CATT TECHNICAL REPORT No. 02-05-01

ROBUST ADAPTIVE CONTROL OF LINEARIZABLE NONLINEAR SINGLE INPUT SYSTEMS WITH GUARANTEED ERROR BOUNDS*

Haojian Xu and Petros A. Ioannou

Department of Electrical Engineering - Systems
University of Southern California
Los Angeles, CA 90089-2562

Abstract: The design of stabilizing controllers for nonlinear plants with unknown nonlinearities is a challenging problem. The inability to identify exactly the nonlinearities on-line or off-line motivates the design of stabilizing controllers based on approximations or on approximate estimates of the plant nonlinearities. The price paid in such case, could be lack of theoretical guarantees for global stability, and non-zero tracking or regulation error at steady state.

In this paper a nonlinear robust adaptive control algorithm is designed and analyzed for a class of single-input nonlinear systems with unknown nonlinearities. The controller employs a single layer neural network to estimate the unknown plant nonlinearities on-line. The proposed controller is continuous and guarantees closed loop semiglobal stability and convergence of the tracking error to a small residual set. Furthermore it deals with the situation where the estimated plant becomes uncontrollable without any restrictive assumptions. In contrast to previous work on the same subject, the size of the residual tracking error can be specified *a priori* and is guaranteed by choosing certain design parameters. A procedure for choosing these parameters is presented. Two examples are used to demonstrate the performance and properties of the proposed scheme.

Key words: Adaptive control; linearizable systems; nonlinear control; robustness; switching functions.

* This work was supported in part by NSF ECS-9877193 and in part by a collaborative linkage grant from the NATO Cooperative Science and Technology Sub-Programme.

1. INTRODUCTION

The traditional way of designing feedback control system is based on the use of Linear Time Invariant (LTI) models for the plant. Off-line frequency domain techniques could be used to fit such an LTI model to experimental data and identify its parameters. In the case, where the parameters of the LTI model change with time, gain scheduling, on-line parameter identification, adaptive control, robust control techniques etc. are developed over the years to address such situations. The reliance on LTI models for control design purposes often puts limitations on the performance improvement that could be achieved for the plant under consideration. For example if the plant consists of strong nonlinearities, its approximation by an LTI model, may considerably reduce the region of attraction for stability in the presence of disturbances and other modeling uncertainties.

During the recent years, considerable research efforts have been made in the area of stabilizing controllers for classes of nonlinear plants. These efforts are described in detail in a recent survey paper (Kokotovic and Arcak, 2001) where a very elegant and informative historical perspective of the evolution of nonlinear control design is presented and discussed. Most of the recent efforts on nonlinear control design assume that the plant nonlinearities are known. In the special case where the plant nonlinearities are products of unknown constant parameters with known nonlinearities, adaptive control techniques are used to estimate the unknown parameters on-line and use them in the control law (Kosmatopoulos and Ioannou, 1999; Jankovic, 1996; Khalil, 1996; Kristic, *et al*., 1995; Seto, *et al*., 1994; Kanellakopoulos, *et al*., 1991; Sastry and Isidori, 1989; Taylor, *et al.*, 1989). Adaptive neural control techniques have been found to be particularly useful for controlling a wide class of highly uncertain or completely unknown nonlinear systems. In this case, neural networks are used as approximation models of the unknown nonlinearities. The control design is then based on the neural network model rather than the actual system. The approach is similar to the certainty equivalence method used in adaptive control for linear plants. As in the linear case, the estimated plant has to satisfy the usual assumptions of controllability otherwise the control law based on the estimated plant may not exist. We clarify this problem for the widely considered class of nonlinear systems represented by

$$x^{(n)} = f(\boldsymbol{x}) + b(\boldsymbol{x})u \tag{1}$$

where $\boldsymbol{x} = [x(t), \dot{x}(t), \cdots, x^{(n-1)}(t)]^T \in \Re^n$ is the state available for measurement, u is the scalar input and f, b are scalar nonlinear continuous functions of the state vector $\boldsymbol{x}$ and $(\cdot)^{(n)} := d^n(\cdot)/dt^n$. A sufficient condition for controllability is that $|b(\boldsymbol{x})| > 0$, $\forall\, \boldsymbol{x}$. If the estimated plant is of the same form, i.e.,

$$\hat{x}^{(n)} = \hat{f}(\boldsymbol{x},t) + \hat{b}(\boldsymbol{x},t)u \tag{2}$$

and u is calculated based on the estimated plant then for the control law u to exist we require $|\hat{b}(\boldsymbol{x},t)| > 0$, $\forall\, \boldsymbol{x}, t$. The on-line estimators that generate $\hat{f}$, $\hat{b}$ have to guarantee that $|\hat{b}(\boldsymbol{x},t)| > 0$, $\forall\, \boldsymbol{x}, t$. This requirement is not guaranteed by the usual estimators unless special modifications are introduced and in some cases additional assumptions are made. Several efforts have been made to deal with this complication. For example in (Polycarpou, 1996; Lewis *et al.*, 1996), it is assumed that $b(\boldsymbol{x})$=1 is known. This assumption is relaxed in (Polycarpou and Mears, 1998) to $b(\boldsymbol{x}) = g^* > 0$ where g^* is an unknown scalar. In (Jankovic, 1996), $b(\boldsymbol{x})$ is a known function. In (Kosmatopoulos and Ioannou, 1999; Khalil, 1996; Seto, *et al.*, 1994; Kanellakopoulos, *et al.*, 1991; Sastry and Isidori, 1989; Taylor, *et al.*, 1989) it is assumed that $b(\boldsymbol{x}) = (\Theta^b)^T \Psi(\boldsymbol{x})$ where $\Psi(\boldsymbol{x})$ is a known function vector and Θ^b is an unknown parameter vector. In this case, no nonlinear modeling errors are included in $b(\boldsymbol{x})$. The procedures proposed in (Khalil, 1996; Seto, *et al.*, 1994; Kanellakopoulos, *et al.*, 1991; Sastry and Isidori, 1989; Taylor, *et al.*, 1989) are applicable only if both $|b(\boldsymbol{x})| > 0$ and $|[\hat{\Theta}^b(t)]^T \Psi(\boldsymbol{x})| > 0$ where $\hat{\Theta}^b(t)$ denotes the estimate of Θ^b. In order to handle the case where the estimated plant violates the condition $|\hat{b}(\boldsymbol{x},t)| > 0$, Kosmatopoulos and Ioannou (1999) proposed a switching adaptive control strategy to deal with the case where the estimated $\hat{b}(\boldsymbol{x},t) = 0$. A drawback of this algorithm is that the controller is discontinuous. In addition, it requires that $[f(\boldsymbol{x})+2]/b(\boldsymbol{x})$ is upper bounded by a known constant. Several investigators (Rovithakis, 2001; Seshagiri and Khalil, 2000; Yesildirek and Lewis, 1995; Chen and Liu, 1994; Liu and Chen 1993; Chen and Khalil, 1992; Sanner and Slotine, 1992) assume that $b(\boldsymbol{x})$ can be approximated by a neural network as $b^a(\boldsymbol{x}, \Theta^b)$ where Θ^b is a constant vector corresponding to some unknown weights of the neural network. The weights Θ^b are then estimated on line generating $\hat{\Theta}^b(t)$, the estimate of Θ^b at each time t, which in turn is used to generate $b^a(\boldsymbol{x}, \hat{\Theta}^b)$, the estimate of $b^a(\boldsymbol{x}, \Theta^b)$ at each time t. In this approach, the control law based on feedback linearization has $\frac{1}{b^a(\boldsymbol{x}, \hat{\Theta}^b)}$ as a term and exists provided $b^a(\boldsymbol{x}, \hat{\Theta}^b) \neq 0 \quad \forall\, \boldsymbol{x}, t$. In fact for computational purposes and uniform boundedness it is required that $|b^a(\boldsymbol{x}, \hat{\Theta}^b)| > \varepsilon > 0$ $\forall\, \boldsymbol{x}, t$, where ε is a small constant. In this case the problem of $|b^a(\boldsymbol{x}, \hat{\Theta}^b)| \approx 0$ has to be dealt with, in order to establish stability and uniform boundedness. Several attempts have been made to deal with this so called "stabilizability problem" in the linear as well as nonlinear case (Ioannou and Sun, 1996). In (Rovithakis, 2001; Seshagiri and Khalil, 2000; Sanner and Slotine, 1992) $b(\boldsymbol{x})$ is approximated by a single layer neural network as $b(\boldsymbol{x}) = (\Theta^b)^T \Psi(\boldsymbol{x}) + d_b(\boldsymbol{x})$ where $\Psi(\boldsymbol{x})$ is a known basis function vector and $d_b(\boldsymbol{x})$ denotes the approximation error. In (Aloliwi and Khalil, 1997), $b(\boldsymbol{x})$ is of the form $b(\boldsymbol{x}) = (\Theta^b)^T \Psi(\boldsymbol{x}) + d_b(\boldsymbol{x})$ where $d_b(\boldsymbol{x})$ denotes the modeling error. The procedure proposed in (Sanner and Slotine, 1992) assumes that the norms of the gradient of $1/b(\boldsymbol{x})$ and $f(\boldsymbol{x})/b(\boldsymbol{x})$ are bounded by some known nonlinear functions. However, for unknown nonlinearities $f(\boldsymbol{x})$ and $b(\boldsymbol{x})$, such *a priori* knowledge is difficult to obtain. In another approach (Seshagiri and Khalil, 2000; Aloliwi and Khalil, 1997; Polycarpou and Ioannou, 1992) it is assumed that a convex set Ω in the space of Θ^b with $\Theta^b \in \Omega$ is known *a priori* so that $b^a(\boldsymbol{x}, \hat{\Theta}^b) \neq 0 \quad \forall\, \boldsymbol{x}, t$ provided $\hat{\Theta}^b(t) \in \Omega$. Projection is then used to guarantee that $\hat{\Theta}^b(t)$ is inside Ω for all $t \geq 0$. The drawback of this approach is that in general such a convex set is almost impossible to calculate even in the linear case (Ioannou and Sun, 1996). Rovithakis (2001) proposed a projection method equipped with a resetting procedure to guarantee $\left|b^a(\boldsymbol{x}, \hat{\Theta}^b)\right| > \varepsilon$. The procedure switches the sign of the elements in $b^a(\boldsymbol{x}, \hat{\Theta}^b)$ such that all elements have the same sign whenever the condition $\left|b^a(\boldsymbol{x}, \hat{\Theta}^b)\right| > \varepsilon$ is violated. This together with projection guarantees that $\left|b^a(\boldsymbol{x}, \hat{\Theta}^b)\right| > \varepsilon \quad \forall\, t \geq 0$. However this procedure is inherited discontinuous. The discontinuous resetting procedure cannot guarantee the existence and uniqueness of solutions (Polycarpou and Ioannou, 1993). Furthermore it requires that the initial value $\left|b^a[\boldsymbol{x}(0), \hat{\Theta}^b(0)]\right| > \varepsilon$ something that is awkward to establish for a neural network with a large number of nodes. In (Ordonez and Passino, 1999; Spooner and Passino, 1996), $b(\boldsymbol{x})$ is approximated by a fuzzy system as $b^a(\boldsymbol{x}, \Theta^b) = z^T(\boldsymbol{x}) \Theta^b \zeta(\boldsymbol{x})$, where $z(\boldsymbol{x})$ and $\zeta(\boldsymbol{x})$ are known function vectors and Θ^b is an unknown parameter matrix. In this case, the structure of $b^a(\boldsymbol{x}, \hat{\Theta}^b)$ and Ω is defined so that Ω is convex and $\hat{\Theta}^b(t) \in \Omega$ implies that $\left|b^a(\boldsymbol{x}, \hat{\Theta}^b)\right| > 0 \quad \forall\, \boldsymbol{x}, t$. Projection is then used to guarantee that $\hat{\Theta}^b(t)$ is always inside Ω. There is no guarantee however that the unknown Θ^b that corresponds to the "optimal" approximation of $b(\boldsymbol{x})$ by $b^a(\boldsymbol{x}, \Theta^b)$ belongs to Ω. Consequently wrong choices for Ω could lead to large approximation errors and possibly instability in certain cases. We should note that the projection methods in (Ordonez and Passino, 1999; Spooner and Passino, 1996) are discontinuous. These discontinuities need to be analyzed in order to guarantee the existence and uniqueness of the solutions of the closed loop system (Polycarpou and Ioannou, 1993). In (Yesildirek and Lewis, 1995; Chen and Liu, 1994; Liu and Chen 1993; Chen and Khalil, 1992), $f(\boldsymbol{x})$, $b(\boldsymbol{x})$ are approximated by multi-layer neural networks as $f^a(\boldsymbol{x}, \Theta^f)$, $b^a(\boldsymbol{x}, \Theta^b)$ respectively, where Θ^f, Θ^b are constant vectors corresponding to some unknown weights of the neural networks. Multi-layer neural networks may have better approximation abilities than single layer neural networks for the same number of nodes. In this case,

