STABILITY ANALYSIS OF A NEURO-IDENTIFICATION SCHEME WITH ASYMPTOTIC CONVERGENCE

José A. R. Vargas [1], Elder M. Hemerly[2] and Elmer R. L. Villarreal[3]

[1]Department of Electrical Engineering, Universidade de Brasília, Brasília, DF, Brazil
vargas@unb.br
[2]Department of System and Control, Instituto Tecnológico de Aeronáutica, SP, Brazil
hemerly@ita.br
[3]Department of Exact and Natural Sciences, Universidade Federal Rural do Semiárido, RN, Brazil
elmerllanos@ufersa.edu.br

ABSTRACT

This paper focuses on the stability and convergence analysis of a neuro-identification scheme for uncertain nonlinear systems. Based on linearly parameterized neural networks and the previous knowledge of upper bounds for the approximation error and disturbances, a robust modification of the descent gradient algorithm is proposed to make the overall identification process stable, and, in addition, the on-line residual prediction error asymptotically null, despite the presence of approximation error and disturbances. A simulation study to show the application and comparative performance of the proposed algorithm is presented.

KEYWORDS

On-line identification, nonlinear estimators, uncertain systems, neural networks.

1. INTRODUCTION

System neuro-identification is important not only to predict the behavior of the system, but also for providing an appealing system parameterization, which can later be used in the synthesis of control algorithms, since mathematical characterization is often a prerequisite to observer and controller design, see for instance [1-10] and the references therein.

Neural identification models usually employed are the dynamic ones, being their weights mainly adjusted using gradient and backpropagation algorithms or their robust modifications [3-10]. Most used robust modifications in neuro-identification are the σ, switching-σ, ε_1, parameter projection, and dead zone [3-10] which avoid the parameter drift. Nevertheless, these modifications, at present, can not ensure that the prediction error converges asymptotically to zero in the presence of approximation error and bounded disturbances.

For instance in [3], the 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

DOI : 10.5121/ijaia.2012.3403

neighborhood of the origin, which can be arbitrarily reduced by setting a sufficiently large number of high-order connections in the RHONN model.

In [4], dynamic NNs with a gradient descent algorithm for weights adjustment were used to identify a general class of uncertain nonlinear systems. Under the crucial assumption that the unknown system can be exactly modeled by a NN model, that is, the disturbance and approximation errors are identically null, it was shown that the state error converges asymptotically to zero, whereas the weight errors remain bounded or, in case that the hidden-layer activation function signal is persistently exciting, the weight errors converge to zero.

Similarly, also others relevant works, such as [5-10], showed that the dead zone, modified $\delta-$rule, and ε_1-modification and others robust modifications can be used in weight adjustment laws to make the entire identification process stable in the presence of approximation error and disturbances.

Although the assumption of free disturbance may be interesting from a theoretical point of view, from a practical perspective it is a restrictive assumption since the presence of approximation errors are, in general, unavoidable, since the structure of the unknown system and the neural model are unrelated. It is well-known that adaptive laws designed for the disturbance or modeling error free case may suffer from parameter drift [11]. In fact, this lack of robustness in adaptive systems in the presence of unmodeled dynamics or bounded disturbances was reported in the early 1980s. Several robust modifications to counteract this [11] have been proposed since then.

Hence, in this paper we propose a robust modification for the weight adaptive law in neuro-identification problems which ensures, in contrast to previous works, that the residual prediction error converges to zero in the presence of approximation error and disturbances, provided that some conditions on the design parameters are satisfied. The adaptive law consists of a leakage modification of a standard gradient descent algorithm. However, in contrast to commonly leakage modifications [3-10] which aim at stability in the presence of approximation errors and disturbances, we introduce the leakage term here for, in addition to stability, ensuring that the residual prediction error converges to zero. More precisely, it is show by using usual Lyapunov arguments and an adaptive bounding technique [12] that the residual prediction error converges asymptotically to zero, whereas the others error signals remain bounded.

Throughout the paper $tr(\cdot)$ denotes the trace operator, $\lambda_{\min}(\cdot)$ denotes the minimum eigenvalue operator, $\|\cdot\|$ denotes the 2-norm and $\|\cdot\|_F$ denotes the Frobenius norm.

2. LINEARLY PARAMETERIZED NEURAL NETWORKS

Linearly parameterized neural networks (LPNNs) can be expressed mathematically as

$$\rho_{nn}(W,\zeta)=W\pi(\zeta) \tag{1}$$

where $W\in\Re^{n\times L_\rho}$, $\zeta\in\Re^{L_\zeta}$, $\pi:\Re^{L_\zeta}\mapsto\Re^{L_\rho}$ is the so-called basis function vector, which can be considered as a nonlinear vector function whose arguments are preprocessed by a scalar function $\pi(\cdot)$, and n, L_ρ, L_ζ are integers strictly positive. Commonly used scalar functions $\pi(\cdot)$ include sigmoid, tanh, gaussian, Hardy's, inverse Hardy's multiquadratic, etc [3,13]. However, here we are only interested in the class of LPNNs for which $\pi(\cdot)$ is bounded, since in this case we have,

$$\|\pi(\zeta)\| \le \pi_0 \tag{2}$$

being π_0 a strictly positive constant.

The class of LPNNs considered in this work includes HONN [3], RBF networks [13], wavelet networks [14], and also others linearly parameterized approximators as Takagi-Sugeno fuzzy systems [15]. Universal approximation results in [3,13,14,15] indicate that:

***Property 1*:** Given a constant $\varepsilon_0 > 0$ and a continuous function $f : \Omega \mapsto \Re^n$, where $\Omega \subset \Re^{L_\zeta}$ is a compact set, there exists a weight matrix $W = W^*$ such that the output of the neural network architecture (where L_ρ may depend on ε_0 and f) satisfies

$$\sup_{\zeta\in\Omega}\left| f(\zeta) - W^*\pi(\zeta)\right| \le \varepsilon_0 \tag{3}$$

where $|\cdot|$ denotes the absolute value if the argument is a scalar. If the argument is a vector function in $\Re^n$ then $|\cdot|$ denotes any norm in $\Re^n$.

3. PROBLEM FORMULATION

Consider the following nonlinear differential equation

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

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 $F : X \times U \times V \times [0,\infty) \mapsto \Re^n$ 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 $X \times U \times V \times [0,\infty)$, such that (4) has a unique solution.

