Identification of Unknown Nonlinear Systems based on Multilayer Neural Networks and Lyapunov Theory

José A. R. Vargas
Department of Electrical Engineering
University of Brasilia
Brasília, Brazil
vargas@unb.br

Emerson Grzeidak
Central Bank of Brazil
Foreign Reserves Department
Accounting and Settlement Division
Brasília, Brazil
emerson.grzeidak@bcb.gov.br

Sadek C.A. Alfaro
Department of Mechanical Engineering
University of Brasilia
Brasília, Brazil
sadek@unb.br

***Abstract*— This paper considers the identification problem of nonlinear systems based on single-hidden-layer neural networks (SHLNNs) and Lyapunov theory. A nonlinearly parameterized neural model, whose weights are adjusted by robust adaptive laws, which are designed via Lyapunov theory, is proposed for ensuring the convergence of the residual state error to an arbitrary neighborhood of zero. In addition, a scaling matrix is used to resize the unknown nonlinearities to be approximated by an SHLNN, which, in turn, provides a simple way to shape the residual state error. It is shown that all estimation errors are uniformly bounded and, in addition, that the residual state error is uniformly ultimately bounded with an ultimate bound that depends directly on some independent design parameters. To validate the theoretical results, the identification of a chaotic system and a comparison study with other work in the literature are performed.**

Keywords—Online identification, Lyapunov methods, Multilayer neural networks

I. INTRODUCTION

It is well known that the mathematical characterization is, often, a prerequisite to observer and controller design. However, in some circumstances, the characterization of the dominant dynamics can be a difficult or even impossible task. In this scenario, the use of online approximators as, for instance, neural networks (NNs) is a possible alternative to parameterization. Basically, the unknown nonlinearities in the system are replaced by NN models, which have a known structure but unknown weights. In the case of supervised learning, the unknown weights are estimated by using an error signal between the outputs of the actual system and the neural identification model.

Commonly-employed neural identification models are linearly and nonlinearly parameterized, which can, by nature, be static or dynamic. Their weights are often adjusted using gradient-based schemes, such as the backpropagation algorithm, or their robust modifications [1-22]. The most widely-used robust modifications in neuro-identification are the σ, switching-σ, ε_1, parameter projection, and dead zone [1-22], which avoid the parameter drift.

For instance, in [2], the neuro-identification of a general class of uncertain continuous-time dynamical systems was proposed, and a σ-modification adaptive law for the weights of recurrent high-order neural networks (RHONNs) was chosen to ensure that the state error converges to the neighborhood of zero. More recently, in [3]-[5], neuro-identification schemes for open loop systems were proposed. In [3]-[4] the conditions to ensure the asymptotical convergence of the residual state error to zero were established, even in the presence of approximation error and bounded internal or external perturbations. The convergence of the state error to zero in both works ([3]-[4]) was based, among other assumptions, on the previous knowledge of bounds for the approximation error and perturbations, which are usually unknown in practice. In [5], an identification scheme based on a dynamical neural model with scaling and a robust weight adaptive law was proposed. The main peculiarity of [5] is that the residual state error is directly related to two design matrices, which allow the residual state error to be arbitrarily and easily reduced.

Despite the remarkable theoretical contribution in these works ([2]-[15]), they are all based on linearly parameterized neural networks and consequently, in general, suffer from "the curse of dimensionality". That is, these models have a poor capability of interpolation and require a large number of basic functions to deal with multi-dimensional inputs. This drawback can be alleviated by using identification models based on SLHNNs. See, for instance, [16]-[22]. In these works, the presence the two weight matrices to be estimated, approximation errors, and perturbations, however, make the problem challenging.

For example, in [16], an online approximator of multi-input multiple output static functions based on SHLNNs is proposed. In [17], a robust scheme based on SHLNNs to identify nonlinear systems was proposed. The weight adaptation laws were based on modified backpropagation algorithms. By using the Lyapunov`s direct method, it was shown that all errors are uniformly bounded and the residual state error converges to a ball whose radius can be reduced by setting some design parameters in adequate values. Nevertheless, the design parameters related to the performance are dependent, and therefore, arbitrary small residual state error could not be achieved. Another disadvantage of [17] is that, due to static approximations assumed in the definition of the adaptive laws, the identification process may not converge in the presence of high frequency perturbations. In [21]-[22], the discrete case is considered and the stability properties of the approximation errors are presented.

In this paper, we extended the results in [16] in order to identify dynamical systems based on SHLNNs. All conditions are established to ensure the convergence of the residual state error to an arbitrary neighborhood of zero, even in the presence of approximation error and internal or external perturbations. Also, the dependence between the residual state error and some independent design parameters is straightforward. Consequently, the residual state error can be arbitrarily and easily reduced. Furthermore, it is not necessary to have any previous knowledge about the ideal weight, approximation error and disturbances, in contrast to [3]-[4]. In addition, the designed methodology is structurally simple, since it does not use a dynamic feedback gain or bounding function employed in [3]. To provide stability, the weight adaptation laws are chosen based on Lyapunov theory. Simulation experiments are performed to illustrate the effectiveness of the proposed method.

