Weight Tracking in Nonlinear System Identification via Fuzzy High Order Neural Network Function Approximation

M.A. Christodoulou[a], D.C. Theodoridis[b], Y.S. Boutalis[b*]

Abstract— **The weight tracking in the identification of varying unknown nonlinear systems is examined in this paper. The unknown nonlinear system is represented and identified by an Adaptive Dynamic Fuzzy Systems (ADFS), which operates in conjunction with High Order Neural Network Functions (HONNFs) and takes the form of a Fuzzy Recurrent High Order Neural Network (F-RHONN). Weight updating laws for the involved HONNFs are given, which guarantee that the identification error reaches zero exponentially fast. The proposed scheme has the ability to track very fast any change in the unknown nonlinear system, that can be reflected in weight changes of its F-RHONN representation. Simulations illustrate the potency of the method especially in tracking the changes made in the unknown system.**

I. Introduction

In dynamical systems the mathematical description of the system is required, so that we are able to control it. Unfortunately, the exact mathematical model of the plant, especially when this is highly nonlinear and complex, is rarely known and thus appropriate identification schemes have to be applied which will provide us with an approximate model of the plant. Recently very efficient techniques have been proposed for the identification of complicated nonlinear systems using neural networks and fuzzy systems.

It has been established that neural networks and fuzzy inference systems are universal approximators [1], [2] that is, they can approximate any nonlinear function to any prescribed 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 modelling [3]-[9], which is the first component of indirect adaptive control schemes [10]-[18], where first the dynamics of the system are identified and then a control input is generated according to the certainty equivalence principle.

In this paper, we consider the problem of approximating general nonlinear dynamical systems of the form

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

using the concept of Fuzzy Dynamical Systems (FDS) operating in conjunction with Recurrent High Order Neural Networks (RHONNs). An ADFS approximates the function of general nonlinear dynamical systems by covering its graph with fuzzy patches in the output state space. Each fuzzy rule defines a fuzzy patch and connects commonsense knowledge with state-space geometry. F-RHONNs (or statistical clustering systems) can approximate the unknown fuzzy patches from training data. These adaptive fuzzy systems approximate a function at two levels. At the local level the F-RHONN approximates and tunes the fuzzy rules. At the global level the rules or patches approximate the function.

Recently [19], [20], higher 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. In this paper HONNFs are also used for the neuro fuzzy identification of unknown nonlinear dynamical systems. This approximation depends on the fact that fuzzy rules could be identified with the help of HONNFs. The same rationale has been employed in [21], [22], where a neuro – fuzzy approach for the indirect control of unknown systems has been introduced.

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 [23], such as fuzzy C-means clustering [6], mountain clustering [8], and subtractive clustering [9]. 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

[a] M.A Chriastodoulou is with Department of Electronic and Computer Engineering, Technical University of Crete,73100, Chania, Crete, Greece,

[b] D.C. Theodoridis and Y.S. Boutalis are with the Automatic Control Systems Laboratory, Department of Electrical & Computer Engineering, Democritus University of Thrace, V. Sofias 12, 67100 Xanthi, Greece, ybout@ee.duth.gr, dtheodo@ee.duth.gr

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. Changes performed in the real system during its operation could be reflected to sudden weight changes in its F-RHONN representation. The fast convergence of the error to zero guarantees also the fast tracking of the weight changes.

The paper is organized as follows. Section II presents the concept of adaptive fuzzy systems (AFS) using rule indicator functions 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, while the adaptive parameter identification is presented in Section V, where the associated weight adaptation laws are given. Finally, simulation results are given in Section VI , which demonstrate the fast weight tracking abilities of the proposed scheme.

II. THE CONCEPT OF FUZZY DYNAMICAL SYSTEMS

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

$$z(k) = f\left(x(k), u(k)\right) \tag{2}$$

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

$$\dot{x}(k) = f\left(x(k), u(k)\right) \tag{3}$$

By setting $y(k) = \left[x(k), u(k)\right]$, Eq. (2) may be rewritten as follows

$$z(k) = f\left(y(k)\right) \tag{4}$$

with $y \subset \Re^{m+n}$.