We assume that the following can be established

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

$$\|h(x,u,v,t)\| \le h_0 \tag{5}$$

where

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

f is an unknown map, *h* are internal or external disturbances, and $\bar{h}_0$, such that $\bar{h}_0 > h_0 \ge 0$, is a known constant.

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 asymptotical state error convergence, that is, the state error converges asymptotically to zero, which will be accomplished despite the presence of approximation error and disturbances.

4. 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.

Let $\bar{f}$ be the best known approximation of *f*, $B \in \Re^{n \times n}$ a scaling matrix defined as $B = diag(b_i)$, $b_i \neq 0$, $\bar{g} = B^{-1} g$, and $g(x,u) = f(x,u) - \bar{f}(x,u)$. Then, by adding and subtracting $\bar{f}(x,u)$, (4) can be rewritten as

$$\dot{x} = \bar{f}(x,u) + B\bar{g}(x,u) + h(x,u,v,t) \quad (7)$$

Remark 1: It should be noted that if the designer has no previous knowledge of *f*, then $\bar{f}$ is simply assumed as being the zero vector.

From (3), by using LPNNs, the nonlinear mapping $\bar{g}(x,u)$ can be replaced by $W^*\pi(x,u)$ plus an approximation error term $\varepsilon(x,u)$. More exactly, (7) becomes

$$\dot{x} = \bar{f}(x,u) + BW^*\pi(x,u) + B\varepsilon(x,u) + h(x,u,v,t) \quad (8)$$

where $W^* \in \Re^{nxL}$ is an "optimal" or ideal matrix, which can be defined as

$$W^* := \arg\min_{\hat{W} \in \Gamma} \left\{ \sup_{\substack{x \in X, \\ u \in U}} \left| \bar{g}(x,u) - \hat{W}\pi(x,u) \right| \right\} \quad (9)$$

with $\Gamma = \left\{ \hat{W} \;\middle|\; \|\hat{W}\| \leq \alpha_{\hat{W}} \right\}$, $\alpha_{\hat{W}}$ is a strictly positive constant, $\hat{W}$ is an estimate of W^*, and $\varepsilon(x,u)$ is an approximation error term, corresponding to W^*, which can be defined as

$$\varepsilon(x,u) := \bar{g}(x,u) - W^*\pi(x,u) \quad (10)$$

The approximation, reconstruction, or modeling error ε is a quantity that arises due to the incapacity of LPNNs to match the unknown map $\bar{g}(x,u)$. Since *X*, *U* are compact sets and from (2), the following can be established

Assumption 2: On a region $X \times U$, the approximation error is upper bounded by

$$\|\varepsilon(x,u)\| \leq \varepsilon_0 \quad (11)$$

where $\bar{\varepsilon}_0$, such that $\bar{\varepsilon}_0 > \varepsilon_0 \geq 0$, is an known constant.

Remark 2: Assumption 1 is usual in identification or robust control literature. Assumption 2 is quite natural since $\bar{g}$ is continuous and their arguments evolve on compact sets. The previous knowledge of upper bounds for approximation error and disturbances is common in the robust on-line parameter estimation literature. For instance, the dead zone algorithm uses a previous knowledge of bounds for the approximation errors, as can be seen in [5,6], or disturbances, as reported in [11].

Remark 3: Note that any $\bar{\pi}_0 > \pi_0$, $\bar{h}_0 > h_0$, and $\bar{\varepsilon}_0 > \varepsilon_0$ also satisfy (2), (5), and (11). Hence, to avoid confusion, we define π_0, h_0, and ε_0 to be the smallest constants such that (2), (5), and (11) are satisfied.

Remark 4: It should be noted that W^* and $\varepsilon(x,u)$ might be nonunique. However, $\|\varepsilon(x,u)\|$ is unique from (9).

Remark 5: It should be noted that W^* was defined as being the value of $\hat{W}$ that minimizes the L_∞- norm difference between $\bar{g}(x,u)$ and $\hat{W}\pi(x,u)$. Hence, the scaling matrix *B* in (7) is introduced to manipulate the magnitude of $\bar{g}(x,u)$, and therefore of $\|W^*\|_F$, since any increasing of $|b_i|$ implies that the corresponding $|\bar{g}_i(x,u)|$ decreases and, eventually, from (9), that $\|W^*\|_F$ decreases too. The matrix *B* provides an additional degree of freedom for shaping the transient performance.

Remark 6: Notice that the proposed neuro-identification scheme is a black-box methodology, hence the external disturbances and approximation error are related. Based on the system input and state measurements, the uncertain system (including the disturbances) is parameterized by a neural network model plus an approximation error term. However, the parameterization (8) is motivated by the fact that neural networks are not adequate for approximating external disturbances, since the basis function depends on the input and states, whereas the disturbances depend on the time and external variables. The aim for presenting the uncertain system in the form (8), where the disturbance *h* is explicitly considered, is also to highlight that the proposed scheme is in addition valid in the presence of unexpected changes in the systems dynamics that can emerge, for instance, due to environment change, aging of equipment or faults.
The structure (8) suggests an identification model of the form

$$\dot{\hat{x}} = \bar{f}(x,u) + B\hat{W}\pi(x,u) - L(\hat{x} - x) \tag{12}$$

where $\hat{x}$ is the estimated state, $L \in \Re^{n\times n}$ is a positive definite feedback gain matrix introduced to attenuate the effect of the nonzero uncertainties and the initial condition x_0. It will be demonstrated that the identification model (12) used in conjunction with a convenient adjustment law for $\hat{W}$, to be proposed in the next section, ensures the asymptotic convergence of the state error to zero, even in the presence of the approximation error and disturbances.

Remark 7: It should be noted that in our formulation, the LPNN is only required to approximate $B^{-1}\left[f(x,u) - \bar{f}(x,u)\right]$ (whose magnitude is often small) instead of the entire function $B^{-1}\left[f(x,u)\right]$. Hence, standard identification methods (to obtain some previous $\bar{f}$) can be used together with the proposed algorithm to improve performance.

By defining the state estimation error as $\tilde{x} := \hat{x} - x$, from (8) and (12), we obtain the state estimation error equation

$$\dot{\tilde{x}} = -L\tilde{x} + B\tilde{W}\pi(x,u) - B\varepsilon(x,u) - h(x,u,v,t) \qquad (13)$$

where $\tilde{W} := \hat{W} - W^*$.