$b^a(\boldsymbol{x},\Theta^b)$ is not linearly dependent on the unknown parameter vector Θ^b. In order to design adaptive laws for this case $b^a(\boldsymbol{x},\Theta^b)$ is linearized. The result of linearization is the introduction of additional modeling error effects that may have an adverse effect on the region of attraction. In (Yesildirek and Lewis, 1995), it requires that the upper bounds on the norms of the unknown parameter vectors Θ^f, Θ^b are known *a priori* and $f(\mathbf{0})=0$. Moreover both the control and adaptive laws are discontinuous. As indicated before, a discontinuous control law leads to a nonlinear system with discontinuities and the existence and uniqueness of solutions cannot be guaranteed (Polycarpou and Ioannou, 1993). Computationally these discontinuities may create chattering at certain boundaries with adverse effects on performance. In (Chen and Liu, 1994; Liu and Chen 1993; Chen and Khalil, 1992), it is assumed that $\hat{\Theta}^b(0)$ is close to the actual values Θ^b and $\hat{\Theta}^b(t)$ is updated slowly by choosing small adaptive gains. Based on this condition it is established that $\hat{\Theta}^b(t)$ stays inside an invariant set where $b^a(\boldsymbol{x},\hat{\Theta}^b) \neq 0$. Since Θ^b is unknown to start with, finding $\hat{\Theta}^b(0)$ close to Θ^b is a difficult task.

Another important issue in adaptive nonlinear control with unknown nonlinearities is that of performance. By performance in this context we mean the size of the region of attraction for signal boundedness and the size of the tracking or regulation error at steady state. Performance issues such as transient behavior are difficult to establish analytically even in the case of known nonlinearities and is not addressed in most of nonlinear control literature at least analytically. In most papers on adaptive nonlinear control with unknown nonlinearities, signal boundedness is established first for some region of attraction within which the assumed neural approximations are valid. Signal boundedness then implies that the approximation or modeling error is also bounded. The upper bound for the tracking or regulation error is then shown to be of the order of the bound on the modeling or approximation error (Rovithakis, 2001; Seshagiri and Khalil, 2000; Aloliwi and Khalil, 1997). In some cases the approximation error is assumed to be upper bounded by some known nonlinearities (Seshagiri and Khalil, 2000; Aloliwi and Khalil, 1997) leading to an upper bound for the tracking error at steady state that is a function of some design parameters. In (Lewis *et al.*, 1996; Yesildirek and Lewis, 1995), the tracking error may be made smaller by increasing the control gain which is a design parameter. It is no clear, however, from the analysis how this gain will affect the region of attraction for signal boundedness. Furthermore, the upper bound for the tracking error cannot be computed and therefore the controller gain cannot be designed *a priori* to achieve a desired tracking error bound at steady state. In (Chen and Liu, 1994; Liu and Chen 1993; Chen and Khalil, 1992), the tracking error converges into a known residual set provided the adaptive gains are small enough and the initial conditions for the estimated parameters are chosen to be close to the actual ones. In (Sanner and Slotine, 1992), a sliding mode technique is used where the design parameters can be chosen *a priori* to guarantee a desired tracking error bound at steady state. As indicated earlier, this result is achieved under the assumption that the gradient of $1/b(\boldsymbol{x})$ and $f(\boldsymbol{x})/b(\boldsymbol{x})$ are bounded form above by some known nonlinear functions. In (Spooner and Passino, 1996), a similar sliding technique is used together with a neural/fuzzy design methodology to establish bounds for the tracking error that can be guaranteed *a priori* by choosing certain design parameters. These results, however, are based on the assumption that the unknown ideal approximations of $f(\boldsymbol{x})$,$b(\boldsymbol{x})$ namely $f^a(\boldsymbol{x},\Theta^f)$, $b^a(\boldsymbol{x},\Theta^b)$ respectively are bounded from above by some known nonlinearities.