In case f in (4) is unknown we may wish to approximate it by using a fuzzy representation. In this case both $y(k)$ 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 measurements, 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 $z(k)$ (Wang, 1994). 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 $\left(j_1, j_2, \ldots, j_{n+m}\right)^{th}$ input fuzzy patch and pointing - through the vector field $f(\cdot)$ - to the subset of $z(k)$, which belong to the $\left(l_1, l_2, \ldots, l_n\right)^{th}$ 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 specific output values.

Also, in this subsection, we are briefly introducing the representation of fuzzy systems using the rule firing indicator functions (RFIF), or simply indicator functions (IF), which is used for the development of the proposed method.

According to the above notation the Indicator Function (IF) connected to $\Omega^{l_1,l_2,\ldots,l_n}_{j_1,j_2,\ldots,j_{n+m}}$ is defined as follows:

$$I^{l_1,l_2,\ldots,l_n}_{j_1,j_2,\ldots,j_{n+m}}\left(x(k),u(k)\right) = \begin{cases} a & if\ \left(x(k),u(k)\right) \in \Omega^{l_1,l_2,\ldots,l_n}_{j_1,j_2,\ldots,j_{n+m}} \\ 0 & otherwise \end{cases} \tag{5}$$

where a denotes the firing strength of the rule.

Define now the following system

$$z(k) = \sum \overline{z}^{l_1,l_2,\ldots,l_n}_{j_1,j_2,\ldots,j_{n+m}} \times I^{l_1,l_2,\ldots,l_n}_{j_1,j_2,\ldots,j_{n+m}}\left(x(k),u(k)\right) \tag{6}$$

Where $\overline{z}^{l_1,l_2,\ldots,l_n}_{j_1,j_2,\ldots,j_{n+m}} \in \Re^n$ be any constant vector consisting of the centres of the membership functions of each output variable z_i and $I^{l_1,l_2,\ldots,l_n}_{j_1,j_2,\ldots,j_{n+m}}\left(x(k),u(k)\right)$ is the IF. Then, according to [19], [20] the system in (6) is a generator for the fuzzy system (FS).

It is obvious that Eq. (6) can be also valid for dynamic systems. In its dynamical form it becomes

$$\dot{x}(k) = \sum \overline{x}^{l_1,l_2,\ldots,l_n}_{j_1,j_2,\ldots,j_{n+m}} \times I^{l_1,l_2,\ldots,l_n}_{j_1,j_2,\ldots,j_{n+m}}\left(x(k),u(k)\right) \tag{7}$$

Where $\overline{x}^{l_1,l_2,\ldots,l_n}_{j_1,j_2,\ldots,j_{n+m}} \in \Re^n$ be again any constant vector consisting of the centres of fuzzy partitions of every variable x_i and $I^{l_1,l_2,\ldots,l_n}_{j_1,j_2,\ldots,j_{n+m}}\left(x(k),u(k)\right)$ is the IF.

III. HONNFS 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 $I^{l_1,l_2,\ldots,l_n}_{j_1,j_2,\ldots,j_{n+m}}$. However, in order the approximation problem to make sense the space $y := x \times u$ must be compact. Thus, our first assumption is the following:

Assumption 1: $y := x \times u$ is a compact set.

Notice that since $y \subset \Re^{n+m}$, the above assumption is identical to the assumption that it is closed and bounded.

Also, it is noted that even if 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 y for all $t<T$; that is if $y_T := \{(x(k),u(k)) \in y, t<T\}$ we may replace assumption 1 by the following assumption.

Assumption 2: y_T is a compact set.

It is worth noticing, that while assumption 1 requires the system in Eq. (3) solutions to be bounded for all $u(k) \in U_C$ and $x(0) \in X$, assumption 2 requires the system in Eq. (6) solutions to be bounded for a finite time period; thus, assumption 1 requires the system in Eq. (3) to be bounded input bounded state (BIBS) stable while assumption 2 is valid for systems that are not BIBS stable and, even more, for unstable systems and systems with finite escape times.