5. ADAPTIVE LAWS AND STABILITY ANALYSIS

Once the identification model and the relevant error equations associated with the problem of estimation were presented, the following step is the design of adaptive laws for the weights to achieve desired stability and convergence properties. Typically, in the literature [3-10], this is performed based on Lyapunov-like analysis via an enlargement process, where the error equation previously determined is used. Therefore, suitable weight adaptive laws are chosen to make $\dot{V}$, a time derivative of a Lyapunov function candidate V, negative semi-definite outside a ball whose radius is proportional to the worst-case of approximation error. However, this procedure ensures only practical stability. In this section we remove the aforementioned drawback by using dynamic leakage gain to make the Lyapunov derivative negative semi-definite in all error space. This choice is motivated by the fact that dynamic leakage gains can be considered as bounding functions and then can be used to dominate positive terms in $\dot{V}$ and hence improve the performance. Dynamic leakage gains have been used previously in [16].

Before presenting the main theorem, we state a fact, remark and lemma, which will be used in the stability analysis.

Fact 1: Let $W^*, W_0, \hat{W}, \tilde{W} \in \Re^{n \times L_\rho}$ and $\overline{C} \in \Re^{n \times n}$ be a diagonal matrix such that $\overline{C}^T \overline{C} = C$, where $C = diag(c_i)$, $c_i > 0$. Then, with the definition of $\tilde{W} = \hat{W} - W^*$, the following equalities are true:

$$2tr\left[\tilde{W}^T C\left(\hat{W} - W_0\right)\right] = \left\|\overline{C}\tilde{W}\right\|_F^2 + \left\|\overline{C}\left(\hat{W} - W_0\right)\right\|_F^2 - \left\|\overline{C}\left(W^* - W_0\right)\right\|_F^2$$

$$2tr\left[\hat{W}^T W_0\right] = \left\|\hat{W}\right\|_F^2 + \left\|W_0\right\|_F^2 - \left\|\hat{W} - W_0\right\|_F^2 \qquad (14)$$

Remark 8: The first equality in (14) leads to the following inequality:

$$2tr\left[\tilde{W}^T C\left(\hat{W} - W_0\right)\right] \geq c_{i\min}\left\|\tilde{W}\right\|_F^2 + c_{i\min}\left\|\hat{W} - W_0\right\|_F^2 - c_{i\max}\left\|W^* - W_0\right\|_F^2 \qquad (15)$$

where $c_{i\max} = \max(c_i)$ and $c_{i\min} = \min(c_i)$.

Lemma 5.1: Let a scalar bounding function be given by

$$\dot{\hat{\psi}} = -\gamma_\psi \left\|\tilde{x}\right\| \left[2\alpha_1 l\left(\hat{\psi}, \psi^*\right)\hat{\psi} - \alpha_2\left(\left\|\hat{W}\right\|_F^2 + \left\|W_0\right\|_F^2\right) - 2\alpha_1 l\left(\hat{\psi}, \psi^*\right)\psi^*\right] \qquad (16)$$

where

$$l\left(\hat{\psi},\psi^{*}\right)=\frac{2l_0}{\hat{\psi}+\psi^{*}} \tag{17}$$

and $\gamma_{\psi}, l_0, \alpha_1, \alpha_2, \psi^{*} > 0$. Then, subject to the condition

$$\hat{\psi}(0) \geq \bar{\delta}\psi^{*} \tag{18}$$

where $\bar{\delta}=\frac{4\alpha_1 l_0+\alpha_2\|W_0\|_F^2}{4\alpha_1 l_0}$, the bounding function is lower bounded, for all $t \geq 0$, by

$$\hat{\psi}(t) \geq \bar{\delta}\psi^{*} \tag{19}$$

Proof: Consider the Lyapunov-like function

$$V_{\psi} = \hat{\psi}\,\gamma_{\psi}^{-1}\hat{\psi}\,/2 \tag{20}$$

By taking the derivative of (20) along (16) we obtain

$$\dot{V}_{\psi} = -\hat{\psi}\|\tilde{x}\|\left[2\alpha_1 l\hat{\psi}-\alpha_2\left(\left\|\hat{W}\right\|_F^2+\|W_0\|_F^2\right)-2\alpha_1 l\psi^{*}\right] \tag{21}$$

Furthermore, based on (16) and (18) it follows that $\hat{\psi}(t)>0$ for all $t \geq 0$. Then, with the definition (17), the Lyapunov derivative (21) can be lower bounded as

$$\dot{V}_{\psi} \geq -2\alpha_1 l\hat{\psi}\|\tilde{x}\|\left[\hat{\psi}-\bar{\delta}\psi^{*}\right] \tag{22}$$

Hence, if $\hat{\psi} \leq \bar{\delta}\psi^{*}$ we have $\dot{V}_{\psi} \geq 0$, which implies that the bounding function is directed towards the outside or boundary of the region $\left\{\hat{\psi}\ \middle|\ \hat{\psi} \leq \bar{\delta}\psi^{*}\right\}$. Consequently, based on (18), it follows that $\hat{\psi} \geq \bar{\delta}\psi^{*}$ for all $t \geq 0$.

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. Let the weight law be given by

$$\dot{\hat{W}} = -\gamma_W\left\{2C\left(\hat{\psi}-\psi^{*}\right)\left[\hat{W}-\left(I-\alpha_2 C^{-1}\right)W_0\right]\|\tilde{x}\|+BK\tilde{x}\pi^{T}(x,u)\right\} \tag{23}$$

where $\hat{\psi}$ is given by (16), $\gamma_W > 0$, I is an identity matrix, $K=P^{T}+P$, P is the unique positive definite solution of the Lyapunov equation

$$L^{T}P+PL=Q \tag{24}$$

where $L>0$ and $Q>0$. Then, subject to the condition (18), and if

$$\psi^* = \frac{2\alpha_4 \|KB\|_F}{\alpha_1 l_0} \tag{25}$$

$$\alpha_2 < c_{i\min} \tag{26}$$

$$W_0 \in \Omega_0 \tag{27}$$

$$l_0 \in [\beta_1, \beta_2] \tag{28}$$

Where

$$\Omega_0 = \{W_0 \mid tr(W^{*T}W_0) \le 0\},\ \beta_1 = \frac{2c_{i\max}}{\alpha_1}\|W^* - W_0\|_F^2,\ \beta_2 = \frac{\sqrt{c_{i\max}\alpha_2}\|W_0\|_F}{4\alpha_1}\|W^* - W_0\|_F,$$
$$\alpha_4 = \bar{\varepsilon}_0 + \|B^{-1}\|_F \bar{h}_0,\ \alpha_3 = \lambda_{\min}(Q) \tag{29}$$

the error signals $\tilde{x}, \tilde{W}, \tilde{\psi}$ are uniformly bounded and $\lim_{t\to\infty} \tilde{x}(t) = 0$.