II. Single Hidden Layer Neural Networks

A class of multilayer NNs used here can be expressed mathematically in matrix form as

$$g_{nn}(W,V,x,u) = W\sigma\left(Vz\right) \tag{1}$$

where $V \in \Re^{n_2\times(n_1+1)}$, n_1 is the number of neurons from the input layer, n_2 is the number of neurons in the hidden layer, $W \in \Re^{n\times n_2}$, $\sigma \in \Re^{n_2}$ is a basis function vector and $z = \left[x_1,\cdots,x_n,u_1,\cdots,u_m,1\right] \in \Re^{n_1+1}$. In standard SHLNNs each entry of $\sigma(\cdot)$ is a linear combination of either an external input or the state passed through a scalar activation function $s(\cdot)$. Commonly used $s(\cdot)$ are the sigmoid and hyperbolic tangent function ([1]). In this paper, we combine SHLNNs with high-order neural networks (HONN) for allowing high order interactions between neurons in the hidden layer. The approximation capacity of RHONN and its superior storage capacity has been shown in several studies (see [1-2] for further details).

Universal approximation results in [1]-[2] indicate that:

Property 1: Let $s(\cdot)$ be a non-constant, bounded and monotone increasing continuous function. Let Ω_z be a compact subset of $\Re^n$, and $g(z)$ be a real valued continuous function on Ω_z. Then for any arbitrary constant $\mu > 0$, there exist an integer n_2 and ideal matrices W^* and V^* such that

$$\max_{z\in\Omega_z}\left|g\left(z\right) - g_{nn}\left(W^*,V^*,z\right)\right| < \mu \tag{2}$$

Based on Property 1, we have

$$g(z) = W^*\sigma(V^*z) + \varepsilon(z) \tag{3}$$

with $\varepsilon(z)$ satisfying $\max_{z\in\Omega_z}\left|\varepsilon(z)\right| < \mu,\ \forall z \in \Omega_z$.

III. Problem Formulation

Consider the following nonlinear differential equation

$$\dot{x} = F\left(x,u,v,t\right),\quad x\left(0\right) = x_0 \tag{4}$$

where $x \in X$ is the *n*-dimensional state vector, $u \in U$ is a *m*-dimensional admissible input vector, $v \in V \subset \Re^q$ is a vector of time varying uncertain variables and $F: X\times U\times V\times\left[0,\infty\right) \mapsto \Re^n$ is a continuous map. In order to have a well-posed problem, we assume that X, U, V are compact sets and F is locally Lipschitzian with respect to x in $X\times U\times V\times\left[0,\infty\right)$, such that (4) has a unique solution.

We assume that the following can be established:

Assumption 1: On a region $X\times U\times V\times\left[0,\infty\right)$

$$\left\|d\left(x,u,v,t\right)\right\| \le d_0 \tag{5}$$

where

$$d\left(x,u,v,t\right) = F\left(x,u,v,t\right) - f\left(x,u\right) \tag{6}$$

f is an unknown map, d is modelling uncertainty, and $d_0 \ge 0$, is an unknown constant. Note that (5) is verified when x e u evolve on compact sets and the temporal disturbances are bounded.

Hence, except for the Assumption 1, we say that $F(x,u,v,t)$ is an unknown map and our aim is to design a NNs-based identifier for (4) to ensure the state error convergence, which will be accomplished despite the presence of approximation error and disturbances.

IV. Identification Model and State Estimate Error Equation

We start by presenting the identification model and the definition of the relevant errors associated with the problem.

By adding and subtracting Ax, where $A \in \Re^{n\times n}$ is an Hurwitz matrix, (4) can be rewritten as

$$\dot{x} = Ax + g\left(x,u\right) + d\left(x,u,v,t\right) \tag{7}$$

where $g(x,u) = f(x,u) - Ax$.

By using SHLNNs, the nonlinear mapping $g(z)$ can be replaced by $W^*\sigma\left(V^*z\right)$ plus an approximation error term $\varepsilon(x,u)$. More exactly, (4) becomes

$$\dot{x} = Ax + BW^*\sigma\left(V^*z\right) + B\varepsilon\left(x,u\right) + d\left(x,u,v,t\right) \tag{8}$$

where $B \in \Re^{n\times n}$ is a scaling matrix, $W^* \in \Re^{n\times n_2}$ and $V^* \in \Re^{n_2\times(n_1+1)}$ are the "optimal" or ideal matrices, which can be defined as

$$\left(W^*,V^*\right) := \underset{(W,V)}{\arg\min}\left\{\sup_{z\in\Omega_z}\left|W\sigma\left(Vz\right) - g(z)\right|\right\} \tag{9}$$

Let $\hat{W}$ and $\hat{V}$ be the estimates of W^* and V^*, respectively, and the weight estimation errors be

$$\begin{aligned}\tilde{W} &= \hat{W} - W^* \\ \tilde{V} &= \hat{V} - V^*\end{aligned} \tag{10}$$

From (3), the following can be established

Assumption 2: On a compact set Ω_z, the ideal neural network weights and the NN approximation error are bounded by

$$\left\|W^*\right\| \le w_m,\ \left\|V^*\right\| \le v_m,\ \left\|\varepsilon\left(x,u\right)\right\| \le \varepsilon_0 \tag{11}$$

with w_m, v_m and ε_0 being positive constants.

Remark 1: Assumption 1 is usual in identification. Assumption 2 is quite natural since g is continuous and their arguments evolve on compact sets.

