONNF approach no a-priori information regarding the inputs is used.

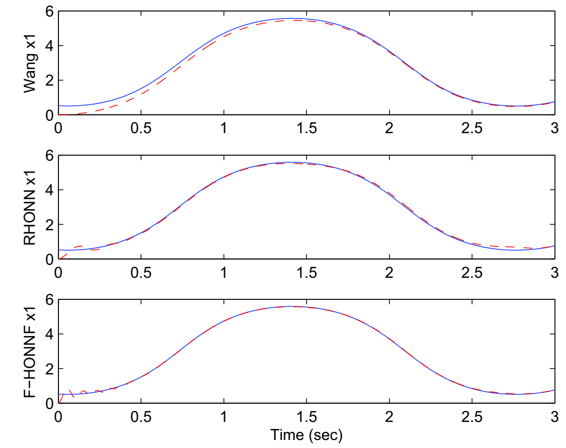

Fig. 1. Evolution of variable x_1 for Wang, RHONN and F-HONNF approach

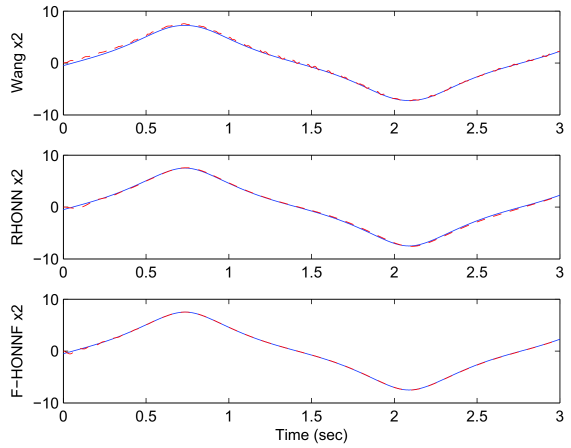

Fig. 2. Evolution of variable x_2 for Wang, RHONN and F-HONNF approach

2) Van der pol: Van der Pol oscillator is usually used as a simple benchmark problem for testing identification and control schemes. It's dynamical equations are given by

$$\dot{x}_1 = x_2$$

$$\dot{x}_2 = x_2 \cdot \left(a - x_1^2\right) \cdot b - x_1 + u \tag{32}$$

The procedure of the approximation was the same as that of Inverted Pendulum. So, using the conventional fuzzy system approach we observe that for very accurate fuzzy representation [33] requires a very large number of fuzzy

TABLE II

COMPARISON OF WANG, RHONN AND FHONNF APPROACHES FOR THE INVERTED PENDULUM WITH 5 HOST.