Proof: Consider the Lyapunov function candidate

$$V = \tilde{x}^T P\tilde{x} + tr\left(\tilde{W}^T \gamma_W^{-1}\tilde{W}\right)/2 + \tilde{\psi}\gamma_\psi^{-1}\tilde{\psi}/2 \tag{30}$$

where $\tilde{\psi} = \hat{\psi} - \psi^*$.

By evaluating (30) along the trajectories of (13), (16) and (23), and using the representation $tr\left(\tilde{W}^T BK\tilde{x}\pi^T\right) = \tilde{x}^T KB\tilde{W}\pi$, we obtain

$$\begin{aligned}\dot{V} = &-\tilde{x}^T\left(L^T P + PL\right)\tilde{x} - \tilde{x}^T K(B\varepsilon + h)\\ &- 2\tilde{\psi}\|\tilde{x}\|\, tr\left[\tilde{W}^T C\left(\hat{W} - W_0\right)\right] - 2\alpha_2\tilde{\psi}\|\tilde{x}\| tr\left(\tilde{W}^T W_0\right)\\ &- 2\alpha_1 l\tilde{\psi}\hat{\psi}\|\tilde{x}\| + \alpha_2\left(\|\hat{W}\|_F^2 + \|W_0\|_F^2\right)\tilde{\psi}\|\tilde{x}\| + 2\alpha_1 l\psi^*\tilde{\psi}\|\tilde{x}\|\end{aligned} \tag{31}$$

By using Fact 1, the representation $2\tilde{\psi}\hat{\psi} = \tilde{\psi}^2 + \hat{\psi}^2 - {\psi^*}^2$, and (24), the Lyapunov derivative can be written as

$$\begin{aligned}\dot{V} = &-\tilde{x}^T Q\tilde{x} - \tilde{x}^T KB\left(\varepsilon + B^{-1}h\right)\\ &-\tilde{\psi}\|\tilde{x}\|\left[\|\bar{C}\tilde{W}\|_F^2 + \|\bar{C}(\hat{W} - W_0)\|_F^2 - \|\bar{C}(W^* - W_0)\|_F^2\right] + 2\alpha_2\tilde{\psi}\|\tilde{x}\| tr\left(W^{*T}W_0\right)\\ &- \alpha_1 l\left(\tilde{\psi}^2 + \hat{\psi}^2 - {\psi^*}^2\right)\|\tilde{x}\| + \alpha_2\|\hat{W} - W_0\|_F^2\tilde{\psi}\|\tilde{x}\| + 2\alpha_1 l\psi^*\tilde{\psi}\|\tilde{x}\|\end{aligned} \tag{32}$$

Furthermore, by using Remark 8, condition (27), Lemma 5.1, and notation (29), the Lyapunov derivative (32) can be upper bounded as

$$\dot{V} \le \|\tilde{x}\| \cdot \Big[-\alpha_3 \|\tilde{x}\| + \alpha_4 \|KB\|_F - \tilde{\psi}\left(c_{i\min}\|\tilde{W}\|_F^2 + c_{i\min}\|\hat{W} - W_0\|_F^2 - c_{i\max}\|W^* - W_0\|_F^2 \right) - \alpha_1 l\left(\tilde{\psi}^2 + \hat{\psi}^2 - \psi^{*2} \right) + \alpha_2 \|\hat{W} - W_0\|_F^2 \tilde{\psi} + 2\alpha_1 l \psi^* \tilde{\psi} \Big] \tag{33}$$

Further using (26) and rearranging terms, we obtain

$$\dot{V} \le \|\tilde{x}\| \cdot \Big[-\alpha_3 \|\tilde{x}\| - c_{i\min}\tilde{\psi}\|\tilde{W}\|_F^2 - \alpha_1 l \tilde{\psi}^2 + \alpha_4 \|KB\|_F + c_{i\max}\tilde{\psi}\|W^* - W_0\|_F^2 - \alpha_1 l\left(\hat{\psi}^2 - \psi^{*2} \right) + 2\alpha_1 l \psi^* \tilde{\psi} \Big] \tag{34}$$

By employ the definition of ψ^*, see (25), recalling that $\tilde{\psi} = \hat{\psi} - \psi^*$, and using Lemma 5.1, (34) reduces to

$$\dot{V} \le \|\tilde{x}\| \cdot \Big[-\alpha_3 \|\tilde{x}\| - c_{i\min}\tilde{\psi}\|\tilde{W}\|_F^2 - \alpha_1 l \tilde{\psi}^2 + \left(\alpha_1 l_0/2 + c_{i\max}\|W^* - W_0\|_F^2 \right)\hat{\psi} - c_{i\max}\psi^*\|W^* - W_0\|_F^2 - \alpha_1 l \hat{\psi}^2 + \alpha_1 l \psi^{*2} + 2\alpha_1 l \psi^* \tilde{\psi} \Big] \tag{35}$$

which, by using (17), implies

$$\dot{V} \le \|\tilde{x}\| \cdot \Big[-\alpha_3 \|\tilde{x}\| - c_{i\min}\tilde{\psi}\|\tilde{W}\|_F^2 + \left(\alpha_1 l_0/2 + c_{i\max}\|W^* - W_0\|_F^2 \right)\hat{\psi} - c_{i\max}\psi^*\|W^* - W_0\|_F^2 - \frac{2\alpha_1 l_0}{\hat{\psi} + \psi^*}\hat{\psi}^2 + 4\alpha_1 l_0 \psi^{*2}/\hat{\psi} \Big] \tag{36}$$