Remark 2: It should be noted that W^* and V^* were defined as being the values of $\hat{W}$ and $\hat{V}$ that minimizes the L_∞ - norm difference between $g\left(x,u\right)$ and $\hat{W}\sigma\left(\hat{V}z\right)$. The scaling matrix B from (8) is introduced to manipulate the magnitude of uncertainties and, hence, the magnitude of the approximation error.

Remark 3: By applying the Taylor series expansion of $\sigma\left(V^*z\right)$ about $\hat{V}z$, we have

$$\sigma\left(V^*z\right) = \sigma\left(\hat{V}z\right) - \hat{\sigma}'\tilde{V}z + \Theta \tag{12}$$

where $\hat{\sigma}' = \left.\dfrac{\partial\sigma\left(V^*z\right)}{\partial V^*z}\right|_{\hat{V}z}$ and Θ represents is the high order terms in the Taylor expansion.

The structure (8) suggests an identification model of the form

$$\dot{\hat{x}} = A\hat{x} + B\hat{W}\sigma\left(\hat{V}z\right) - l_0\tilde{x} - l \tag{13}$$

where $l_0 > 0$, l is a vector function to be defined afterwards, $\hat{x}$ is the estimated state, and $\tilde{x} := \hat{x} - x$ is the state estimation error. It will be demonstrated that the identification model (13) used in conjunction with convenient adjustment laws for $\hat{W}$ and $\hat{V}$, to be proposed in the next section, ensures the convergence of the state error to a neighborhood of the origin, even in the presence of the approximation error and disturbances, whose radius depends on some design parameters.

From (8) and (13), we obtain the state estimation error equation

$$\begin{aligned}\dot{\tilde{x}} = {} & A\tilde{x} + B\hat{W}\sigma\left(\hat{V}z\right) - BW^*\sigma\left(V^*z\right) \\ & - B\varepsilon(x,u) - d(x,u,v,t) - l_0\tilde{x} - l\end{aligned} \tag{14}$$

From (12), we can rewrite the state estimation error equation as

$$\begin{aligned}\dot{\tilde{x}} = {} & A\tilde{x} + B\tilde{W}\sigma\left(\hat{V}z\right) + B\hat{W}\hat{\sigma}'\tilde{V}z \\ & - B\tilde{W}\sigma'\left(\hat{V}z\right)\tilde{V}z - \Lambda - l_0\tilde{x} - l\end{aligned} \tag{15}$$

where $\Lambda = -BW^*\Theta - B\varepsilon(x,u) - d(x,u,v,t)$ is a residual term.

V. Adaptive Laws and Stability Analysis

We now state and prove the main theorem of the paper.

Theorem 5.1: Consider the class of general nonlinear systems described by (4), which satisfies Assumptions 1-2, the identification model (13) with

$$l = \frac{\gamma_0\ \tilde{x}}{\lambda_{\min}(K)\left[\left\|\tilde{x}\right\| + \gamma_1\exp(-\gamma_2 t)\right]} \tag{16}$$

Let the weights adaptation laws be given by

$$\begin{aligned}\dot{\hat{W}} &= -\gamma_W\left[2\alpha_W\left\|\tilde{x}\right\|\left(\hat{W} - W_0\right) + BK\tilde{x}\hat{\sigma}^T - BK\tilde{x}\left(\hat{\sigma}'\hat{V}z\right)^T\right] \\ \dot{\hat{V}} &= -\gamma_v\left[2\alpha_V\left\|\tilde{x}\right\|(\hat{V} - V_0) + \hat{\sigma}'\hat{W}^T BK\tilde{x}z^T\right]\end{aligned} \tag{17}$$

where $\gamma_0 \ge 0, \gamma_1 > 0, \gamma_2 > 0, \gamma_W > 0, \gamma_v > 0, \alpha_W > 0, \alpha_V > 0,$ W_0 and V_0 are constant matrices, K is a matrix such that

$$K = P + P^T \tag{18}$$

and P is a positive definite matrix. Then, if $\gamma_0 = 0$, the estimation errors $\tilde{W}$ and $\tilde{V}$ are bounded, and $\tilde{x}$ is uniformly ultimately bounded with an ultimate bound $\alpha_{\tilde{x}}$. If $\gamma_0 > \alpha_3$, $\alpha_3 > 0$, the state error converges to zero, i.e., $\lim_{t\to\infty} \tilde{x}(t) = 0$.

Proof: Consider the Lyapunov function candidate

$$\bar{V} = \tilde{x}^T P \tilde{x} + \frac{\|\tilde{W}\|_F^2}{2\gamma_w} + \frac{\|\tilde{V}\|_F^2}{2\gamma_v} \quad (19)$$