In this paper, we consider the same class of nonlinear systems described by (1) and considered by several investigators in the past. Our approach is based on the use of sliding mode ideas and neural network approximation techniques. We propose a control law that bypasses the stabilizability problem without using discontinuities. The adaptive laws use a new continuous σ-modification function instead of complicated projection techniques that require *a priori* knowledge of unknown parameters. Furthermore, it provides a clear procedure for choosing the design parameters to guarantee a desired tracking error at steady state without having to make any assumption about known upper bounds for the unknown nonlinearities. Consequently, our approach relaxes previous restrictive assumptions that are made to by pass the stabilizability problem (Rovithakis, 2001; Seshagiri and Khalil, 2000; Kosmatopoulos and Ioannou, 1999; Ordonez and Passino, 1999; Aloliwi and Khalil, 1997; Khalil, 1996; Spooner and Passino, 1996; Yesildirek and Lewis, 1995; Sanner and Slotine, 1992) and does that without using discontinuities as was done in (Kosmatopoulos and Ioannou, 1999; Ordonez and Passino, 1999; Spooner and Passino, 1996; Yesildirek and Lewis, 1995). Furthermore in contrast to (Kosmatopoulos and Ioannou, 1999; Ordonez and Passino, 1999; Spooner and Passino, 1996; Yesildirek and Lewis, 1995; Sanner and Slotine, 1992), we guarantee a desired tracking error at steady state by properly choosing certain design parameters without having to assume knowledge of nonlinear functions that upper bound the unknown nonlinearities plant functions.

This paper is organized as follows: In section 2 the problem formulation and preliminaries are presented. In section 3 a new robust adaptive control scheme for a class of nonlinear plants is presented and analyzed. Two examples are presented in section 4 to demonstrate the properties of the proposed adaptive control scheme.

2. PROBLEM FORMULATION

Consider the nth-order single input feedback linearizable nonlinear system (1), i.e.,

$$x^{(n)}(t) = f(\boldsymbol{x}) + b(\boldsymbol{x})u \qquad (3)$$

$$y(t) = x(t)$$

where y is the scalar system output, $f(\boldsymbol{x})$, $b(\boldsymbol{x})$ are completely unknown functions.

In order for system (1) to be controllable and feedback linearizable we assume that

Assumption 1: $b(\boldsymbol{x})$ is bounded from below by a constant $\overline{b}$, i.e., $|b(\boldsymbol{x})| \geq \overline{b}$, $\forall\, \boldsymbol{x} \in \Re^n$, and the sign of $b(\boldsymbol{x})$ is known.

The problem is to design a control law u such that the output $y(t)$ tracks a given desired trajectory $y_d(t)$, a known smooth function of time, as close as possible.

Using some ideas from sliding mode control literature, we define the scalar function $S(t)$ as the metric that describes the tracking error dynamics:

$$S(t) = (d/dt + \lambda)^{n-1} e(t) \tag{4}$$

$$e(t) = y(t) - y_d(t)$$

where λ is a positive constant defining the bandwidth of the error dynamics. The error metric $S(t) = 0$ represents a linear differential equation whose solution implies that $e(t)$ converges to zero with time constant $(n-1)/\lambda$ (Slotine and Li, 1991).

Differentiating $S(t)$ with respect to time, we obtain:

$$\begin{aligned} \dot{S} &= e^{(n)} + \alpha_{n-1} e^{(n-1)} + \cdots + \alpha_1 \dot{e} \\ &= f(\boldsymbol{x}) + b(\boldsymbol{x})u - y_d^{(n)} + (\alpha_{n-1} e^{(n-1)} + \cdots + \alpha_1 \dot{e}) \end{aligned} \tag{5}$$

where, $\alpha_{n-1}, \cdots, \alpha_1$ represent the coefficients in the binomial expansion of (4). Let

$$v(t) = -y_d^{(n)} + \alpha_{n-1} e^{(n-1)} + \cdots + \alpha_1 \dot{e} \tag{6}$$

Then, $\dot{S}$ can be written in the compact form:

$$\dot{S} = f(\boldsymbol{x}) + v(t) + b(\boldsymbol{x})u \tag{7}$$

If $f(\boldsymbol{x})$ and $b(\boldsymbol{x})$ were completely known functions, then the control law

$$u = \frac{1}{b(\boldsymbol{x})}\left[-f(\boldsymbol{x}) - v(t) - k_S S(t)\right] \tag{8}$$

could be used to meet the control objective provided of course that the controllability condition $b(\boldsymbol{x}) \neq 0$ for all $\boldsymbol{x}$ is satisfied (guaranteed by Assumption 1).

Using (8) we obtain

$$\dot{S} = -k_S S \tag{9}$$

which implies that $S(t)$ and therefore $e^{(i)}$, i=0,1,2,…,n-1, converge to zero exponentially fast.

In the case where, f, b are unknown, (8) can no longer be used. Let $f^a(\Theta^f, \boldsymbol{x})$ and $b^a(\Theta^b, \boldsymbol{x})$ denote the "optimal" approximations of the unknown functions $f(\boldsymbol{x})$ and $b(\boldsymbol{x})$ respectively, represented by a single layer neural network on a compact set $\boldsymbol{x} \in \Omega \subset \Re^n$:

$$f^a(\Theta^f, \boldsymbol{x}) = \sum_{i=1}^{l_f} \theta_i^f \zeta_i^f(\boldsymbol{x}) \tag{10a}$$

$$b^a(\Theta^b, \boldsymbol{x}) = \sum_{i=1}^{l_b} \theta_i^b \zeta_i^b(\boldsymbol{x}) \tag{10b}$$

where $\Theta^f := [\theta_1^f, \theta_2^f, \cdots, \theta_{l_f}^f]^T$, $\Theta^b := [\theta_1^b, \theta_2^b, \cdots, \theta_{l_b}^b]^T$, θ_i^f, θ_i^b are unknown parameters, $\zeta_i^f(\boldsymbol{x})$ and $\zeta_i^b(\boldsymbol{x})$ are selected basis functions, and l_f, l_b are the number of the nodes respectively. The neural network approximation errors $d_f(\boldsymbol{x})$, $d_b(\boldsymbol{x})$ are given by

$$d_f(\boldsymbol{x}) := f(\boldsymbol{x}) - f^a(\Theta^f, \boldsymbol{x}) \tag{11a}$$

$$d_b(\boldsymbol{x}) := b(\boldsymbol{x}) - b^a(\Theta^b, \boldsymbol{x}) \tag{11b}$$

By "optimal" approximation we mean the weights $\Theta^f \in \Re^{l_f}$, $\Theta^b \in \Re^{l_b}$ are chosen to minimize $d_f(\boldsymbol{x})$, $d_b(\boldsymbol{x})$ for all $\boldsymbol{x} \in \Omega$ respectively, i.e.,

$$\Theta^f := \arg \min_{\Theta \in \Re^{l_f}} \{\sup_{\boldsymbol{x} \in \Omega} \left| f(\boldsymbol{x}) - f^a(\Theta, \boldsymbol{x}) \right|\} \tag{12a}$$

$$\Theta^b := \arg \min_{\Theta \in \Re^{l_b}} \{\sup_{\boldsymbol{x} \in \Omega} \left| b(\boldsymbol{x}) - b^a(\Theta, \boldsymbol{x}) \right|\} \tag{12b}$$

Assumption 2: The approximation errors $d_f(\boldsymbol{x}), d_b(\boldsymbol{x})$ are upper bounded by some known constants $\psi_f > 0$ and $\psi_b > 0$ over a compact set $\Omega \subset \Re^n$, i.e.,

$$\sup_{\boldsymbol{x} \in \Omega} \left| d_f(\boldsymbol{x}) \right| \le \psi_f \tag{13a}$$

$$\sup_{\boldsymbol{x} \in \Omega} \left| d_b(\boldsymbol{x}) \right| \le \psi_b \tag{13b}$$

Let us now assume that the approximation functions f^a, b^a instead of the actual ones $f(\boldsymbol{x})$,$b(\boldsymbol{x})$ are known. In this case replacing $f(\boldsymbol{x})$, $b(\boldsymbol{x})$ with f^a, b^a in the control law (8) will be a straight forward approach. The stability and performance analysis of the closed loop system, however, is difficult without additional assumptions. Instead, we modify (8) and propose the control law

$$u = \frac{1}{b^a(\Theta^b, \boldsymbol{x})}\left[-k_S S(t) - \sigma_v |v(t)| S(t) - \sigma_f \left|f^a(\Theta^f, \boldsymbol{x})\right| S(t) - v(t) - f^a(\Theta^f, \boldsymbol{x})\right] \tag{14}$$

where $k_S > 0$, $\sigma_v > 0$, $\sigma_f > 0$ are design constants.

The following theorem establishes the stability and performance properties of the closed-loop system with (14) as the control law.

Theorem 1: Consider the system (3) and the control law (14). Let ψ_f, ψ_b are some positive constants such that (13a-b) are satisfied in a compact set $\Omega \subset \Re^n$ for some $\Theta^f \in \Re^{l_f}, \Theta^b \in \Re^{l_b}$. Moreover $\overline{b}$ satisfies the condition $\overline{b} > 2\psi_b$. Then given any arbitrary small positive number Φ, there exist a positive constant $\delta_0 < 1$, such that if $k_S \geq \frac{1}{\Phi}\frac{\psi_f}{1-\delta_0}$, $\sigma_v \geq \frac{1}{\Phi}\frac{\delta_0}{1-\delta_0}$, $\sigma_f \geq \frac{1}{\Phi}\frac{\delta_0}{1-\delta_0}$, for any $\boldsymbol{x}(0) \in \Omega_x \subset \Omega$, then all signals in the closed-loop system are bounded. Furthermore, the tracking error and its derivatives are bounded and $\lim_{t\to\infty}\left|e^{(i)}(t)\right| \leq 2^i \lambda^{i-n+1}\Phi$, $i = 0, 1, \cdots, n-1$.