Thus by using Lemma 5.1 and rearranging terms in (36), we finally obtain

$$\dot{V} \le \|\tilde{x}\| \cdot \Big\{ -\alpha_3 \|\tilde{x}\| - c_{i\min}\tilde{\psi}\|\tilde{W}\|_F^2 - \psi^*\left(c_{i\max}\|W^* - W_0\|_F^2 - \frac{(4\alpha_1 l_0)^2}{\alpha_2 \|W_0\|_F^2} \right) - \frac{\alpha_1 \hat{\psi}^2}{\hat{\psi} + \psi^*}\left[l_0 - \left(l_0/2 + c_{i\max}\|W^* - W_0\|_F^2 / \alpha_1 \right) \right] - \frac{\alpha_1 \hat{\psi}}{\hat{\psi} + \psi^*}\left[l_0 \hat{\psi} - \left(l_0/2 + c_{i\max}\|W^* - W_0\|_F^2 / \alpha_1 \right)\psi^* \right] \Big\} \tag{37}$$

It addition, we note from (28) that

$$\left\|W^* - W_0\right\|_F^2 \geq \frac{(4\alpha_1 l_0)^2}{\alpha_2 c_{i\max}\left\|W_0\right\|_F^2}, \quad \frac{l_0}{2} \geq c_{i\max}\left\|W^* - W_0\right\|_F^2 \Big/ \alpha_1 \tag{38}$$

By substituting (38) into (37), and using Lemma 5.1, we arrive at

$$\dot{V} \leq -\alpha_3 \left\|\tilde{x}\right\|^2 \tag{39}$$

Hence, the error signals $\tilde{x}, \tilde{W}, \tilde{\psi}$ are uniformly bounded. Further, since V is bounded from below and non increasing with time, we have

$$\lim_{t\to\infty} \int_0^t \left\|\tilde{x}(\tau)\right\|^2 d\tau \leq \frac{V(0) - V_\infty}{\alpha_3} < \infty \tag{40}$$

where $\lim_{t\to\infty} V(t) = V_\infty < \infty$. Notice that with the bounds on $\tilde{x}, \tilde{W}, \tilde{\psi}, \varepsilon$ and h, $\left\|\tilde{x}\right\|^2$ is uniformly continuous. Thus from (13), it follows that $\dot{\tilde{x}}$ is bounded. Hence by Barbalat´s lemma [11], we conclude that $\lim_{t\to\infty} \tilde{x}(t) = 0$.

Remark 9: Conditions (18), (24), and (26) are trivial since them are defined by the user according to a desired performance. Condition (25) implies the previous knowledge of upper bounds for the approximation error and disturbances. Conditions (27) and (28) require at least that the sign of some entry of W^* and bounds for the ideal weights are known. The previous knowledge of bounds for the modeling error and ideal weights is not peculiar to the proposed scheme. Most robust modifications in the literature, as for example, switching-σ, parameter projection, and dead-zone require *a priori* information on the plant or modeling error for ensuring stability, as reported in [3,5,6,11]. Despite the aforementioned drawback, the relevance of the proposed algorithm consist in the introduction of a new robust modification and the establishment of conditions, on the design parameters, for which the residual prediction error converges to zero, even in the presence of approximation error and disturbances.

Remark 10: There is at least one way of selecting the design parameters to satisfy the interval condition (28): by selecting a conservative $\left\|W_0\right\|_F$ (large enough) and, in the sequence, by adjusting $\left\|W^* - W_0\right\|_F$ (small enough), what can be achieved by appropriate selection of the scaling matrix *B*.

6. SIMULATIONS

To illustrate the application of the proposed scheme, we consider an engine model operating under idle [17,18] described by

$$\dot{P} = k_p\left(\dot{m}_{ai} - \dot{m}_{ao}\right) \tag{41}$$

$$\dot{N} = k_n\left(T_i - T_l\right) \tag{42}$$

where

$\dot{m}_{ai} = \left(1 + k_{m1}\theta + k_{m2}\theta^2\right)g(P)$,

$\dot{m}_{ao} = -k_{m3}N - k_{m4}P + k_{m5}NP + k_{m6}NP^2$,

$$g(P) = \begin{cases} 1 & P \le 50.6625 \\ 0.0197\sqrt{101.325P - P^2} & P > 50.6625 \end{cases},$$

$T_i = -39.22 + 325024m_{ao} - 0.0112\delta^2 + 0.635\delta$
$+ (2\pi/60)(0.0216 + 0.000675\delta)N - (2\pi/60)^2 0.000102N^2$,

$T_l = (N/263.17)^2 + T_d$, $m_{ao} = \dot{m}_{ao}(t-\tau)/(120N)$, $k_p = 42.40$, $k_n = 54.26$, $k_{m1} = 0.907$, $k_{m2} = 0.0998$, $k_{m3} = 0.0005968$, $k_{m4} = 0.1336$, $k_{m5} = 0.0005341$, $k_{m6} = 0.000001757$,

P is the manifold pressure (kPa),

N is engine speed (rpm),

δ is the spark advance (degrees),

θ is the throttle angle (degrees),

$\dot{m}_{ai}$ is the mass flow rate into the manifold,

$\dot{m}_{ao}$ is the mass flow rate out of the manifold and into the cylinder,

T_d are disturbances which act on the engine as unmeasured accessory torque (N-m)

T_i is the internally developed torque (N.m),

T_l is the load torque made up of accessory torque T_d and shaft torque (N.m),

$g(P)$ is a manifold pressure influence function,

m_{ao} is the air mass in the cylinder, and

τ is a dynamic transport time delay,

The meaning of the main variables of the model is shown in Fig. 1 ([18]).

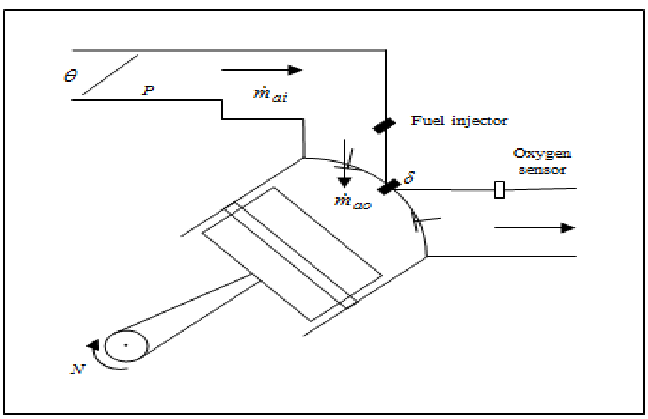

Figure 1- The main engine subsystem.