By evaluating the time derivative of (19) along the trajectories of (15) and (17), we obtain

$$\begin{aligned}\dot{\bar{V}} = & -\tilde{x}^T Q \tilde{x} + \tilde{x}^T KB\tilde{W}\hat{\sigma} - \tilde{x}^T KB\tilde{W}\hat{\sigma}'\tilde{V}z \\ & + \tilde{x}^T KB\hat{W}\hat{\sigma}'\tilde{V}z - \tilde{x}^T K\Lambda \\ & - tr\left\{\tilde{W}^T\left[2\alpha_W \|\tilde{x}\|\left(\hat{W} - W_0\right) + BK\tilde{x}\hat{\sigma}^T\right.\right. \\ & \left.\left. -BK\tilde{x}\left(\hat{\sigma}'\hat{V}z\right)^T\right]\right\} - tr\left\{\tilde{V}^T\left[2\alpha_V \|\tilde{x}\|\left(\hat{V} - V_0\right)\right.\right. \\ & \left.\left. +\hat{\sigma}'\hat{W}^T BK\tilde{x}z^T\right]\right\} - l_0\tilde{x}^2 - \tilde{x}^T Kl\end{aligned} \quad (20)$$

where $A^T P + PA = -Q$, -Q is a Hurwitz matrix, $\dot{\tilde{W}} = \dot{\hat{W}}$ and $\dot{\tilde{V}} = \dot{\hat{V}}$, since $\dot{W}^* = 0$ and $\dot{V}^* = 0$.

Furthermore, by using the following representations

$$\begin{aligned} & tr\left\{\tilde{W}^T BK\tilde{x}\hat{\sigma}^T\right\} = \tilde{x}^T KB\tilde{W}\hat{\sigma} \\ & tr\left\{\tilde{W}^T BK\tilde{x}\left(\hat{\sigma}'\hat{V}z\right)^T\right\} = \tilde{x}^T KB\tilde{W}\hat{\sigma}'\hat{V}z \\ & tr\left\{\tilde{V}\hat{\sigma}'\hat{W}^T BK\tilde{x}z^T\right\} = \tilde{x}^T KB\hat{W}\hat{\sigma}'\tilde{V}z \end{aligned} \quad (21)$$

and further rearranging terms, (20) results

$$\begin{aligned}\dot{\bar{V}} = & -\tilde{x}^T Q \tilde{x} - \tilde{x}^T K\Lambda - tr\left\{\tilde{W}^T\left[2\alpha_W \|\tilde{x}\|\left(\hat{W} - W_0\right)\right]\right\} \\ & - tr\left\{\tilde{V}^T\left[2\alpha_V \|\tilde{x}\|\left(\hat{V} - V_0\right)\right]\right\} + \tilde{x}^T KB\tilde{W}\hat{\sigma}'V^* z \\ & - l_0\tilde{x}^2 - \tilde{x}^T Kl\end{aligned} \quad (22)$$

In the case that $\gamma_0 = 0$, by considering the facts

$$\begin{aligned} & -\tilde{V} + \hat{V} = V^*, \\ & 2tr\left[\tilde{W}^T\left(\hat{W} - W_0\right)\right] = \|\tilde{W}\|_F^2 + \left\|\left(\hat{W} - W_0\right)\right\|_F^2 - \left\|\left(W^* - W_0\right)\right\|_F^2, \\ & 2tr\left[\tilde{V}^T\left(\hat{V} - V_0\right)\right] = \|\tilde{V}\|_F^2 + \left\|\left(\hat{V} - V_0\right)\right\|_F^2 - \left\|\left(V^* - V_0\right)\right\|_F^2, \\ & \|\hat{\sigma}'(t)\| \le \sigma_{0d},\ \|\Lambda(t)\| \le \Lambda_0, \|z(t)\| \le z_0, \forall t \ge 0 \text{ and for some} \\ & \text{positive constants } \sigma_{0l}, \Lambda_0 \text{ and } z_0. \end{aligned} \quad (23)$$

and by completing the square, (22) implies

$$\dot{\bar{V}} = -\|\tilde{x}\|\left[l_0\|\tilde{x}\| + \alpha_W\left(\|\tilde{W}\|_F - \frac{\alpha_1}{2\alpha_W}\right)^2 + \alpha_W\|\tilde{V}\|_F^2 - \alpha_0\right] \quad (24)$$

where $\alpha_0 = \alpha_1^2/4\alpha_W + \alpha_2 + \alpha_W\|W^* - W_0\|_F^2 + \alpha_W\|V^* - V_0\|_F^2$, $\alpha_1 = z_0\sigma_{0d}\|KB\|_F\|W^*\|_F$ and $\alpha_2 = \Lambda_0\|K\|_F$.

Hence, $\dot{\bar{V}} < 0$ outside the compact set $\Omega = \left\{\left(\tilde{x}, \tilde{W}, \tilde{V}\right) \mid \|\tilde{x}\| \le \alpha_{\tilde{x}}\ or \|\tilde{W}\|_F \le \alpha_{\tilde{W}}\ or \|\tilde{V}\|_F \le \alpha_{\tilde{V}}\right\}$ where $\alpha_{\tilde{x}} = \alpha_0/l_0$, $\alpha_{\tilde{W}} = \left(\alpha_0/l_0\right)^{1/2} + \alpha_1/2\alpha_W$ and $\alpha_{\tilde{V}} = \left(\alpha_0/\alpha_V\right)^{1/2}$. Thus, since $\alpha_{\tilde{x}}$, $\alpha_{\tilde{W}}$ and $\alpha_{\tilde{V}}$ are positive constants, by employing usual Lyapunov arguments [24], we concluded that all error signals are uniformly bounded. In addition, since l_0, $\alpha_{\tilde{W}}$ and $\alpha_{\tilde{V}}$ can be arbitrarily selected, $\tilde{x}(t)$ is uniformly ultimately bounded with an ultimate bound $\alpha_{\tilde{x}}$.