Proof: Let us consider the following Lyapunov function:

$$V(t) = \frac{1}{2} S_\Delta^2(t) \tag{15}$$

where,

$$S_\Delta(t) = S(t) - \Phi\,\mathrm{sat}(S(t)/\Phi) \tag{16}$$

$$\mathrm{sat}(S/\Phi) = \begin{cases} 1, & \text{if } S > \Phi \\ S/\Phi, & \text{if } |S| \leq \Phi \\ -1, & \text{if } S < -\Phi \end{cases} \tag{17}$$

$S_\Delta(t)$ defines an error metric with a dead zone of width Φ. Φ>0 is a design parameter.

The time derivative $V(t)$ is then given by

$$\dot{V}(t) = S_\Delta \dot{S}_\Delta(t) \tag{18}$$

where $\dot{S}_\Delta = 0$ for $|S(t)| \leq \Phi$ and $\dot{S}_\Delta = \dot{S}$ for $|S(t)| > \Phi$. Therefore $\dot{V} = 0$ for $|S| \leq \Phi$ and the remaining of this proof deals strictly with the case of $|S| > \Phi$.

In view of equation (7) and $\dot{S}_\Delta = \dot{S}$ for $|S(t)| > \Phi$, $\dot{S}_\Delta$ can be written as:

$$\dot{S}_\Delta = f(\boldsymbol{x}) + v(t) + b(\boldsymbol{x})u \tag{19}$$

By substituting the control input (14) into (19), we obtain:

$$\begin{aligned}\dot{S}_\Delta &= f(\boldsymbol{x}) + v(t) + b^a(\Theta^b, \boldsymbol{x})u + [b(\boldsymbol{x}) - b^a(\Theta^b, \boldsymbol{x})]u \\ &= -k_S S(t) - \sigma_v |v(t)| S(t) - \sigma_f \left| f^a(\Theta^f, \boldsymbol{x}) \right| S(t) + [f(\boldsymbol{x}) - f^a(\Theta^f, \boldsymbol{x})] + [b(\boldsymbol{x}) - b^a(\Theta^b, \boldsymbol{x})]u \\ &= -k_S S(t) - \sigma_v |v(t)| S(t) - \sigma_f \left| f^a(\Theta^f, \boldsymbol{x}) \right| S(t) + d_f(\boldsymbol{x}) + d_b(\boldsymbol{x})u\end{aligned} \tag{20}$$

Using $\left|b^a(\Theta^b, \boldsymbol{x})\right| \geq |b(\boldsymbol{x})| - |d_b(\boldsymbol{x})| \geq \bar{b} - \psi_b$ and $|S| \leq |S_\Delta| + \Phi$ we have

$$\begin{aligned}|d_f(\boldsymbol{x}) + d_b(\boldsymbol{x})u| &\leq \psi_f + \delta_0 k_S |S| + \delta_0(\sigma_v |S| + 1)|v| + \delta_0(\sigma_f |S| + 1)\left|f^a\right| \\ &\leq \delta_0 k_S |S_\Delta| + \delta_0 \sigma_v |v| |S_\Delta| + \delta_0 \sigma_f \left|f^a\right| |S_\Delta| + \psi_f + \delta_0 k_S \Phi + \delta_0(\sigma_v \Phi + 1)|v| + \delta_0(\sigma_f \Phi + 1)\left|f^a\right|\end{aligned} \tag{21}$$

where δ_0 is defined as:

$$\delta_0 = \frac{\psi_b}{\bar{b} - \psi_b} \tag{22}$$

Using (20), (16) and $S_\Delta \, \mathrm{sat}(S/\Phi) = |S_\Delta|$, $\dot{V}$ becomes:

$$\dot{V} = -k_S S_\Delta^2 - \sigma_v |v(t)| S_\Delta^2 - \sigma_f \left|f^a(\Theta^f, \boldsymbol{x})\right| S_\Delta^2 - k_S \Phi |S_\Delta| - \sigma_v \Phi |v(t)| |S_\Delta| - \sigma_f \Phi \left|f^a(\Theta^f, \boldsymbol{x})\right| |S_\Delta| + S_\Delta \{d_f(\boldsymbol{x}) + d_b(\boldsymbol{x})u\} \tag{23}$$

Using (21), we obtain:

$$\begin{aligned}\dot{V} &\leq -(1-\delta_0) k_S S_\Delta^2 - (1-\delta_0)\sigma_v |v(t)| S_\Delta^2 - (1-\delta_0)\sigma_f \left|f^a(\Theta^f, \boldsymbol{x})\right| S_\Delta^2 \\ &\quad - \{(1-\delta_0) k_S \Phi - \psi_f\} |S_\Delta| - \{\sigma_v \Phi - \delta_0(1 + \sigma_v \Phi)\} |v(t)| |S_\Delta| - \{\sigma_f \Phi - \delta_0(1 + \sigma_f \Phi)\} \left|f^a(\Theta^f, \boldsymbol{x})\right| |S_\Delta|\end{aligned} \tag{24}$$

Assume that $\bar{b} > 2\psi_b$, then

$$\delta_0 = \frac{\psi_b}{\bar{b} - \psi_b} < 1 \tag{25}$$

Given (25), we can choose $k_S, \Phi, \sigma_v, \sigma_f, \sigma_b$ such that the following conditions are satisfied,

$$k_S \Phi \geq \frac{\psi_f}{1 - \delta_0} \tag{26a}$$

$$\sigma_v \Phi \geq \frac{\delta_0}{1 - \delta_0} \tag{26b}$$

$$\sigma_f \Phi \geq \frac{\delta_0}{1 - \delta_0} \tag{26c}$$

We obtain

$$\dot{V} = 0 \,, \text{ for } |S| \le \Phi \tag{27a}$$

$$\dot{V} \le -(1-\delta_0) k_S S_\Delta^2 \le 0 \,, \text{ for } |S| > \Phi \tag{27b}$$

The results (27a-b) are valid provided (13a-b) hold. Since (13a-b) hold on a compact set, i.e., $\boldsymbol{x} \in \Omega$, all states need to remain in this compact set for all $t \ge 0$ in order for the results to be valid. Consider the set

$$\Omega_{\boldsymbol{x}} = \{\boldsymbol{x} \mid V(\boldsymbol{x}) \le V_0\} \tag{28}$$

where $V_0 > V(0)$ and $V_0 > \Phi$ is chosen so that $\Omega_{\boldsymbol{x}} \subset \Omega$. Then for $\forall\, \boldsymbol{x}(0) \in \Omega_{\boldsymbol{x}}$ it follows from (15), (27a-b) that $V(t)$ is bounded from above by V_0 for all $t \ge 0$ which implies that $\boldsymbol{x} \in \Omega_{\boldsymbol{x}} \subset \Omega$, $\forall\; t \ge 0$.

The boundedness of $V(t)$ implies that S_Δ is bounded for all $t > 0$. Since $V(t)$ is bounded from below and is non-increasing with time, it has a limit, i.e., $\lim_{t\to\infty} V(t) = V_\infty$. Using (27b) and the fact that $S_\Delta = 0$ for $|S| \le \Phi$, we have $\lim_{t\to\infty} \int_0^t k_S S_\Delta^2(\tau) d\tau = k_S \int_0^\infty S_\Delta^2(t) dt \le \dfrac{V(0) - V_\infty}{1-\delta_0} < \infty$ which implies that $S_\Delta \in L_2$. From $S_\Delta \in L_\infty$, it follows that all signals are bounded which implies that $\dot{S}_\Delta \in L_\infty$. From $\dot{S}_\Delta \in L_\infty$ and $S_\Delta \in L_2$ we have $S_\Delta(t) \to 0$ as $t \to \infty$ (Ioannou and Sun, 1996). This implies that $S(t)$ converges to the region $|S| \le \Phi$ which in turn implies that the tracking error converges to a small residual set whose size is characterized by the size of the design parameter Φ. We can also establish that the tracking error and its derivatives are bounded from above and $\lim_{t\to\infty} \left| e^{(i)}(t) \right| \le 2^i \lambda^{i-n+1} \Phi$, $i = 0, 1, \cdots, n-1$ (Slotine and Li, 1991; Slotine and Coetsee, 1986)