Since the engine operation is highly nonlinear system, because it presents time delays that vary inversely with engine speed and is time-varying due to aging of components and environment changes such as engine warm-up after a cold star [17], and their dynamical equations entail a great deal of uncertainty [18], the neural parameterization of the engine is relevant and useful. The neural model is valid for any operation points, in contrast to others models derived from steady-state map data and other empirical information.

We define $x = [P \quad N]^T$, $u = [\theta \quad \delta]^T$, and select, as in [19], θ as being a square wave with amplitude 20^{o}, frequency 0.25 rad/s, δ as a sine wave with amplitude 30^{o}, frequency 0.5 rad/s, T_d as a sawtooth wave with amplitude 10 N.m, frequency 0.5 rad/s, $\tau = 0$, and $x(0) = [10 \quad 500]^T$. It should be noted that system (41)-(42) is only used for generation of state trajectory, which is used in the implementation of the algorithm.

To identify the uncertain system (41)-(42) the proposed identification model (12) and the adaptive laws (16) and (23) were implemented. The initial eigenvalues of the gain matrix *L* were freely chosen since the system dynamic is not accessible. The basis function vector was chosen to be very simple, to evaluate the performance of the proposed method in the presence of several approximation errors. Hence, a conservative bound α_4 was selected. In the sequence, by a trial and error procedure, the others design parameters were adjusted for providing an adequate transient and residual state error performance. The design parameter were chosen as

$$\alpha_4 = 3535.53,\ \hat{x}(0) = -[5 \quad 5]^T,\ \hat{W}(0) = 0,\ \hat{\psi}(0) = \bar{\delta}\psi^*,\ \psi^* = 20,\ l_0 = 0.5,\ \bar{f} = 0,$$
$$L = 20I,\ B = I,\ K = 000.1I,\ W_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix},\ \pi = [s(x_1) \quad s(x_2) \quad s(u_1) \quad s(u_2)]^T,$$
$$s(\cdot) = 10/[1 + \exp - 0.5(\cdot)],\ \gamma_W = \gamma_\psi = 1,\ C = 100I,\ \alpha_1 = 1, \text{ and } \alpha_2 = 75. \qquad (43)$$

Notice that the design parameters in (43) satisfy the conditions (18) and (26). It is supposed that condition (25) is satisfied when $\psi^* = 20$. Another arbitrary large upper bound can be obtained by reducing the entries of *K* and, correspondingly, for shaping the transient performance, by increasing the entries of *L*. Hence, condition (25) is mild, since α_4 can be arbitrarily enlarged by reducing *K*, as already mentioned.

Based on (43), condition (28) becomes

$$0.02 \le \left\| W^* - W_0 \right\|_F \le 0.05 \qquad (44)$$

It should be noted that others large upper and small lower bounds for (44) can be obtained by adjusting the design parameters.

The performance in the estimation of the manifold pressure and engine speed are shown in Figures 2-3. We can see that the simulations confirm the theoretical results, that is, the algorithm is stable and the residual state error is small. Figure 4 shows the behavior of the weights, where a logarithmic scale is used for the time axis.

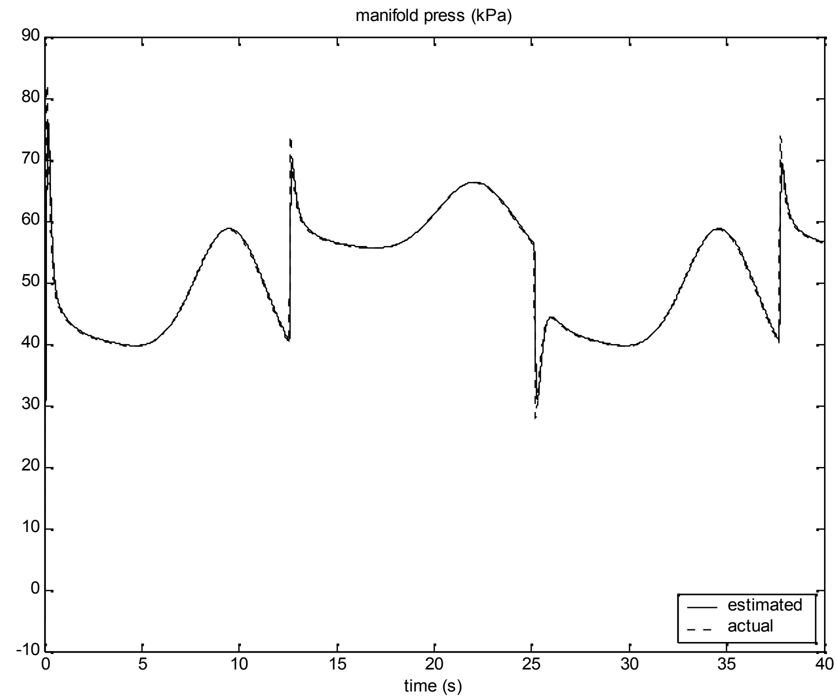

Figure 2- Performance in the estimation of *P*.

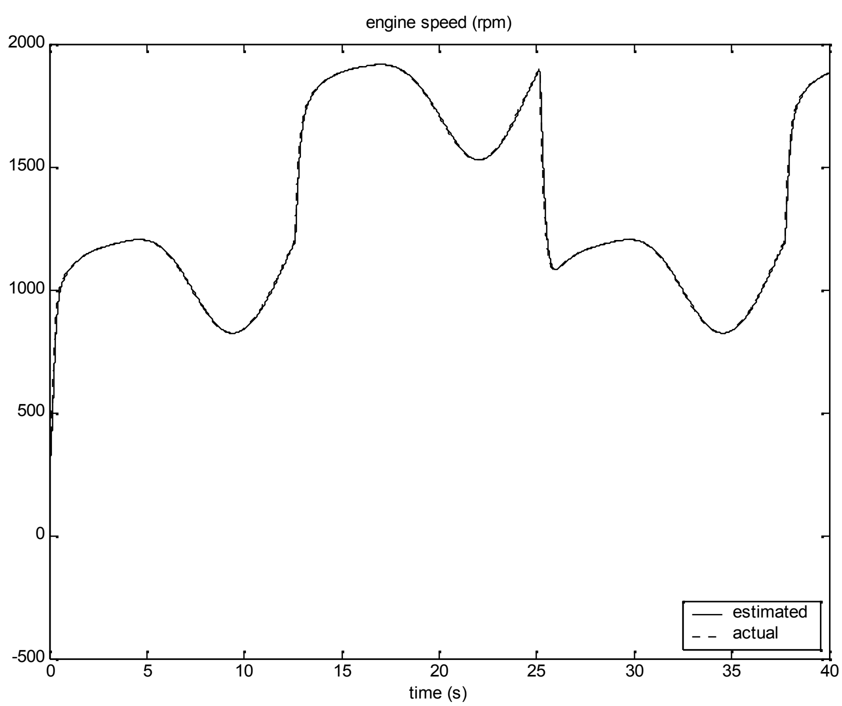

Figure 3- Performance in the estimation of *N*.