Based on the fact that functions of high order neurons are capable of approximating discontinuous functions [19] and [20] use high order neural network functions HONNFs in order to approximate the IF $I^{l_1,l_2,\ldots,l_n}_{j_1,j_2,\ldots,j_{n+m}}$, a HONNF is defined as:

$$N\left(x(k),u(k);w,L\right) = \sum_{hot=1}^{L} w_{hot} \prod_{j\in I_{hot}} \Phi_j^{d_j(hot)} \tag{8}$$

where $I_{hot} = \{I_1, I_2, \ldots, I_L\}$ (hot: high order terms) is a collection of L not-ordered subsets of $\{1,2,\ldots,m+n\}$, $d_j(hot)$ are non-negative integers, Φ_j are sigmoid functions of the state or the input, which are the elements of the following vector

$$\Phi = \begin{bmatrix} \Phi_1 \\ \vdots \\ \Phi_n \\ \Phi_{n+1} \\ \vdots \\ \Phi_{m+n} \end{bmatrix} = \begin{bmatrix} s(x_1) \\ \vdots \\ s(x_n) \\ s(u_1) \\ \vdots \\ s(u_m) \end{bmatrix} \tag{9}$$

Where $S(x) = \alpha \dfrac{1}{1+e^{-\beta x}} - \gamma$ and $w := [w_1 \cdots w_L]^T$ are the HONNF weights. Eq. (8) can also be written

$$N\left(x(k),u(k);w,L\right) = \sum_{hot=1}^{L} w_{hot} S_{hot}\left(x(k),u(k)\right) \tag{10}$$

where $S_{hot}\left(x(k),u(k)\right)$ are high order terms of sigmoid functions of the state and/or input.

The next lemma [19] states that a HONNF of the form in Eq. (10) can approximate the indicator function $I^{l_1,l_2,\ldots,l_n}_{j_1,j_2,\ldots,j_{n+m}}$.

Lemma 1: Consider the indicator function $I^{l_1,l_2,\ldots,l_n}_{j_1,j_2,\ldots,j_{n+m}}$ and the family of the HONNFs $N\left(x(k),u(k);w,L\right)$. Then for any $\varepsilon > 0$ 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

$$\sup_{(x(k),u(k))\in \overline{y}} \{ I^{l_1,l_2,\ldots,l_n}_{j_1,j_2,\ldots,j_{n+m}}\left(x(k),u(k)\right) - N\left(x(k),u(k);w^{j_1,\ldots,j_{n+m};l_1,\ldots,l_n}, L^{j_1,\ldots,j_{n+m};l_1,\ldots,l_n}\right)\} < \varepsilon$$

where $\overline{y} \equiv y$ if assumption 1 is valid and $\overline{y_T} \equiv y$ if assumption 2 is valid.

Let us now keep $L^{j_1,\ldots,j_{n+m};l_1,\ldots,l_n}$ constant, that is let us preselect the number of high order connections, and let us define the optimal weights of the HONNF with $L^{j_1,\ldots,j_{n+m};l_1,\ldots,l_n}$ high order connections as follows

$$w^* := \arg \min_{w\in R^{j_1,\ldots,j_{n+m};l_1,\ldots,l_n}} \times \sup_{(x(k),u(k))\in \overline{y}} \{| I^{l_1,l_2,\ldots,l_n}_{j_1,j_2,\ldots,j_{n+m}}\left(x(k),u(k)\right) - N\left(x(k),u(k);w,L^{j_1,\ldots,j_{n+m};l_1,\ldots,l_n}\right)|\} < \varepsilon$$

and the modelling error as follows

$$v^{l_1,l_2,\ldots,l_n}_{j_1,j_2,\ldots,j_{n+m}}\left(x(k),u(k)\right) = I^{l_1,l_2,\ldots,l_n}_{j_1,j_2,\ldots,j_{n+m}}\left(x(k),u(k)\right) - N\left(x(k),u(k);w^{j_1,\ldots,j_{n+m};l_1,\ldots,l_n}, L^{j_1,\ldots,j_{n+m};l_1,\ldots,l_n}\right)$$