In the case where f^a, b^a are unknown, the control law (14) cannot be used. In this case we follow the certainty equivalence approach and replace the unknown functions $f^a(\Theta^f, \boldsymbol{x})$, $b^a(\Theta^b, \boldsymbol{x})$ in (14) with their estimates. Let the estimates of the functions $f^a(\Theta^f, \boldsymbol{x})$, $b^a(\Theta^b, \boldsymbol{x})$ at time t be formed as

$$f^a(\hat{\Theta}^f, \boldsymbol{x}) = \sum_{i=1}^{l_f} \hat{\theta}_i^f(t) \zeta_i^f(\boldsymbol{x}) \tag{29a}$$

$$b^a(\hat{\Theta}^b, \boldsymbol{x}) = \sum_{i=1}^{l_b} \hat{\theta}_i^b(t) \zeta_i^b(\boldsymbol{x}) \tag{29b}$$

where $\hat{\theta}_i^f(t)$, $\hat{\theta}_i^b(t)$ are the estimates of θ_i^f, θ_i^b respectively at time t. The difference between the estimated and actual parameter values results in the estimation errors

$$f^a(\tilde{\Theta}^f, \boldsymbol{x}) = f^a(\hat{\Theta}^f, \boldsymbol{x}) - f^a(\Theta^f, \boldsymbol{x}) = \sum_{i=1}^{l_f} \tilde{\theta}_i^f(t)\zeta_i^f(\boldsymbol{x}) \quad (30a)$$

$$b^a(\tilde{\Theta}^b, \boldsymbol{x}) = b^a(\hat{\Theta}^b, \boldsymbol{x}) - b^a(\Theta^b, \boldsymbol{x}) = \sum_{i=1}^{l_b} \tilde{\theta}_i^b(t)\zeta_i^b(\boldsymbol{x}) \quad (30b)$$

where

$$\tilde{\theta}_i^f(t) = \hat{\theta}_i^f(t) - \theta_i^f, \ \tilde{\theta}_i^b(t) = \hat{\theta}_i^b(t) - \theta_i^b \quad (31)$$

are the parameter errors. The estimator and parameter errors are not available for measurement, therefore equations (30a)-(31) are used only for analysis.

Using the Certainty Equivalence (CE) approach we replace the unknown nonlinearities $f^a(\Theta^f, \boldsymbol{x})$, $b^a(\Theta^b, \boldsymbol{x})$ in (14) with their on-line estimates $f^a(\hat{\Theta}^f, \boldsymbol{x})$, $b^a(\hat{\Theta}^b, \boldsymbol{x})$ to come up with an initial guess of the adaptive control law:

$$u = \frac{1}{b^a(\hat{\Theta}^b, \boldsymbol{x})}\left[-k_S S(t) - \sigma_v |v(t)| S(t) - \sigma_f \left| f^a(\hat{\Theta}^f, \boldsymbol{x}) \right| S(t) - v(t) - f^a(\hat{\Theta}^f, \boldsymbol{x})\right] \quad (32)$$

However, the CE control law (32) cannot be used to stabilize the closed loop system for the case where the estimated plant becomes uncontrollable, i.e., $b^a(\hat{\Theta}^b, \boldsymbol{x}) = 0$ at some time t due to the division by $b^a(\hat{\Theta}^b, \boldsymbol{x})$. In the following section we show how to bypass this problem by modifying (32).

3. ROBUST ADAPTIVE CONTROL LAW

In order to take care of the possible loss of controllability of the estimated plant at some points in time, we modify the CE control law (32) as:

$$u = \frac{b^a(\hat{\Theta}^b, \boldsymbol{x})}{(b^a(\hat{\Theta}^b, \boldsymbol{x}))^2 + \delta_b}\left[-k_S S(t) - \sigma_v |v(t)| S(t) - \sigma_f \left| f^a(\hat{\Theta}^f, \boldsymbol{x}) \right| S(t) - v(t) - f^a(\hat{\Theta}^f, \boldsymbol{x})\right] \quad (33)$$

where $\delta_b > 0$ is a small design constant.

The parameters $\hat{\theta}_i^f(t)$, $\hat{\theta}_i^b(t)$ in $f^a(\hat{\Theta}^f, \boldsymbol{x})$, $b^a(\hat{\Theta}^b, \boldsymbol{x})$ respectively are updated as follows:

$$\dot{\hat{\theta}}_i^f(t) = k_f S_\Delta \zeta_i^f(\boldsymbol{x}) \quad (34a)$$

$$\dot{\hat{\theta}}_i^b(t) = k_b S_\Delta u \zeta_i^b(\boldsymbol{x}) + \rho(t) k_b \sigma_b \operatorname{sgn}(b(\boldsymbol{x})) |S_\Delta| (|u'| + |u|) \zeta_i^b(\boldsymbol{x}) \quad (34b)$$

where

$$u' = \frac{1}{(b^a(\hat{\Theta}^b, \boldsymbol{x}))^2 + \delta_b}\left[-k_S S(t) - \sigma_v |v(t)| S(t) - \sigma_f \left| f^a(\hat{\Theta}^f, \boldsymbol{x}) \right| S(t) - v(t) - f^a(\hat{\Theta}^f, \boldsymbol{x})\right] \tag{35}$$

k_f , $k_b > 0$ are the adaptive gains chosen by the designer, $\sigma_b > 0$ is a small design parameters, sgn(·) is the sign function (sgn(x)=1, if $x \geq 0$ and sgn(x)=-1 otherwise), and $\rho(t)$ is a continuous switching function given by:

$$\rho(t) = \begin{cases} 0, & \text{if } \left|b^a(\hat{\Theta}^b, \boldsymbol{x})\right| \geq \bar{b} - \psi_b \\ (\bar{b} - \psi_b - \left|b^a(\hat{\Theta}^b, \boldsymbol{x})\right|)/\Delta, & \text{if } \bar{b} - \psi_b - \Delta < \left|b^a(\hat{\Theta}^b, \boldsymbol{x})\right| < \bar{b} - \psi_b \\ 1, & \text{if } \left|b^a(\hat{\Theta}^b, \boldsymbol{x})\right| \leq \bar{b} - \psi_b - \Delta \end{cases} \tag{36}$$

where $\Delta > 0$ is a design parameter used to avoid discontinuity in $\rho(t)$ as shown graphically in Figure 1. The continuous switching function $\rho(t)$ (shown in Figure 1), instead of a discontinuous one, is used in order to guarantee the existence and uniqueness of solutions of the closed-loop system (Polycarpou and Ioannou, 1993).

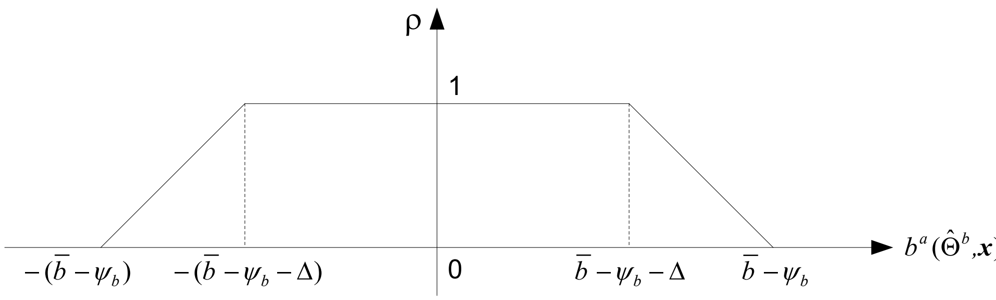

Fig. 1. Continuous Switching Function $\rho(t)$

By design, the control law (33) will never involve division by zero since $(b^a(\hat{\Theta}^b, \boldsymbol{x}))^2 + \delta_b > \delta_b > 0$, $\forall\, \boldsymbol{x}, t$. Therefore, the proposed controller overcomes the difficulty encountered in implementing many adaptive control laws where the identified model becomes uncontrollable at some points in time. It is also interesting to note that $u \to 0$ with the same speed as $b^a(\hat{\Theta}^b, \boldsymbol{x}) \to 0$. Thus, when the estimate $b^a(\hat{\Theta}^b, \boldsymbol{x})$ approaches zero, the control input remains bounded and also reduces to zero. In other words in such case it is pointless to control what appears to the controller as uncontrollable plant. It is worth noting that the proposed control law is continuous, which as indicated above does not create any problem with respect to guaranteeing existence and uniqueness of solutions of the closed loop system (Polycarpou and Ioannou, 1993).

The control law (33)-(34a-b) is designed using stability and Lyapunov type arguments and its properties are described by the following theorem.