In the case that $\gamma_0 > \alpha_3$, (22) implies

$$\dot{\bar{V}} \le -\left[\lambda_{\min}(Q) + l_0\right]\|\tilde{x}\|^2 - \frac{\alpha_3\gamma_1}{\eta}\frac{\left[\|\tilde{x}\| - \eta\exp(-\gamma_2 t)\right]}{\|\tilde{x}\| + \gamma_1\exp(-\gamma_2 t)} \quad (25)$$

where $\alpha_3 = \alpha_1 + \alpha_2 + \alpha_W\|W^* - W_0\|_F^2 + \alpha_V\|V^* - V_0\|_F^2$ and $\eta = \gamma_1\alpha_3/(\gamma_0 - \alpha_3)$.

Define now

$$\Omega = \left\{\left(\tilde{x}, \tilde{W}\right) \mid \|\tilde{x}(t)\| \le \eta\exp(-\gamma_2 t)\right\} \quad (26)$$

Note that the numerator in the bracket of (25) is greater than zero for $\|\tilde{x}\| > \eta\exp(-\gamma_2 t)$ (or $\tilde{x} \in \Omega^c$); hence,

$$\dot{\bar{V}} \le -\left[\lambda_{\min}(Q) + l_0\right]\|\tilde{x}\|^2 \quad (27)$$

Further, since $\bar{V}$ is bounded from below and non-increasing with time, we have

$$\lim_{t\to\infty}\int_0^t \|\tilde{x}(\tau)\|^2 d\tau \le \frac{\bar{V}(0) - \bar{V}_\infty}{\lambda_{\min}(Q) + l_0} < \infty \quad (28)$$

where $\lim_{t\to\infty} \bar{V}(t) = \bar{V}_\infty < \infty$. Notice that, based on (15), with the bounds on $\tilde{x}, \tilde{W}, \tilde{V}$, and l, $\dot{\tilde{x}}$ is also bounded. Thus, $\dot{\bar{V}}$ is uniformly continuous. Hence, by applying the Barbalat's Lemma [24], we conclude that $\lim_{t\to\infty} \tilde{x}(t) = 0$ for all $\tilde{x} \in \Omega^c$.

Once the synchronization error $\tilde{x}(t)$ has entered Ω, it will remain in Ω forever, due to (26). Consequently, we conclude that $\lim_{t\to\infty} \tilde{x}(t) = 0$ holds in the large, i.e., whatever the initial value of $\left(\tilde{x}(t), \tilde{W}(t), \hat{V}(t)\right)$ (inside or outside Ω).

Corollary 1: Consider the class of general nonlinear systems described by (4), which satisfies Assumptions 1-2, the identification model (13), (16)-(17) with $\gamma_2 = 0$. Then, the state error $\tilde{x}(t)$ converges to the residual set

$$\Xi = \left\{ \tilde{x} \;\; \middle| \left\| \tilde{x}(t) \right\| \le \gamma_1 \alpha_t \right\}$$

where $\alpha_t = \exp(-\gamma_2 t_s)$ and t_s is the time in which the exponential function in (16) is turned off.

Remark 4: Corollary 1 establishes an interesting peculiarity of the proposed method. The exponential function used in the identification model can be turned off when the residual state error has entered into any desired neighborhood of the origin. It is important to overcome numerical errors that can appear when the exponential function on the right-hand side of (16) has practically decayed to zero.

VI. SIMULATION

Consider the unified chaotic system [23], which is described by

$$\begin{aligned} \dot{x} &= (25\alpha + 10)(y - x) \\ \dot{y} &= (28 - 35\alpha)x - xz + (29\alpha - 1)y \\ \dot{z} &= xy - \left(\frac{8+\alpha}{3}\right)z \end{aligned} \tag{29}$$

where x, y and z are state variables and it is always chaotic in the whole interval $\alpha \in [0,1]$. It should be also noted that system (29) becomes the Lorenz system for $\alpha = 0$ and the Chen system for $\alpha = 1$. In the following simulations, we consider the Chen system.

To identify the chaotic system (29), the proposed identification model (13) and the adaptive law (16) and (17) were implemented. The design parameters were chosen as $\gamma_V = 1$, $\gamma_W = 0.02$, $l_0 = 0.001$, $\alpha_W = 0.5$, $\alpha_V = 0.5$, $s(\cdot) = 85/\left[1 + \exp{-1(\cdot)}\right]$,

$$A = \begin{bmatrix} -7.8 & 0 & 0 \\ 0 & -7.8 & 0 \\ 0 & 0 & -7.8 \end{bmatrix}, \; B = \begin{bmatrix} 121 & 0 & 0 \\ 0 & 127{,}6 & 0 \\ 0 & 0 & 143 \end{bmatrix}$$

$P = 0.05I$, $V_0 = 0$ and $W_0 = 0$.

The chosen initial conditions for the system and the identification model are $x(0) = 2$, $y(0) = 1$, $z(0) = 2$, $\hat{x}(0) = 5$, $\hat{y}(0) = 5$, $\hat{z}(0) = 5$, $\hat{W}(0) = 0$ and $\hat{V}(0) = 0$.