It is worth noticing that from Lemma 1, we have that $\sup_{(x(k),u(k))\in \overline{y}} \left| v^{l_1,l_2,\ldots,l_n}_{j_1,j_2,\ldots,j_{n+m}}\left(x(k),u(k)\right)\right|$ can be made arbitrarily small by simply selecting appropriately the number of high order connections.

Using the approximation Lemma 1, it is natural to approximate system in Eq. (7) by the following dynamical system

$$z(k+1) = \sum \overline{x}^{l_1,l_2,\ldots,l_n}_{j_1,j_2,\ldots,j_{n+m}}\left(x(k),u(k)\right) \times N\left(z(k),u(k);w^{j_1,\ldots,j_{n+m};l_1,\ldots,l_n}, L^{j_1,\ldots,j_{n+m};l_1,\ldots,l_n}\right)$$

Let now $x(k)\left(x(0),u(k)\right)$ denote the solution in Eq. (7) given that the initial state at $t=0$ is equal to $x(0)$ and the input is $u(k)$. Similarly we define $z(k)\left(z(0),u(k)\right)$. Also let

$$\nu\left(z(k),u(k)\right) = \sum \left(\overline{x}^{l_1,l_2,\ldots,l_n}_{j_1,j_2,\ldots,j_{n+m}}\left(x(k),u(k)\right) \times \nu\left(z(k),u(k)\right)\right)$$

Then, it can be easily shown that

$$z(k)\left(z(0),u(k)\right) = x(k)\left(x(0),u(k)\right) + \nu\left(z(k),u(k)\right) \tag{11}$$

Note now that from the approximation Lemma 1, and the definition of $\nu\left(z(k),u(k)\right)$ we have that modelling error can be made arbitrarily small provided that $\left(z(k),u(k)\right)$ remain in a compact set (for example $\overline{y}$).

Theorem 1: [19], [20] Consider the FDS in (7) and suppose that system in Eq. (3) is its underlying system. Assume that either assumptions 1 or 2 hold. Also consider the RHONN in [20]. Then, for any $\varepsilon > 0$ there exists a matrix W^* and a number L^* of high order connections such as $W = W^*$ is a generator for the FDS.

IV. Neuro-fuzzy representation

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

$$\dot{x} = f(x) + G(x)\cdot u \tag{12}$$

where the state $x \in \Re^n$ is assumed to be completely measured, the control u is in $\Re^m$, f is an unknown smooth vector field called the drift term and G is a matrix with columns the unknown smooth controlled vector fields $g_i, i = 1,\ldots,n$ and $G = [g_1, g_2, \ldots, g_n]$.

We are using an affine in the control fuzzy dynamical system, which approximates the system in (12) and uses two fuzzy subsystem blocks for the description of $f(x)$ and $G(x)$ as follows

$$\hat{f}(x) = A\hat{x} + \sum \overline{f}^{l_1,\ldots,l_n}_{j_1,\ldots,j_n} \times I_{f\,j_1,\ldots,j_n}^{\;l_1,\ldots,l_n}(x) \quad (13)$$

$$\hat{g}_{ij}(x) = \sum (\overline{g}_{ij})^{l_1,\ldots,l_n}_{j_1,\ldots,j_n} \times I_{g\,j_1,\ldots,j_n}^{\;l_1,\ldots,l_n}(x) \quad (14)$$

Where A is a $n \times n$ stable matrix which for simplicity can be taken to be diagonal as $A = diag[-a_1, -a_2, \ldots, -a_n], a_i > 0$ and the summation is carried out over the number of all available fuzzy rules, I_f, I_g are appropriate fuzzy rule indicator functions and the meaning of indices $\square^{l_1,\ldots,l_n}_{j_1,\ldots,j_n}$ has already been described in the previous section.