Theorem 2: Consider the system (3), the control law (33) and the adaptive laws (34a-b). Let ψ_f, ψ_b are some positive constants such that (13a-b) are satisfied in a compact set $\Omega \subset \Re^n$ for some $\Theta^f \in \Re^{l_f}, \Theta^b \in \Re^{l_b}$. Moreover $\overline{b}$ satisfies the condition $\overline{b} > \sqrt{\delta_b} + \Delta + 3\psi_b$ for arbitrary small positive constants δ_b and Δ. Then given any arbitrary small positive number Φ, there exist positive constants $\delta_1 < 1$, δ_2, δ_3 such that if $k_S \geq \frac{1}{\Phi}\frac{\psi_f}{1-\delta_1}$, $\sigma_v \geq \frac{1}{\Phi}\frac{\delta_1}{1-\delta_1}$, $\sigma_f \geq \frac{1}{\Phi}\frac{\delta_1}{1-\delta_1}$, $\sigma_b \geq \max\{\delta_2, \delta_3\}$, for any $\{\widetilde{\Theta}^f(0), \widetilde{\Theta}^b(0)\} \in \Omega_\Theta$ and $\boldsymbol{x}(0) \in \Omega_{\boldsymbol{x}} \subset \Omega$, where $\Omega_\Theta \subset \Re^{l_f + l_b}$ is a compact set in the space of $\widetilde{\Theta}^f, \widetilde{\Theta}^b$, all signals in the closed-loop system are bounded. Furthermore, the tracking error and its derivatives are bounded and $\lim_{t\to\infty}\left|e^{(i)}(t)\right| \leq 2^i \lambda^{i-n+1}\Phi$, $i = 0, 1, \cdots, n-1$.

Proof: Let us consider the following Lyapunov-like function:

$$V(t) = \frac{1}{2}S_\Delta^2(t) + \frac{1}{2k_f}\sum_{i=1}^{l_f}\left(\widetilde{\theta}_i^f(t)\right)^2 + \frac{1}{2k_b}\sum_{i=1}^{l_b}\left(\widetilde{\theta}_i^b(t)\right)^2 \tag{37}$$

The time derivative of $V(t)$ is then given by

$$\dot{V}(t) = S_\Delta \dot{S}_\Delta(t) + \frac{1}{k_f}\sum_{i=1}^{l_f}\widetilde{\theta}_i^f \dot{\hat{\theta}}_i^f(t) + \frac{1}{k_b}\sum_{i=1}^{l_b}\widetilde{\theta}_i^b \dot{\hat{\theta}}_i^b(t) \tag{38}$$

where, $\dot{S}_\Delta = 0$ for $|S(t)| \leq \Phi$ and $\dot{S}_\Delta = \dot{S}$ for $|S(t)| > \Phi$. In view of the adaptive laws (34a-b), $\dot{V} = 0$ for $|S| \leq \Phi$. Therefore, the remaining of this proof deals strictly with the case of $|S| > \Phi$. First, we analyze the first term in $\dot{V}$ in (38). Let us rewrite the control law (33) as

$$u = \frac{b^a(\hat{\Theta}^b, \boldsymbol{x})}{(b^a(\hat{\Theta}^b, \boldsymbol{x}))^2 + \delta_b}\overline{u} \tag{39}$$

where $\overline{u}$ is given by:

$$\overline{u} = -k_S S(t) - \sigma_v |v(t)| S(t) - \sigma_f \left|f^a(\hat{\Theta}^f, \boldsymbol{x})\right| S(t) - v(t) - f^a(\hat{\Theta}^f, \boldsymbol{x}) \tag{40}$$

By substituting the control input (39) and (40) into (19), $\dot{S}_\Delta$ can be written as

$$\begin{aligned}
\dot{S}_\Delta &= f(\boldsymbol{x}) + v(t) + b^a(\hat{\Theta}^b, \boldsymbol{x})u + [b(\boldsymbol{x}) - b^a(\hat{\Theta}^b, \boldsymbol{x})]u \\
&= f(\boldsymbol{x}) + v(t) + \frac{(b^a(\hat{\Theta}^b, \boldsymbol{x}))^2}{(b^a(\hat{\Theta}^b, \boldsymbol{x}))^2 + \delta_b}\overline{u} + [b(\boldsymbol{x}) - b^a(\hat{\Theta}^b, \boldsymbol{x})]u \\
&= f(\boldsymbol{x}) + v(t) + \overline{u} - \frac{\delta_b}{(b^a(\hat{\Theta}^b, \boldsymbol{x}))^2 + \delta_b}\overline{u} + [b(\boldsymbol{x}) - b^a(\hat{\Theta}^b, \boldsymbol{x})]u \\
&= -k_S S(t) - \sigma_v |v(t)| S(t) - \sigma_f \left| f^a(\hat{\Theta}^f, \boldsymbol{x}) \right| S(t) \\
&\quad + [f(\boldsymbol{x}) - f^a(\hat{\Theta}^f, \boldsymbol{x})] + [b(\boldsymbol{x}) - b^a(\hat{\Theta}^b, \boldsymbol{x})]u - \delta_b u'
\end{aligned} \tag{41}$$

Using the identities,

$$\begin{aligned}
f(\boldsymbol{x}) - f^a(\hat{\Theta}^f, \boldsymbol{x}) &= [f(\boldsymbol{x}) - f^a(\Theta^f, \boldsymbol{x})] - [f^a(\hat{\Theta}^f, \boldsymbol{x}) - f^a(\Theta^f, \boldsymbol{x})] \\
&= d_f(\boldsymbol{x}) - f^a(\widetilde{\Theta}^f, \boldsymbol{x})
\end{aligned} \tag{42}$$

$$\begin{aligned}
b(\boldsymbol{x}) - b^a(\hat{\Theta}^b, \boldsymbol{x}) &= [b(\boldsymbol{x}) - b^a(\Theta^b, \boldsymbol{x})] - [b^a(\hat{\Theta}^b, \boldsymbol{x}) - b^a(\Theta^b, \boldsymbol{x})] \\
&= d_b(\boldsymbol{x}) - b^a(\widetilde{\Theta}^b, \boldsymbol{x})
\end{aligned} \tag{43}$$

$\dot{S}_\Delta$ becomes:

$$\dot{S}_\Delta = -k_S S - \sigma_v |v(t)| S - \sigma_f \left| f^a(\hat{\Theta}^f, \boldsymbol{x}) \right| S - f^a(\hat{\Theta}^f, \boldsymbol{x}) - b^a(\widetilde{\Theta}^b, \boldsymbol{x})u + d_f(\boldsymbol{x}) + \{d_b(\boldsymbol{x})u - \delta_b u'\} \tag{44}$$

The term $\{d_b(\boldsymbol{x})u - \delta_b u'\}$ in (44) is a modeling error term representing the effect of the design constant δ_b in the control law and the approximation error $d_b(\boldsymbol{x})$. The price paid to avoid the singularity in the adaptive control law (33) is that the design parameter δ_b appears as a disturbance in the system. The disturbance term $\{d_b(\boldsymbol{x})u - \delta_b u'\}$ will become dominant when the estimate $b^a(\hat{\Theta}^b, \boldsymbol{x}) \to 0$.

Define:

$$\delta_1 := \frac{\psi_b}{\overline{b} - \psi_b - \Delta} + \frac{\delta_b}{(\overline{b} - \psi_b - \Delta)^2 + \delta_b} \tag{45a}$$

$$\delta_2 := \frac{2\delta_b}{\overline{b} - \psi_b} \tag{45b}$$

$$\delta_3 := \frac{\psi_b}{\overline{b} - \psi_b} \tag{45c}$$

Then, as shown in the Appendix, the absolute value of the last term in $\dot{S}_\Delta$ can be expressed as:

$$|d_b(\boldsymbol{x})u - \delta_b u'| \le \delta_1 |\overline{u}| + \rho\delta_2 \left| b^a(\widetilde{\Theta}^b, \boldsymbol{x}) \right| |u'| + \rho\delta_3 \left| b^a(\widetilde{\Theta}^b, \boldsymbol{x}) \right| |u| \tag{46}$$

Since $|S| \le |S_\Delta| + \Phi$, we have

$$|\bar{u}| \le k_S|S_\Delta| + \sigma_v|v(t)||S_\Delta| + \sigma_f\left|f^a(\hat{\Theta}^f,\boldsymbol{x})\right||S_\Delta| + k_S\Phi + (\sigma_v\Phi+1)|v(t)| + (\sigma_f\Phi+1)\left|f^a(\hat{\Theta}^f,\boldsymbol{x})\right| \tag{47}$$

and (46) can be rewritten as:

$$\begin{aligned}|d_b(\boldsymbol{x})u - \delta_b u'| \le\ & \delta_1 k_S|S_\Delta| + \delta_1\sigma_v|v(t)||S_\Delta| + \delta_1\sigma_f\left|f^a(\hat{\Theta}^f,\boldsymbol{x})\right||S_\Delta| + \delta_1 k_S\Phi + \delta_1(\sigma_v\Phi+1)|v(t)| \\ & + \delta_1(\sigma_f\Phi+1)\left|f^a(\hat{\Theta}^f,\boldsymbol{x})\right| + \rho\delta_2\left|b^a(\tilde{\Theta}^b,\boldsymbol{x})\right||u'| + \rho\delta_3\left|b^a(\tilde{\Theta}^b,\boldsymbol{x})\right||u|\end{aligned} \tag{48}$$