In the simulations of the proposed algorithm, the design matrices *A*, *B*, and *P* were initially chosen as identity matrices. In the sequence, these values were adjusted, by a trial and error procedure.

To check the robustness of the proposed method, we consider the emergence at $t = 5$s of disturbances of the form:

$$\begin{aligned} d(x,u,v,t) &= 3\sin(t)\left(x^2 + y^2 + z^2\right)^{1/2} + 50\sin(200t) \\ &\quad + 10\cos(400t) \end{aligned} \tag{30}$$

To illustrate the advantages of the proposed methodology, the identification model introduced in [17] is used here for comparison. Consider the online identification multilayer neural network algorithm proposed in [17] described as

$$\begin{aligned} \dot{\hat{x}} &= A\hat{x} + \hat{W}\sigma\left(\hat{V}\hat{\bar{x}}\right) \\ \dot{\hat{W}} &= -\eta_1 \left(\tilde{x}^T A^{-1}\right)^T \left(\sigma\left(\hat{V}\hat{\bar{x}}\right)\right)^T - \rho_1 \left\|\tilde{x}\right\| \hat{W} \\ \dot{\hat{V}} &= -\eta_2 \left(\tilde{x}^T A^{-1} \hat{W} \left(I - \bar{\Lambda}\left(\hat{V}\hat{\bar{x}}\right)\right)\right)^T \hat{\bar{x}}^T - \rho_2 \left\|\tilde{x}\right\| \hat{V} \end{aligned} \tag{31}$$

where $\bar{\Lambda}\left(\hat{V}\hat{\bar{x}}\right) = diag\left\{\sigma_i^2\left(\hat{V}_i\hat{\bar{x}}\right)\right\}, i = 1,2,...,m.$ The design parameters are chosen as $\eta_1 = 25$, $\eta_2 = 0.4$, $\rho_1 = 0.00012$, $\rho_2 = 0.00012$,

$$A = \begin{bmatrix} -0.0078 & 0 & 0 \\ 0 & -0.0078 & 0 \\ 0 & 0 & -0.0078 \end{bmatrix} \text{ and } s(\cdot) = {}^{500}\!/\!_{\left[1+\exp{-0.5(\cdot)}\right]} .$$

Other design parameters and initial conditions were chosen as before.

The performances in the estimation of state variables are shown in Fig. 1-3. It can be seen that the simulations confirm the theoretical results, that is, the algorithm is stable and the residual state error is small. From Figs. 1-5, it can be concluded that the identification scheme is robust in the presence of perturbation without, practically, any degradation of performance.

Figs. 1-3 show the state error norms comparisons for each state variable. It should be pointed out that the adjustment of the design parameters in [17] was not trivial, perhaps due to the mutual dependence between the design matrices *P* and *Q*.

The comparison between the Frobenius norms associated with the estimated weight matrices W and V are shown in Figs. 4-7. After a transient phase, due to large initial uncertainty, these norms seem to converge in our case, indicating that most of the state estimation error has been removed.

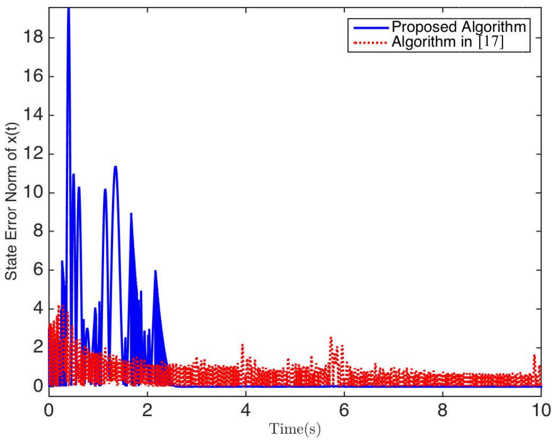

Fig. 1: Performance comparison in the estimation of x.

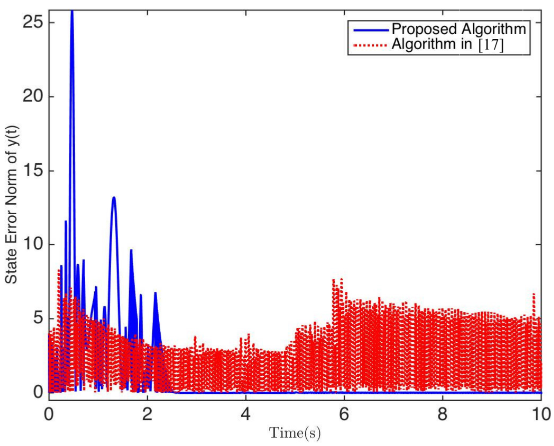

Fig. 2: Performance comparison in the estimation of y.

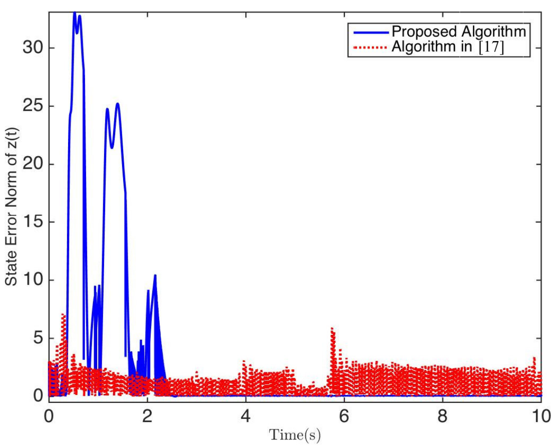

Fig. 3: Performance comparison in the estimation of z.