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

$$\hat{f}(x \mid W_f) = A\hat{x} + \sum \overline{f}^{l_1,\ldots,l_n}_{j_1,\ldots,j_n} \times N_{f\,j_1,\ldots,j_n}^{\;l_1,\ldots,l_n}(x) \quad (15)$$

$$\hat{g}_{ij}(x \mid W_g) = \sum (\overline{g}_{ij})^{l_1,\ldots,l_n}_{j_1,\ldots,j_n} \times N_{g\,j_1,\ldots,j_n}^{\;l_1,\ldots,l_n}(x) \quad (16)$$

where W_f, W_g are weights that results from adaptive laws which will discussed later, and N_f, N_g are appropriate HONNFs.

So, the optimal approximation of $f(x)$ and $G(x)$ sub- functions of the dynamical system becomes

$$f(x \mid W_f^*) = Ax + \sum \overline{f}^{l_1,\ldots,l_n}_{j_1,\ldots,j_n} \times N^{*\,l_1,\ldots,l_n}_{f\,j_1,\ldots,j_n}(x) \quad (17)$$

$$g_{ij}(x \mid W_g^*) = \sum (\overline{g}_{ij})^{l_1,\ldots,l_n}_{j_1,\ldots,j_n} \times N^{*\,l_1,\ldots,l_n}_{g\,j_1,\ldots,j_n}(x) \quad (18)$$

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 (15), (16) are carried out over the number of the corresponding output partitions. Therefore, the affine in the control fuzzy dynamical system in (17), (18) is replaced by the following equivalent affine Recurrent High Order Neural Network (RHONN), which depends on the centres of the fuzzy output partitions $\overline{f_l}$ and $\overline{g}_{ij,l}$

$$\dot{\hat{x}} = A\hat{x} + \sum_{l=1}^{Npf} \overline{f} \times N_{f_l}(x) + \sum_{j=1}^{q} \left(\sum_{l=1}^{Npg_{ij}} \left(\overline{g}_{ij}\right)_l \times N_{g_l}(x) \right) \cdot u_j \quad (19)$$

Or in a more compact form

$$\dot{\hat{x}} = A\hat{x} + X_f W_f S_f(x) + X_g W_g S_g(x) u \quad (20)$$

Where X_f, X_g are matrices containing the centres of the partitions of every fuzzy output variable of $f(x)$ and $g(x)$ respectively, $S_f(x), S_g(x)$ are matrices containing high order combinations of sigmoid functions of the state x and W_f, W_g are matrices containing respective neural weights according to (10) and (19). The dimensions and the contents of all the above matrices are chosen so that $X_f W_f S_f(x)$ is a $n \times 1$ vector and $X_g W_g S_g(x)$ is a $n \times n$ matrix. Without compromising the generality of the model we assume that the vector fields in (14) 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_f is a $n \times n \cdot m$ block diagonal matrix of the form $X_f = diag\left(X_{f_1}, X_{f_2}, \ldots, X_{f_n}\right)$ with each X_{f_i} being an m-dimensional raw vector of the form

$$X_{f_i} = \begin{bmatrix} \overline{x}^1_{f_i} & \overline{x}^2_{f_i} & \cdots & \overline{x}^m_{f_i} \end{bmatrix}$$

where $\overline{x}^p_{f_i}$ with $p = 1,\ldots,m$ denotes the centre of the p-th partition of f_i. Also, $S_f(x) = [s_1(x) \;\; \ldots \;\; s_k(x)]^T$, where each $s_i(x)$ is a high order combination of sigmoid functions of the state variables and W_f is a $n \cdot m \times k$ matrix with neural weights. W_f assumes the form $W_f = \begin{bmatrix} W_{f_1} & \cdots & W_{f_n} \end{bmatrix}^T$, where each W_{f_i} is a matrix $\left[w^{pl}_{f_i}\right]_{m \times k}$. X_g is a $n \times n \cdot m$ block diagonal matrix $X_g = diag\left(X_{g_1}, X_{g_2}, \ldots, X_{g_n}\right)$ with each X_{g_i} being an m-dimensional raw vector of the form