Then in view of (44), and using $S_\Delta\ \mathrm{sat}(S/\Phi) = |S_\Delta|$, the first term in $\dot{V}$ is expressed as:

$$\begin{aligned}S_\Delta\dot{S}_\Delta =\ & -k_S S_\Delta^2 - \sigma_v|v(t)|S_\Delta^2 - \sigma_f\left|f^a(\hat{\Theta}^f,\boldsymbol{x})\right|S_\Delta^2 - k_S\Phi|S_\Delta| - \sigma_v\Phi|v(t)||S_\Delta| - \sigma_f\Phi\left|f^a(\hat{\Theta}^f,\boldsymbol{x})\right||S_\Delta| \\ & - S_\Delta f^a(\tilde{\Theta}^f,\boldsymbol{x}) - S_\Delta b^a(\tilde{\Theta}^b,\boldsymbol{x})u + S_\Delta d_f(\boldsymbol{x}) + S_\Delta\{d_b(\boldsymbol{x})u - \delta_b u'\}\end{aligned} \tag{49}$$

Using (48) and (13a), we obtain:

$$\begin{aligned}S_\Delta\dot{S}_\Delta \le\ & -(1-\delta_1)k_S S_\Delta^2 - (1-\delta_1)\sigma_v|v(t)|S_\Delta^2 - (1-\delta_1)\sigma_f\left|f^a(\hat{\Theta}^f,\boldsymbol{x})\right|S_\Delta^2 \\ & - \{(1-\delta_1)k_S\Phi - \psi_f\}|S_\Delta| - \{\sigma_v\Phi - \delta_1(1+\sigma_v\Phi)\}|v(t)||S_\Delta| - \{\sigma_f\Phi - \delta_1(1+\sigma_f\Phi)\}\left|f^a(\hat{\Theta}^f,\boldsymbol{x})\right||S_\Delta| \\ & - S_\Delta f^a(\tilde{\Theta}^f,\boldsymbol{x}) - S_\Delta b^a(\tilde{\Theta}^b,\boldsymbol{x})u + \rho\delta_2|S_\Delta|\left|b^a(\tilde{\Theta}^b,\boldsymbol{x})\right||u'| + \rho\delta_3|S_\Delta|\left|b^a(\tilde{\Theta}^b,\boldsymbol{x})\right||u|\end{aligned} \tag{50}$$

In view of (34a), the second term in $\dot{V}$ can be expressed as:

$$\begin{aligned}\frac{1}{k_f}\sum_{i=1}^{l_f}\tilde{\theta}_i^f(t)\dot{\hat{\theta}}_i^f(t) &= \frac{1}{k_f}\sum_{i=1}^{l_f}\tilde{\theta}_i^f(t)\{k_f S_\Delta\zeta_i^f(\boldsymbol{x})\} \\ &= S_\Delta f^a(\tilde{\Theta}^f,\boldsymbol{x})\end{aligned} \tag{51}$$

Finally, using (34b) the last term in $\dot{V}$ can also be expanded as:

$$\begin{aligned}\frac{1}{k_b}\sum_{i=1}^{l_b}\tilde{\theta}_i^b(t)\dot{\hat{\theta}}_i^b(t) &= \frac{1}{k_b}\sum_{i=1}^{l_b}\tilde{\theta}_i^b(t)\{k_b S_\Delta u\zeta_i^b(\boldsymbol{x}) + \rho k_b\sigma_b\,\mathrm{sgn}(b(\boldsymbol{x}))|S_\Delta|(|u'|+|u|)\zeta_i^b(\boldsymbol{x})\} \\ &= S_\Delta b^a(\tilde{\Theta}^b,\boldsymbol{x})u - \rho\sigma_b|S_\Delta|\left|b^a(\tilde{\Theta}^b,\boldsymbol{x})\right||u'| - \rho\sigma_b|S_\Delta|\left|b^a(\tilde{\Theta}^b,\boldsymbol{x})\right||u|\end{aligned} \tag{52}$$

Here, we have used the identity $\rho b^a(\tilde{\Theta}^b,\boldsymbol{x})\,\mathrm{sgn}(b(\boldsymbol{x})) = -\rho\left|b^a(\tilde{\Theta}^b,\boldsymbol{x})\right|$. Since $b^a(\tilde{\Theta}^b,\boldsymbol{x}) = b^a(\hat{\Theta}^b,\boldsymbol{x}) - b^a(\Theta^b,\boldsymbol{x}) = (b^a(\hat{\Theta}^b,\boldsymbol{x}) + d_b(\boldsymbol{x})) - b(\boldsymbol{x})$ and $\rho \ne 0$ only if $\left|b^a(\hat{\Theta}^b,\boldsymbol{x})\right| \le \bar{b} - \psi_b$ implies that $\left|b^a(\hat{\Theta}^b,\boldsymbol{x}) + d_b(\boldsymbol{x})\right| \le \bar{b}$, then, for $\rho \ne 0$, the sign of $\rho b^a(\tilde{\Theta}^a,\boldsymbol{x})$ is always the opposite of sign $b(\boldsymbol{x})$ $\forall\ t \ge 0$.

Combining (50)-(52), $\dot{V}$ satisfies:

$$\dot{V} \le -(1-\delta_1)k_S S_\Delta^2 - (1-\delta_1)\sigma_v |v(t)| S_\Delta^2 - (1-\delta_1)\sigma_f \left| f^a(\hat{\Theta}^f, \boldsymbol{x}) \right| S_\Delta^2$$
$$- \{(1-\delta_1)k_S\Phi - \psi_f\} |S_\Delta| - \{(1-\delta_1)\sigma_v\Phi - \delta_1\} |v(t)| |S_\Delta| - \{(1-\delta_1)\sigma_f\Phi - \delta_1\} \left| f^a(\hat{\Theta}^f, \boldsymbol{x}) \right| |S_\Delta| \qquad (53)$$
$$- \rho(\sigma_b - \delta_2) |S_\Delta| \left| b^a(\tilde{\Theta}^b, \boldsymbol{x}) \right| |u'| - \rho(\sigma_b - \delta_3) |S_\Delta| \left| b^a(\tilde{\Theta}^b, \boldsymbol{x}) \right| |u|$$

Assume that $\bar{b} > \sqrt{\delta_b} + \Delta + 3\psi_b$, then

$$\delta_1 = \frac{\psi_b}{\bar{b} - \psi_b - \Delta} + \frac{\delta_b}{(\bar{b} - \psi_b - \Delta)^2 + \delta_b} < 1 \qquad (54)$$

Given (54), we can choose $k_S, \Phi, \sigma_v, \sigma_f, \sigma_b$ such that the following conditions are satisfied:

$$k_S\Phi \ge \frac{\psi_f}{1-\delta_1} \qquad (55a)$$

$$\sigma_v\Phi \ge \frac{\delta_1}{1-\delta_1} \qquad (55b)$$

$$\sigma_f\Phi \ge \frac{\delta_1}{1-\delta_1} \qquad (55c)$$

$$\sigma_b \ge \max\{\delta_2, \delta_3\} = \max\left(\frac{2\delta_b}{\bar{b} - \psi_b}, \frac{\psi_b}{\bar{b} - \psi_b} \right) \qquad (55d)$$

We obtain

$$\dot{V} = 0, \text{ for } |S| \le \Phi \qquad (56a)$$

$$\dot{V} \le -(1-\delta_1)k_S S_\Delta^2 \le 0, \text{ for } |S| > \Phi \qquad (56b)$$

The results (56a-b) are valid provided (13a-b) hold. Since (13a-b) hold on a compact set, i.e., $\boldsymbol{x} \in \Omega$, all states need to remain in this compact set for all $t \ge 0$ in order for the results to be valid. Consider the set

$$M(\boldsymbol{x}, \tilde{\Theta}^f, \tilde{\Theta}^b) = \{\boldsymbol{x}, \tilde{\Theta}^f, \tilde{\Theta}^b \mid V(\boldsymbol{x}, \tilde{\Theta}^f, \tilde{\Theta}^b) \le V_0\} \qquad (57)$$

where $V_0 > V(0)$ and $V_0 > \Phi$ is chosen as the largest constant for which $M = \Omega_{\boldsymbol{x}} \times \Omega_\Theta$, where $\Omega_{\boldsymbol{x}} \subset \Omega$. Then for $\forall\, \boldsymbol{x}(0) \in \Omega_{\boldsymbol{x}}$ and $\{\tilde{\Theta}^f(0), \tilde{\Theta}^b(0)\} \in \Omega_\Theta$ it follows from (37), (56a-b) that $V(t)$ is bounded from above by V_0 for all $t \ge 0$ which implies that $\boldsymbol{x} \in \Omega_{\boldsymbol{x}} \subset \Omega$, $\forall\ t \ge 0$.