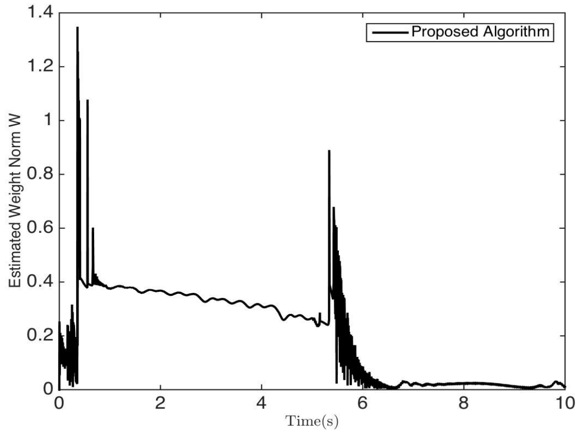

Fig. 4: Frobenius norm of the estimated weight matrix W.

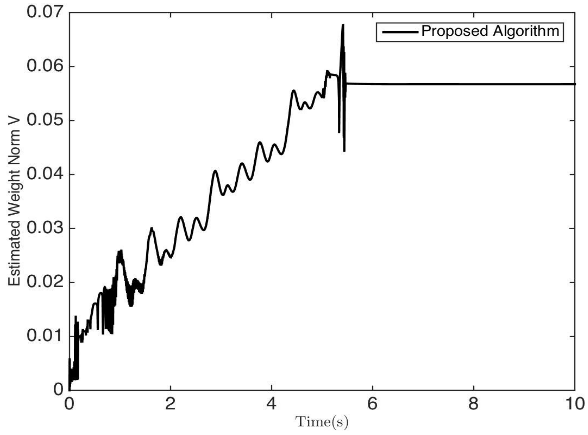

Fig. 5: Frobenius norm of the estimated weight matrix V.

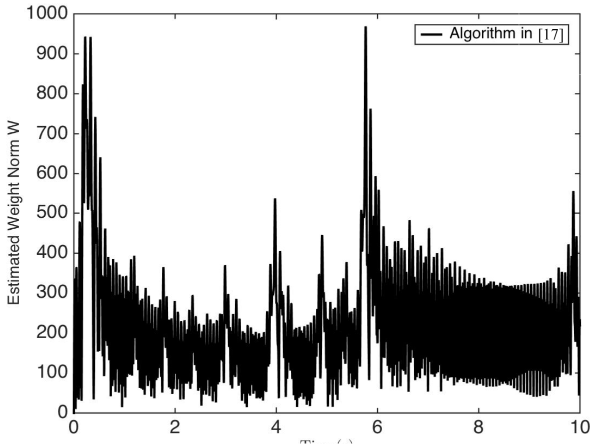

Fig. 6: Frobenius norm of the estimated weight matrix W.

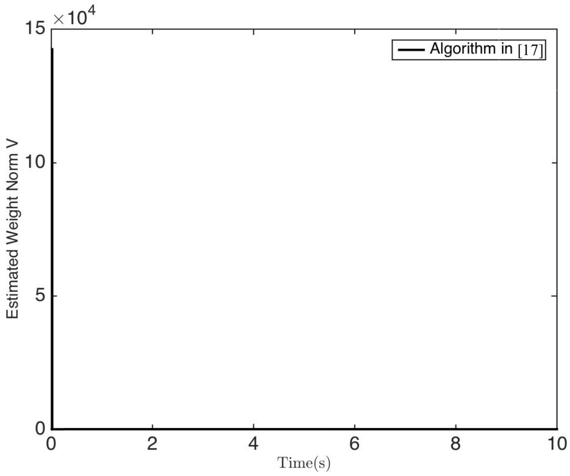

Fig. 7: Frobenius norm of the estimated weight matrix V.

VII. CONCLUSION

In this paper, we proposed a novel identification scheme for the approximation of nonlinear dynamical systems. The scheme is based on SHLNNs, to parameterize the unknown nonlinearities, whose weights are adjusted by adaptive laws designed using Lyapunov theory. It was shown that the residual state error can be adjusted via independent design constants. The use of a scaling matrix to adjust the size of the unknown nonlinearities and an SHLNN, with an activation vector function with high order terms, allowed for the easy manipulation of the residual error performance. Simulation results were performed to show the effectiveness and performance of the proposed approach in the presence of perturbations.

REFERENCES

[1] S.S. Ge, C.C. Hang, T.H. Lee, and X. Li, *Stable adaptive neural network control*, Boston, MA: Kluwer, 2001.

[2] E.B. Kosmatopoulos, M.M. Polycarpou, M.A. Christodoulou, and P.A. Ioannou, "High-order neural network structures for identification of dynamical systems," *IEEE Trans. Neural Netw.*, vol. 6, no. 2, pp. 422–431, 1995.