$$X_{g_i} = \begin{bmatrix} \overline{x}^1_{g_i} & \overline{x}^2_{g_i} & \cdots & \overline{x}^m_{g_i} \end{bmatrix}$$

where $\overline{x}^k_{g_i}$ denotes the centre of the k-th partition of g_{ii}. W_g is a $m \cdot n \times n$ block diagonal matrix $W_g = diag\left(W_{g_1}, W_{g_2}, \ldots, W_{g_n}\right)$, where each W_{g_i} is a column vector $\left[w^p_{g_i}\right]_{m \times 1}$ of neural weights. Finally, $S_g(x)$ is a $n \times n$ diagonal matrix with each diagonal element $s_i(x)$ being a high order combination of sigmoid functions of the state variables.

V. Adaptive parameter identification

We assume the existence of only parameter uncertainty, so, we can take into account that the actual system (12) can be modelled by the following neural form

$$\dot{x} = Ax + X_f W_f^* S_f(x) + X_g W_g^* S_g(x) u \quad (21)$$

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

$$e = \hat{x} - x \quad (22)$$

Then from (20) and (21) we obtain the error equation

$$\dot{e} = Ae + X_f \tilde{W}_f S_f(x) + X_g \tilde{W}_g S_g(x) u \quad (23)$$

Where $\tilde{W}_f = W_f - W_f^*$ and $\tilde{W}_g = W_g - W_g^*$. Regarding the identification of W_f and W_g in (20) we are now able to state the following theorem.

Theorem 2: Consider the identification scheme given by (23). The learning law

a) For the elements of W_{f_i}

$$\dot{w}_{f_i}^{pl}(x) = -\bar{x}_{f_i}^p p_i e_i s_l(x) \qquad (24)$$

b) For the elements of W_{g_i}

$$\dot{w}_{g_i}^{p}(x) = -\bar{x}_{g_i}^p p_i e_i u_i s_i(x) \qquad (25)$$

or equivalently $\dot{W}_{g_i}(x) = -\left(X_{g_i}\right)^T p_i e_i u_i s_i(x)$ with $i = 1,\dots,n$, $p = 1,\dots,m$ and $l = 1,\dots,k$ guarantees the following properties.

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

Proof: Consider the Lyapunov function candidate,

$$V\left(e, \tilde{W}_f, \tilde{W}_g\right) = \frac{1}{2} e^T P e + \frac{1}{2\gamma_1} tr\left\{\tilde{W}_f{}^T \tilde{W}_f\right\} + \frac{1}{2\gamma_2} tr\left\{\tilde{W}_g{}^T \tilde{W}_g\right\}$$

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

$$PA + A^T P = -I$$

Taking the derivative of the Lyapunov function candidate we get

$$\dot{V}\left(e, \tilde{W}_f, \tilde{W}_g\right) = \frac{1}{2}\dot{e}^T P e + \frac{1}{2} e^T P \dot{e} + \\ + \frac{1}{\gamma_1} tr\left\{\dot{\tilde{W}}_f{}^T \tilde{W}_f\right\} + \frac{1}{\gamma_2} tr\left\{\dot{\tilde{W}}_g{}^T \tilde{W}_g\right\}$$

which after substituting Eq. (23) becomes

$$\dot{V} = \frac{1}{2} e^T \left(A^T P + PA\right) e + \left(\frac{1}{2} e^T P X_f \tilde{W}_f S_f + S_f^T \tilde{W}_f^T X_f^T P e\right) + \\ + \left(\frac{1}{2} e^T P X_g \tilde{W}_g S_g u + u^T S_g^T \tilde{W}_g^T X_g^T P e\right) + \frac{1}{\gamma_1} tr\left\{\dot{\tilde{W}}_f{}^T \tilde{W}_f\right\} + \\ + \frac{1}{\gamma_2} tr\left\{\dot{\tilde{W}}_g{}^T \tilde{W}_g\right\}$$