	Wang	RHONN	FHONNF
$MSEx_1$	0.0456	0.0069	0.0021
$MSEx_2$	0.0530	0.0098	0.0020

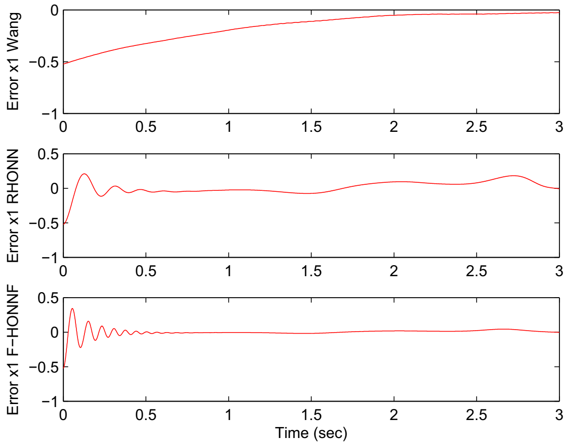

Fig. 3. Approximation Error of variable x_1 for Wang, RHONN and F-HONNF approach

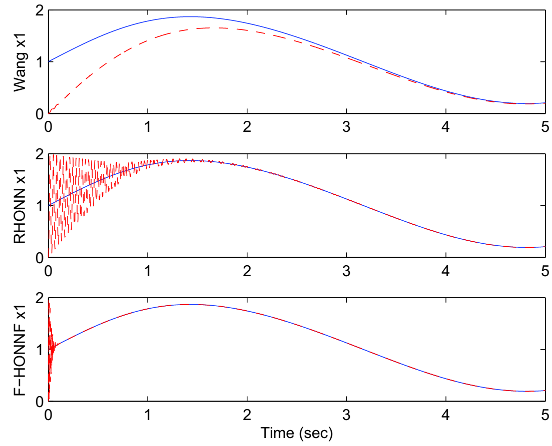

Fig. 5. Evolution of variable x_1 for Wang, RHONN and F-HONNF approach

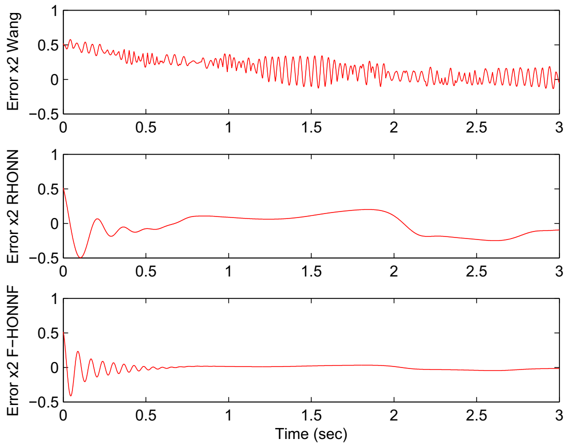

Fig. 4. Approximation Error of variable x_2 for Wang, RHONN and F-HONNF approach

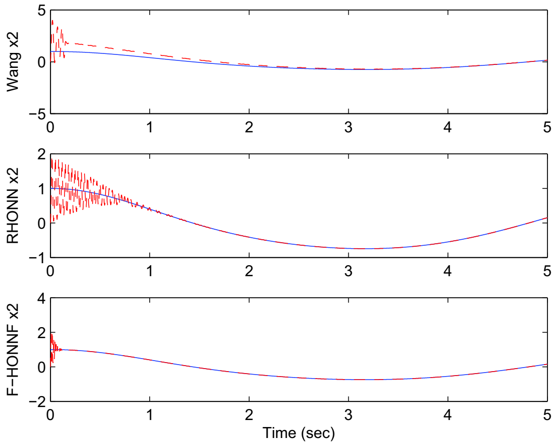

Fig. 6. Evolution of variable x_2 for Wang, RHONN and F-HONNF approach

rules almost 1500 (27 membership functions used). This in turn would lead to a parameter explosion like that of "Inverted Pendulum" case which were discussed before.

The proposed Neuro-Fuzzy model was chosen to have the same parameters as before except the initial conditions, $[x_1(0) \quad x_2(0)] = [1 \quad 1]$. Fig. (5) and (6) shows the approximation of states x_1 and x_2, while fig. (7) and (8) presents the evolution of errors x_1 and x_2 respectively.

The mean squared error (MSE) for Wang's, RHONN and F-HONNF approaches were measured and are shown in Tables III and IV, demonstrating as before (Inverted Pendulum case), a significant (order of magnitude) increase in the approximation performance , although no a-priori information regarding fuzzy partitions and membership functions of the inputs were used.

Conclusively, the comparison between Wang and F-HONNF's leads to a large superiority of F-HONNF's regarding the number of adjustable parameters and the approximation abilities. With respect to the RHONN approach the proposed F-HONNF approach is also much better although it does not present the same large difference as with the Wang's approach.

VI. CONCLUSIONS AND FUTURE WORKS

The identification of unknown nonlinear dynamical systems using a new definition of Adaptive Neuro Fuzzy Systems was presented in this paper. The proposed scheme uses the concept of Adaptive Fuzzy Systems operating in conjunction with High Order Neural Network Functions (F-HONNFs). Under this scheme the identification is driven to a Recurrent High Order Neural Network, which however takes into account the fuzzy output partitions of the initial

TABLE III

COMPARISON OF WANG, RHONN AND FHONNF APPROACHES FOR VAN DER POL OSCILLATOR WITH 2 HOST.