[3] J.A.R. Vargas, E.M. Hemerly, "Neural adaptive observer with asymptotic convergence in the presence of time-varying parameters and disturbances," (in Portuguese) *Sba Controle Autom.*, vol. 19, no. 1, pp. 18–29, 2008. doi:10.1590/S0103-17592008000100002

[4] J.A.R. Vargas, E.M. Hemerly, and E.R.L. Villarreal, "Stability analysis of a neuro-identification scheme with asymptotic convergence," *Int. J. Artif. Intell. Appl.*, vol. 3, no. 4, pp. 35-50, 2012.

[5] J.A.R. Vargas, K.H.M. Gularte, and E.M. Hemerly, "On-line neuro-identification of uncertain systems based on scaling and explicit feedback," *Control Autom. Electr. Syst.*, vol. 24, pp. 753–763, 2013.

[6] I. Chairez, "Neural networks network nonparametric identifier with normalized learning laws," *IEEE Trans. Neural Netw. Learn. Syst.*, doi: 10.1109/TNNLS.2015.2505090, to be published.

[7] J.H. Pérez-Cruz, I. Chairez, and J. Pacheco, "Identification and control of class of non-linear systems with non-symmetric deadzone using recurrent neural networks," *IET Control Theory Appl.*, vol. 8, no. 3, pp. 183-192, 2014.

[8] V.G. Lopez, A.Y. Alanis, E.N. Sanchez, and J. Rivera, "Real-time implementation of neural optimal control and state estimation for a linear induction motor," *Neurocomputing*, vol. 152, pp. 403-412, 2015.

[9] A.Y. Alanis, J.D. Rios, J. Rivera, N. Arana-Daniel, and C. Lopez-Franco, "Real-time discrete neural control applied to a Linear Induction Motor," *Neurocomputing*, vol. 164, pp. 240-251, 2015.

[10] K.J. Gurubel, A.Y. Alanis, E.N. Sánchez, and S. Carlos-Hernandez, "A Neural observer with time-varying learning rate: analysis and applications," *Int. J. Neural Syst.*, vol. 24, no. 1, pp. 1-15, 2014.

[11] J. H., Pérez-Cruz, J.J. Rubio, J. Pacheco, and E. Soriano, "State estimation in MIMO nonlinear systems subject to unknown deadzones using recurrent neural networks," *Neural Comput. Appl.*, vol. 25, no. 3-4, pp. 693-701, 2014.

[12] Z.J. Fu, W.F. Xie, J. Na, "Robust adaptive nonlinear observer design via multi-time scales neural network," *Neurocomputing*, vol. 190, no. 217-225, 2016.

[13] J.A. Vargas, K.H. Gularte, and E.M. Hemerly, "An improved on-line neuro-identification scheme," in *Proc. 2012 UKACC IEEE*, Cardiff, UK, Sep 2012, pp. 1088-1093.

[14] H.A. Abbas, B. M. Belkheiri, and B. Zegnini, "Feedback linearisation control of an induction machine augmented by single-hidden layer neural networks," *Int. J. Control*, vol. 89, no. 1, pp. 140-155, 2016.

[15] V.M. Janakiraman and D. Assanis, "Lyapunov method based online identification of nonlinear systems using extreme learning machines," *Computing Research Repository* (CoRR):1211.1441: pp. 1-8, 2012.

[16] A. Bazaei and M. Moallem, "Online neural identification of multi-input multi-output systems," *IET Control Theory Appl.*, vol.1, no 1, pp. 44-50, 2007.

[17] F. Abdollahi, H. A. Talebi, and R.V. Patel, "Stable Identification of Nonlinear Systems Using Neural Networks: Theory and Experiments," *Trans. Mechatronics*, vol. 11, no. 4, pp. 488-495, 2006.

[18] S. Bhasin, R. Kamalapurkar, H.T. Dinh, and W.E. Dixon, "Robust identification-based state derivative estimation for nonlinear systems," *IEEE Trans. Autom. Control*, vol. 58, no. 1, pp. 187-192, 2013.

[19] D.D. Zheng, Z.J. Fu, and W.F. Xie, "Indirect adaptive control of nonlinear system via dynamic multilayer neural networks with multi - time scales," *Int. J. Adapt. Control Sign. Proc.*, vol. 29, no. 4, pp. 505-523, 2015.

[20] K. Esfandiari, F. Abdollahi, and H.A. Talebi, "Stable adaptive output feedback controller for a class of uncertain non-linear systems," *IET Control Theory Appl.*, vol. 9, no. 9, pp. 1329-1337, 2015.

[21] J.D.J. Rubio, P. Angelov, and J. Pacheco. "Uniformly stable backpropagation algorithm to train a feedforward neural network," *IEEE Trans. Neural Netw.,* vol. 22, no. 3, pp. 356-366, 2011.

[22] W. Yu and J.D.J. Rubio, "Recurrent neural networks training with stable bounding ellipsoid algorithm," *IEEE Trans. Neural Netw.*, vol. 20, no. 6, pp. 983-991, 2009.

[23] J. Lü, G. Chen, D. Cheng, and S. Celikovsky, "Bridge the gap between the Lorenz system and the Chen system," *Int. J. Bifurcat. Chaos*, vol. 12, no. 12, pp. 2917-2926, 2002.

[24] P.A. Ioannou, J. Sun, *Robust adaptive control*, Mineola, NY: Dover Publications, Inc., 2012.