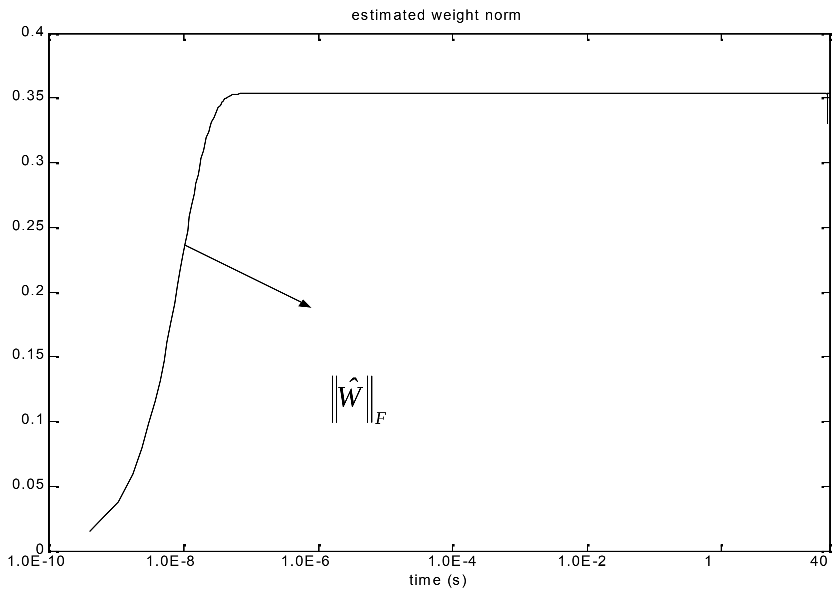

Figure 4- Frobenius norm of the estimated weights.

In Figures 5-6, we compare the performance and robustness, in the presence of uncertainties, of the proposed algorithm and that in [19]. These figures illustrate several points: 1) the proposed algorithm exhibit robustness in the presence of uncertainties, 2) concerning the residual estimation state error, the proposed algorithm presents good performance, and 3) it is not necessary to use arbitrary high gains to obtain arbitrary small residual estimation state errors, as can be seen from the design parameters (43).

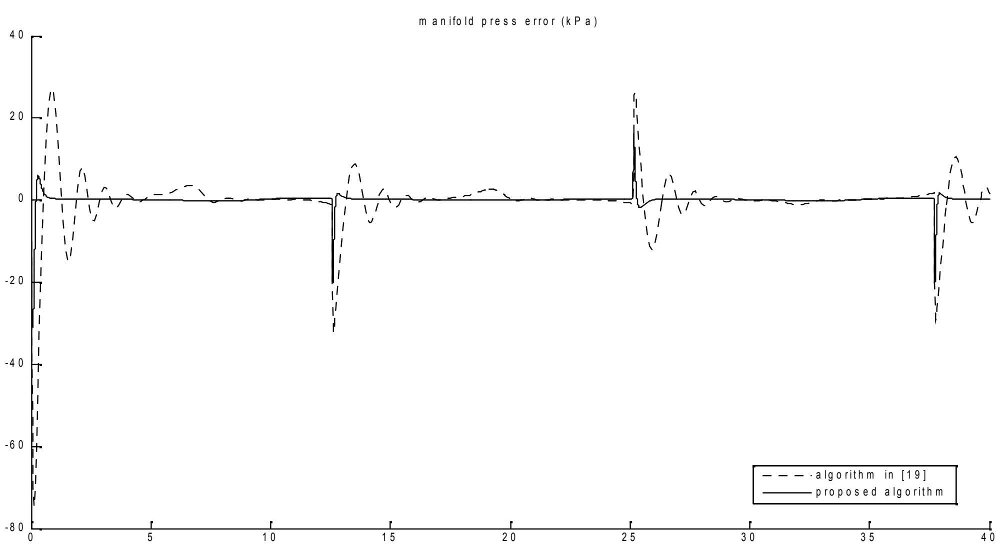

Figure 5- Performance comparison in the estimation of P.

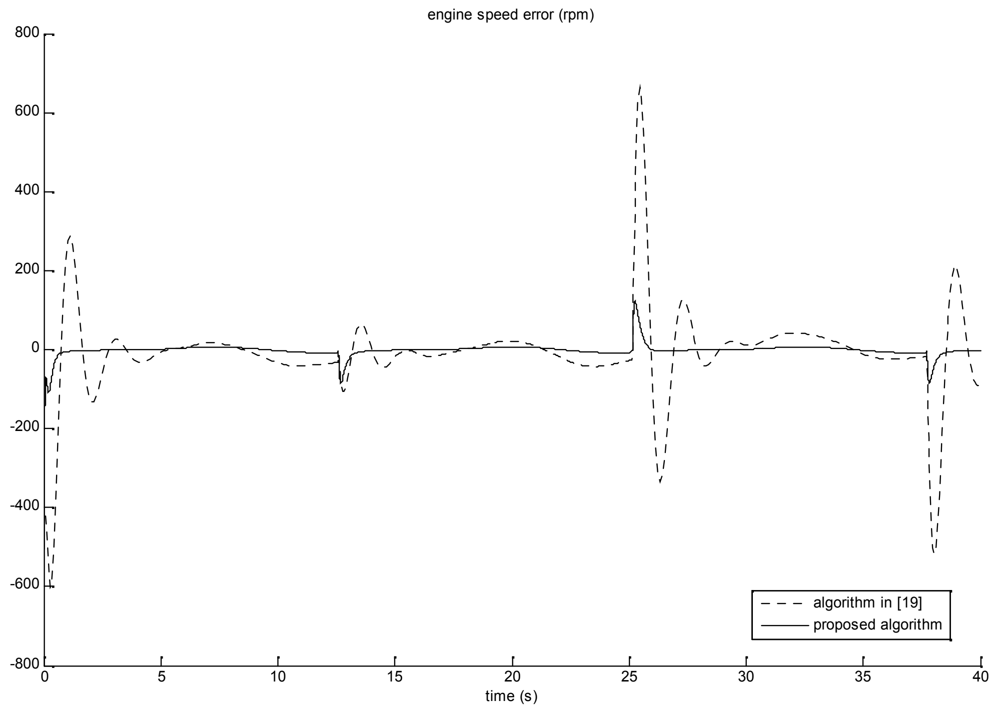

Figure 6- Performance comparison in the estimation of *N*.