	Wang	RHONN	FHONNF
$MSEx_1$	0.1038	0.0303	0.0058
$MSEx_2$	0.1401	0.0259	0.0087

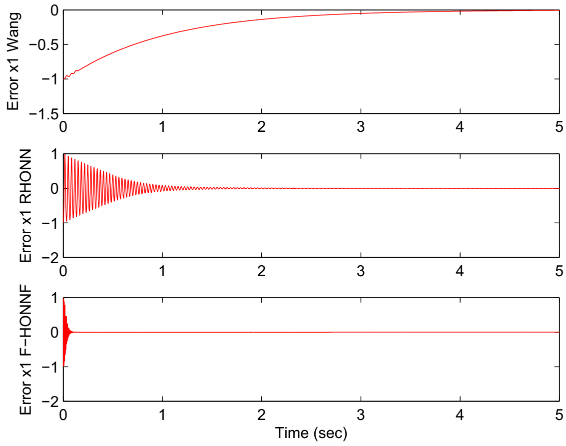

Fig. 7. Approximation Error of variable x_1 for Wang, RHONN and F-HONNF approach

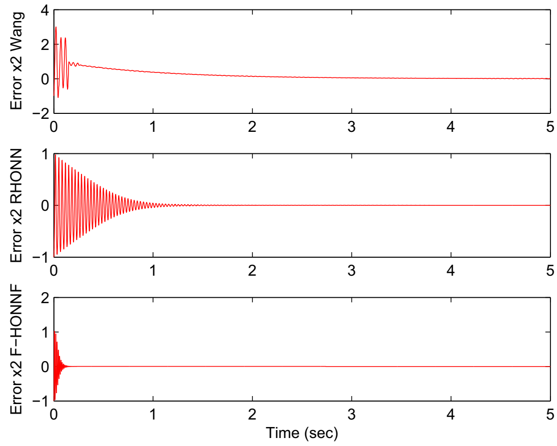

Fig. 8. Approximation Error of variable x_2 for Wang, RHONN and F-HONNF approach

AFS. The proposed scheme does not require a-priori experts' information on the number and type of input variable membership functions making it less vulnerable to initial design assumptions. Weight updating laws for the involved HONNFs are provided, which guarantee that the identification error reaches zero exponentially fast. Simulations illustrate the potency of the method by comparing its performance with this of other well known approaches. Future work will include the use of the proposed identification scheme as the first part in control algorithms.

TABLE IV
COMPARISON OF WANG, RHONN AND FHONNF APPROACHES FOR VAN DER POL OSCILLATOR WITH 5 HOST.

	Wang	RHONN	FHONNF
$MSEx_1$	0.1038	0.0180	0.0013
$MSEx_2$	0.1401	0.0149	0.0018

REFERENCES

[1] K. Hornic, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators", *Neural Networks*, vol. 2, pp. 359-366, 1989.

[2] L. Wang, *Adaptive Fuzzy Systems and Control*, Prentice Hall, Englewood Cliffs, NJ, 1994.

[3] K. M. Passino and S. Yurkovich, *Fuzzy Control*, Addison Wesley Longman, Menlo Park, CA, 1998.

[4] J. S. R. Jang, "ANFIS: Adaptive-network-based fuzzy inference system", *IEEE Trans. Syst. Man. Cyber.*, vol. 23, pp. 665-684, 1993.

[5] C.T. Lin, "A neural fuzzy control system with structure and parameter learning", *Fuzzy Sets and Systems*, vol. 70, pp. 183-212, 1995.

[6] K.B. Cho and B.H.Wang, "Radial basis function based adaptive fuzzy systems and their applications to system identification and prediction", *Fuzzy Sets and Systems*, vol. 83, pp. 325-339, 1996.

[7] C.F.Juang and C. T. Lin, "An on-line self-constructing neural fuzzy inference network and its applications", *IEEE Trans. Fuzzy Syst.*, vol. 6, pp. 12-32, 1998.

[8] R. P. Li and M. Mukaidono, "A new approach to rule learning based on fusion of fuzzy logic and neural networks", *IEICE Trans. Fuzzy Syst*, vol. E78-d, pp. 1509-1514, 1995.

[9] S. L. Chiu, "Fuzzy model identification based on cluster estimation", *Journal of Intelligent and Fuzzy Systems*, vol. 2, 1994.

[10] Y. H. Lin and G. A. Cunningham, "A new approach to fuzzy-neural system modelling", *IEEE Trans. Fuzzy Syst.*, vol. 3, pp. 190-197, 1995.

[11] C. F. Jang and C. T. Lin, "An online self-constructing neural fuzzy inference network and its applications", *IEEE Trans. Fuzzy Syst.*, vol. 6, pp. 12-32, 1998.

[12] S. Mitra and Y. Hayashi, "Neuro-fuzzy rule generation: survey in soft computing framework", *IEEE Trans. On Neural Networks*, vol. 11, pp. 748-768, 2000.

[13] B. S. Chen, C. H. Lee, and Y. C. Chang, "Tracking design of uncertain nonlinear siso systems: Adaptive fuzzy approach", *IEEE Trans. Fuzzy Syst.*, vol. 4, pp. 32-43, Feb. 1996.

[14] J. T. Spooner and K. M. Passino, "Stable adaptive control using fuzzy systems and neural networks", *IEEE Trans. Fuzzy Syst.*, vol. 4, pp. 339-359, June 1996.

[15] K. S. Narendra and K. Parthasarathy, "Identification and control of dynamical systems using neural networks", *IEEE Trans. Neural Networks*, vol. 1, pp. 4-27, Feb.1990.

[16] M. M. Polycarpou and M. J. Mears, "Stable adaptive tracking of uncertain systems using nonlinearly parameterized online approximators", *Int. J. Control*, vol. 70, no. 3, pp. 363-384, 1998.