Now since $e^T P X_f \tilde{W}_f S_f$ and $e^T P X_g \tilde{W}_g S_g u$ are scalars, we have that

$$e^T P X_f \tilde{W}_f S_f = \left(S_f^T \tilde{W}_f^T X_f^T P e\right)^T$$

$$e^T P X_g \tilde{W}_g S_g u = \left(u^T S_g^T \tilde{W}_g^T X_g^T P e\right)^T$$

Therefore, $\dot{V}$ becomes

$$\dot{V} = -\frac{1}{2} e^T e + e^T P X_f \tilde{W}_f S_f + e^T P X_g \tilde{W}_g S_g u + \\ + \frac{1}{\gamma_1} tr\left\{\dot{\tilde{W}}_f{}^T \tilde{W}_f\right\} + \frac{1}{\gamma_2} tr\left\{\dot{\tilde{W}}_g{}^T \tilde{W}_g\right\}$$

We assume that

$$\frac{1}{\gamma_1} tr\left\{\dot{\tilde{W}}_f{}^T \tilde{W}_f\right\} = -e^T P X_f \tilde{W}_f S_f$$

$$\frac{1}{\gamma_2} tr\left\{\dot{\tilde{W}}_g{}^T \tilde{W}_g\right\} = -e^T P X_g \tilde{W}_g S_g u$$

For extracting the adaptive law of the weights. Then, taking into account the form of W_f and W_g the above equations result in the element wise learning laws given in (24), (25). These laws can also be written in the following compact form

$$\dot{W}_f = -\gamma_1 X_f^T P e S_f^T \qquad (26)$$

$$\dot{W}_g = -\gamma_2 X_g^T P E U S_g^T \qquad (27)$$

Where E and U are diagonal matrices such that $E = diag\left(e_1,\dots,e_n\right)$ and $U = diag\left(u_1,\dots,u_n\right)$. Finally, the Lyapunov function candidate results in

$$\dot{V} = -\frac{1}{2} e^T e \le 0$$

Since $\dot{V}$ is negative semi definite then we conclude that $V \in L_\infty$, which implies that $e, \hat{x}, \tilde{W}_f, \tilde{W}_g \in L_\infty$. Furthermore, $W_f = \tilde{W}_f + W_f^*$, $W_g = \tilde{W}_g + W_g^*$ are also bounded. Since V is a non-increasing function of time and bounded from below, the $\lim_{t\to\infty} V = V_\infty$ exists; therefore, by integrating $\dot{V}$ from 0 to ∞ we have

$$\int_0^\infty \|e\|^2 \, dt \le \left[V(0) - V_\infty\right] < \infty$$

which implies that $e \in L_2$.

Since $e \in L_2 \cap L_\infty$, using Barbalat's Lemma we conclude that $\lim_{t\to\infty} e(t) = 0$.

Now, using the bounded ness of u, S_f, S_g and the convergence of $e(t)$ to zero, we have that $\dot{\tilde{W}}_f$, $\dot{\tilde{W}}_g$ also converges to zero [16].

VI. Simulation Results

Our aim is to test the performance of the proposed F-HONNF scheme in approximating a time-varying model. This can be equivalently measured by monitoring the fast tracking of any weight changes performed in the F-RHONN representation of the time varying unkown system.

We assume the existence of only parametric uncertainty. Therefore we select an initial second order model of the form (21) with optimal weights W_f^* and W_g^* shown in the following tables 1 and 2 respectively, and membership centre values randomly selected. The inputs assume the persistently exciting form $u = 1 + 0.8\sin\left(0.001t\right)$.