7. CONCLUSIONS

In this work, by using Lyapunov-like analysis, we have proved that a robust modification based on a dynamic leakage gain can also ensure residual state error convergence to zero in neuro-identification algorithms, even in the presence of approximation error and disturbances. The proposed algorithm is based on a dynamic ε_1-modification and relies on the previous knowledge of upper bounds for the ideal weight, approximation error and disturbances to ensure asymptotic convergence. Although these bounds are usually unknown in practice, the proposed scheme has stability properties similar to others robust modification since it is based on a ε_1-modification. However, in contrast to previous works, the proposed algorithm has better convergence properties, since it can also guarantee asymptotic convergence even in the presence of approximation error and disturbances, if some conditions on the design parameters are verified.

REFERENCES

[1] K. S. Narendra, K. Parthasarathy, (1990) Identification and control of dynamical systems using neural networks, IEEE Trans. Neural Networks, Vol. 1 no. 1, pp. 4-27.

[2] F. O. Souza, R. M. Palhares, (2010) Interval time-varing delay stability for neural networks, Neurocomputing, Vol. 73, pp. 2789-2792.

[3] E. B. Kosmatopoulos, M. M. Polycarpou, M. A. Christodoulou, P. A. Ioannou, (1995) High-order neural network structures for identification of dynamical systems, IEEE Trans. Neural Networks, Vol. 6, no. 2, pp. 422-431.

[4] R. R. Selmic, F. L. Lewis, (2001) Multimodel neural networks identification and failure detection of nonlinear systems, Proc. 2001 CDC, Orlando, Florida USA, pp. 3128-3133.
[5] Q. Song, (1998) Robust training algorithms of multilayered neural networks for identification of nonlinear dynamical systems, IEE Proc.- Control Theory Appl., Vol. 145, no. 1, pp. 41-46.
[6] S. Jagannathan, F. L. Lewis, (1996) Identification on nonlinear dynamical systems using multilayered neural networks, Automatica, Vol. 32 no. 12, pp. 1707-1712.
[7] F. Abdollahi, H. A. Talebi, R. V. Patel, (2006) Stable identification of nonlinear systems using neural networks: Theory and experiments, IEEE/ASME Trans. on Mechatronics, Vol. 11, no. 4, pp. 488-495.
[8] A. Bazaei, M. Moallem, (2007) Online neural identification of multi-input multi-output systems, IET Control Theory Appl., Vol. 1, no. 1, pp. 44-50.
[9] I. Chairez, A. Poznyak, T. Poznyak, (2009) Stable weights dynamics for a class of differential neural network observer, IEE Proc.- Control Theory Appl., Vol. 3, no. 10, pp. 1437-1447.
[10] A. Y. Alanis, E. N. Sanchez, A. G. Loukianov, E. A. Hernandez, (2010) Discrete-time recurrent high order neural networks for nonlinear identification, Journal of the Franklin Institute, Vol. 347, pp. 1253-1265.
[11] P.A. Ioannou, J. Sun, (1996) Robust adaptive control, Prentice-Hall, Inc., Englewood Cliffs, New Jersey.
[12] M. Corless, (1993) Control of uncertain nonlinear systems, Trans. ASME J. Dyn. Sys., Meas., Control, Vol. 115, pp. 362-372.
[13] S. S. Ge, C. C. Hang, T. H. Lee, T. Zhang, (2002) Stable adaptive neural network control, Kluwer academic publishers.
[14] Q. Zhang and A. Benveniste, (1992) Wavelet networks, IEEE Trans. Neural Networks, Vol. 3, no. 6, pp. 889-898.
[15] L. X. Wang, (1994) Adaptive fuzzy system and control: design and stability analysis, Prentice-Hall, Inc., Englewood Cliffs, New Jersey.
[16] T.Y. Chai, Z. Tao, (1994) A new model reference robust adaptive controller in the presence of unmodeled dynamics and bounded disturbances, Automatica, Vol. 30, no. 5, pp. 865-869.
[17] G. V. Puskorius, L. A. Feldcamp, (1994) Neurocontrol of nonlinear dynamical systems with Kalman filter trained recurrent networks, IEEE Trans. Neural Networks, Vol. 5, no. 2, , pp. 279-297.
[18] G. Vachtsevanos, S. S. Farinwata, D. K. Pirovolou, D. K., (1993) Fuzzy logic control of an automotive engine, IEEE Control Systems Magazine, Vol. 13, no. 3, pp. 62-68.
[19] W. Yu, X. Li, (2001) Some new results on system identification with dynamic neural networks, IEEE Trans. Neural Networks, Vol. 12, no. 2, pp. 412-417.

Authors

José A. R. Vargas is Professor of Control Systems in the Electrical Engineering Department, University of Brasília, DF-Brazil. He received the Dr.Sc. degree in electronics and computer engineering from Instituto Tecnológico de Aeronáutica, in 2003. His current research interests include on-line identification, state estimation and adaptive control.

Elder M. Hemerly is Professor of Control Systems in the Electronics Division, Technological Institute of Aeronautics, SP-Brazil. He received the Ph.D. degree in electrical engineering from Imperial College, London, in 1989. His current research interests include robotics, system identification and adaptive control.

Elmer R. L. Villarreal is Professor of Control Systems in the Exact and Natural Sciences Department, Univesidade Federal Rural do Semi-árido, RN-Brazil. He received the Dr.Sc. degree in electrical engineering from University of Santa Catarina, in 2002. His current research interests include system identification and nonlinear control.