[17] G. A. Rovithakis and M. A. Christodoulou, "Adaptive control of unknown plants using dynamical neural networks", *IEEE Trans. Syst, Man, Cybern.*, vol. 24, , pp. 400-412 Mar. 1994.

[18] N. Golea, A. Golea, K. Benmahammed, "Stable indirect fuzzy adaptive control", *Fuzzy Sets and Systems*, vol. 137, pp. 353-366, 2003.

[19] T. Lina, C. Wang, H. Liub, "Observer-based indirect adaptive fuzzy-neural tracking control for nonlinear SISO systems using VSS and H_1 approaches", *Fuzzy Sets and Systems*, vol. 143, pp. 211-232, 2004.

[20] Y. Diao, K.M. Passino, Adaptive Neural/Fuzzy Control for Interpolated Nonlinear Systems, *IEEE Trans. on Fuzzy Systems*, Vol. 10, No. 5, pp. 583-595, Oct. 2002.

[21] R. Ordonez, K.M. Passino, "Adaptive Control for a Class of Nonlinear Systems with Time-Varying Structure", *IEEE Trans. on Automatic Control*, Vol. 46, No. 1, pp. 152-155, Jan. 2001.

[22] D. Sbarbaro, T.A. Johansen, "Analysis of Artificial Neural Networks for Pattern-Based Adaptive Control", *IEEE Trans. on Neural Networks*, vol. 17, no. 5, pp. 1184 - 1193, Sept. 2006.

[23] P. Ioannou and B. Fidan, *Adaptive control tutorial*, SIAM: Advances in Design and Control Series, 2006.

[24] T. Hayakawa, W.M. Haddad, N. Hovakimyan, V. Chellaboina, "Neural network adaptive control for nonlinear nonnegative dynamical systems", *IEEE Trans. on Neural Networks*, vol. 16, no. 2, pp. 399 - 413, March 2005.

[25] S. G. Shuzhi, T. H. Thanh, "Semiglobal ISpS Disturbance Attenuation With Output Tracking via Direct Adaptive Design", *IEEE Trans. on Neural Networks*, vol. 18, no. 4, pp. 1129 - 1148, July 2007.

[26] N. Hovakimyan, E. Lavretsky, J. Y. Bong, A.J. Calise, "Coordinated decentralized adaptive output feedback control of interconnected systems", *IEEE Trans. on Neural Networks*, vol. 16, no. 1, pp. 185 - 194, Jan. 2005.

[27] S.S. Ge, W. Jing, "Robust adaptive neural control for a class of perturbed strict feedback nonlinear systems", *IEEE Trans. on Neural Networks*, vol. 13, no. 6, pp. 1409 - 1419, Nov. 2002.

[28] C. Lee and C. Teng, "Identification and Control of Dynamic Systems Using Recurrent Fuzzy Neural Networks", *IEEE Trans. Fuzzy Syst.*, vol. 8, pp. 349-366, 2000.

[29] M. Hojati and S. Gazor, "Hybrid Adaptive Fuzzy Identification and Control of Nonlinear Systems", *IEEE Trans. Fuzzy Syst.*, vol. 10, pp. 198-210, April 2002.

[30] H.N. Nounou, K.M. Passino , Stable Auto-Tuning of Hybrid Adaptive Fuzzy/Neural Controllers for Nonlinear Systems, Engineering Applications of Artificial Intelligence, Vol. 18, No. 3, pp. 317-334, April 2005.

[31] E. B. Kosmatopoulos and M. A. Christodoulou, "Recurrent neural networks for approximation of fuzzy dynamical systems", *Int. Journal of Intelligent Control and Systems*, vol. 1, pp. 223-233, 1996.

[32] M. A. Christodoulou, D. C. Theodoridis, and Y. S. Boutalis, "Building Optimal Fuzzy Dynamical Systems Description Based on Recurrent Neural Network Approximation", *in Proc. Int. Conf. of Networked Distributed Systems for Intelligent Sensing and Control*, Kalamata, Greece, June, 2007.

[33] H. L. Hiew, C. P. Tsang, *Fuzzy Chaos and Recursive Partitioning*, in B. Kosko (Ed.), Fuzzy Engineering, Prentice-Hall International Inc. New Jersey, 1997

[34] J. E. slotine, W. Li, *Applied Nonlinear Control*, Prentice-Hall International Inc. New Jersey, 1997

[35] G.A. Rovithakis and M.A.Christodoulou, *Adaptive Control with Recurrent High Order Neural Networks (Theory and Industrial Applications)*, in Advances in Industrial Control, M.A.Grimble and M.A.Johnson, Springer Verlag London Limited, 2000.