Table 1

0.6669	0.7848	0.2271	0.3132	0.8488
0.1654	0.5776	0.3831	0.6733	0.0048
0.6068	0.5640	0.1076	0.9665	0.6119
0.0152	0.6603	0.8959	0.4103	0.1181
0.5613	0.3616	0.3845	0.4701	0.5335
0.3254	0.3147	0.2569	0.9380	0.0337
0.7551	0.6089	0.3899	0.6808	0.9781
0.7822	0.5819	0.8589	0.7045	0.9365
0.5822	0.4470	0.8773	0.8601	0.3356
0.5127	0.9641	0.3694	0.9679	0.7290

Table 2

0.9993	0.6491
0.3032	0.5357
0.4929	0.6519
0.8544	0.6480
0.3693	0.1582
0.4829	0.2186

The weights of the model change during the simulation every 1 second by multiplying the weights with a constant value (in this example the weights are doubled every second).

In the sequel, we used our Fuzzy-RHONN approach given with equation (20) and the appropriate adaptive laws for the weights, in order to capture the initial model and monitored its performance during its changes.

For that purpose, we selected a second order Fuzzy-HONNF with learning rates $g_1 = 0.01$, $g_2 = 50$ and the parameters of the sigmoidal terms being $a_1 = 0.1, a_2 = 6$, $b_1 = b_2 = 1$ and $c_1 = c_2 = 0$. The evolution of one state is shown in figure 4. It can be observed that by following the proposed approach we can have a quite perfect approximation of the initial system despite the changes.

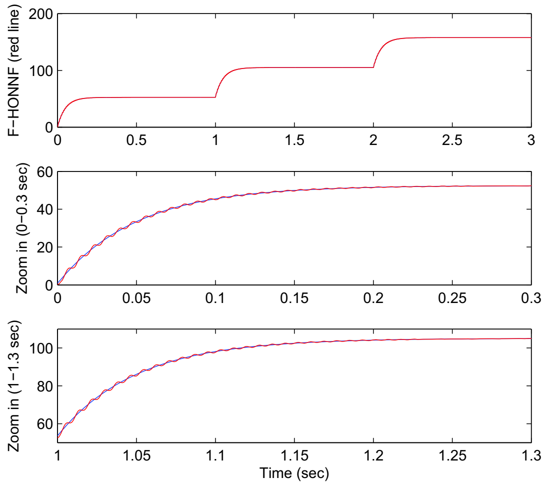

Figure 1. Approximation of one state of the initial Fuzzy-HONNF model with the proposed approach

VII. CONCLUSIONS

The weight tracking performance in the identification of unknown nonlinear dynamical systems was presented in this paper. The identification scheme is based on a new definition of Adaptive Fuzzy Systems (AFS) operating in conjunction with High Order Neural Network Functions (F-HONNFs). Under this scheme the identification is driven to a Fuzzy-Recurrent Higher Order Neural Network, which however takes into account the fuzzy output partitions of the initial AFS. The proposed scheme does not require a-priori expert's 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 in tracking changes performed in a time varying nonlinear artificial model.

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.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.

[4] 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.

[5] 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.

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

[7] 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.

[8] 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.

[9] 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.

[10] 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.

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

[12] 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.

[13] 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.

[14] 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.

[15] 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.

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

[17] 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.

[18] J. T. Spooner, M. Maggiore, R. Ordonez, and K. M. Passino, *Stable Adaptive Control and Estimation for Nonlinear Systems: Neural and Fuzzy Approximator Techniques*, NY: John Wiley and Sons, 2002.

[19] 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.

[20] 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, pp. 82-93.

[21] Theodoridis, D. C., Christodoulou, M. A. & Boutalis, Y. S., “Indirect Adaptive Neuro - Fuzzy Control Based On High Order Neural Network Function Approximators”, *in Proc. 16th Mediterranean Conference on Control and Automation - MED08*, Ajaccio, Corsica, France, 2008, pp. 386-393.

[22] Boutalis, Y. S., Theodoridis, D. C. & Christodoulou, M. A., “A new Neuro FDS definition for indirect adaptive control of unknown nonlinear systems using a method of parameter hopping”, *IEEE Trans. On Neural Networks*, to be published.

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

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