The sets $\Omega_{\boldsymbol{x}}$, Ω_Θ, Ω, and M can be illustrated using $\boldsymbol{x}, \Theta \in \Re$ as shown in Figure 2.

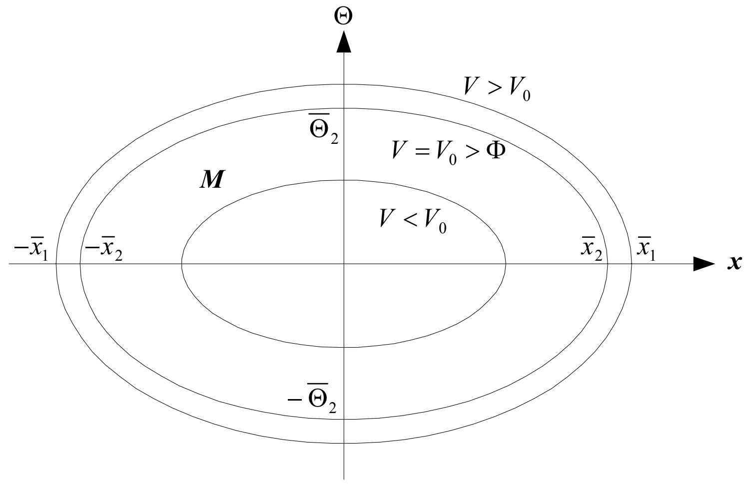

Fig 2. The sets $\Omega_{\boldsymbol{x}}$, Ω_{Θ}, Ω, *M* in two dimensional case

In Figure 2, $\Omega = \{\boldsymbol{x} \in \Re \mid -\overline{x}_1 < \boldsymbol{x} < \overline{x}_1\}$, $\Omega_{\boldsymbol{x}} = \{\boldsymbol{x} \in \Re \mid -\overline{x}_2 < \boldsymbol{x} < \overline{x}_2\}$, $\Omega_{\Theta} = \{\Theta \in \Re \mid -\overline{\Theta}_2 < \Theta < \overline{\Theta}_2\}$.

The boundness of *V*(*t*) implies that S_{Δ} and $\tilde{\theta}_i^f, \tilde{\theta}_i^f$ are bounded for all $t > 0$. Applying the similar arguments as in the proof of *Theorem* 1, we can establish that the tracking error and its derivatives are bounded from above and $\lim_{t\to\infty} \left| e^{(i)}(t) \right| \le 2^i \lambda^{i-n+1} \Phi$, $i = 0, 1, \cdots, n-1$ ■.

Design Parameter Procedure:

The design parameters can be chosen to guarantee that the tracking error is within a desired bound at steady state by using the following procedure.

(1) Using the upper bound of the approximation error ψ_b, check if the lower bound of *b*(***x***) satisfies $\overline{b} > 3\psi_b$. If so, choose the design parameters δ_b, Δ such that $\sqrt{\delta_b} + \Delta + 3\psi_b < \overline{b}$. If not then the number of nodes l_b of the neural network for *b*(***x***) has to be increased in order to obtain a better approximation.

(2) Calculate $\delta_1, \delta_2, \delta_3$ using (45a-c) and the knowledge of $\overline{b}$, ψ_b, δ_b, Δ, i.e.,

$$\delta_1 = \frac{\psi_b}{\overline{b} - \psi_b - \Delta} + \frac{\delta_b}{(\overline{b} - \psi_b - \Delta)^2 + \delta_b}, \quad \delta_2 = \frac{2\delta_b}{\overline{b} - \psi_b}, \quad \delta_3 = \frac{\psi_b}{\overline{b} - \psi_b}$$

(3) Set the desired upper bound for the tracking error at steady state equal to $\lambda^{-n+1}\Phi$ and choose λ, Φ to satisfy the equation.

(4) Choose σ_b in the adaptive law (34b) such that $\sigma_b \ge \max(\delta_2, \delta_3)$.

(5) Choose k_S, σ_v, σ_f such that $k_S \geq \frac{1}{\Phi}\frac{\psi_f}{1-\delta_1}$, $\sigma_v \geq \frac{1}{\Phi}\frac{\delta_1}{1-\delta_1}$, $\sigma_f \geq \frac{1}{\Phi}\frac{\delta_1}{1-\delta_1}$.

Remark1: It is worth noting that the ratio, $\psi_b / \bar{b} < 1/3$, gives us an upper bound for the approximation error $d_b(\boldsymbol{x})$ that can be tolerated by the closed loop system. Larger $\bar{b}$ implies that the closed loop system can tolerate larger approximation error $d_b(\boldsymbol{x})$. On the other hand, small $\bar{b}$ requires more accurate approximation of $b(\boldsymbol{x})$ by $b^a(\Theta^b, \boldsymbol{x})$, which normally implies more nodes in the neural network for $b^a(\Theta^b, \boldsymbol{x})$.

Remark 2: The tracking error is guaranteed to converge inside the residual set of size $\lambda^{-n+1}\Phi$ as $t \to \infty$. The smaller the design parameter Φ is, the smaller the tracking error is guaranteed to be at steady state. From (55a-c) the design parameters k_S, σ_v, σ_f are inversely proportional to Φ. Given a desired Φ, the value of k_S depends on δ_1 and is proportional to ψ_f. The values of σ_v, σ_f depend only on δ_1. It is easy to check that δ_1 is a crucial parameter used to determine k_S, σ_v, σ_f. Let us rewrite δ_1 as $\delta_1 = \frac{\psi_b / \bar{b}}{1 - \psi_b / \bar{b} - \Delta / \bar{b}} + \frac{(\sqrt{\delta_b} / \bar{b})^2}{(1 - \psi_b / \bar{b} - \Delta / \bar{b})^2 + (\sqrt{\delta_b} / \bar{b})^2}$. The value of δ_1 is a function of the 3 ratios $\psi_b / \bar{b}$, $\sqrt{\delta_b} / \bar{b}$, $\Delta / \bar{b}$. Also, rewrite $\delta_2 = \frac{2(\delta_b / \bar{b})}{1 - (\psi_b / \bar{b})}$, $\delta_3 = \frac{(\psi_b / \bar{b})}{1 - (\psi_b / \bar{b})}$. From $\psi_b / \bar{b} < 1/3$, we have $\delta_2 < 3\delta_b / \bar{b}$, $\delta_3 < 1/2$. The design parameter σ_b in the adaptive law (34b) requires $\sigma_b \geq \max\{\delta_2, \delta_3\}$, which depends on the ratios $\psi_b / \bar{b}$, $\delta_b / \bar{b}$. Obviously the design parameters, δ_b, Δ, appear as a disturbance. This is the price paid to overcome the singularity of the control law. Smaller ratios $\psi_b / \bar{b}$, $\delta_b / \bar{b}$, $\Delta / \bar{b}$ imply smaller values for k_S, σ_v, σ_f and σ_b. However, a smaller ψ_b needs more accurate approximation or better knowledge of the unknown function $b(\boldsymbol{x})$; a smaller δ_b implies a relative larger control input when $b(\hat{\Theta}^b, \boldsymbol{x})$ becomes smaller; a smaller Δ may imply a faster switching action. Therefore, there is a tradeoff between the design parameters, k_S, σ_v, σ_f, σ_b, and ψ_f, $\psi_b / \bar{b}$, $\delta_b / \bar{b}$, $\Delta / \bar{b}$. These tradeoffs can be taken into account in the design parameter procedure presented above in order to improve the properties of the closed loop system further.

Remark 3: As shown in the proof of *Theorem* 2, the new σ-modification $\rho(t)k_b\sigma_b\,\mathrm{sgn}(b(\boldsymbol{x}))|S_\Delta|(|u'|+|u|)\zeta_i^b(\boldsymbol{x})$ in the adaptive law (34b) is crucial in avoiding the potentially unstable situation resulting from the estimate $b^a(\hat{\Theta}^b,\boldsymbol{x})$ going to zero. In the expression for $\dot{S}_\Delta$, the modeling error term $\{d_b(\boldsymbol{x})u-\delta_b u'\}$ appears because of the design parameter, δ_b, and the approximation error $d_b(\boldsymbol{x})$. This term may become dominant when $b^a(\hat{\Theta}^b,\boldsymbol{x})$ approaches zero. This implies the tracking error metric S_Δ may become unbounded when $b^a(\hat{\Theta}^b,\boldsymbol{x})\to 0$. The σ-modification in $\dot{\hat{\theta}}_i^b$ is activated to guarantee stability when $b^a(\hat{\Theta}^b,\boldsymbol{x})$ becomes smaller than the lower bound of b^a. From the adaptive learning point of view, when the σ-modification is activated, $\hat{\theta}_i^b$ will increase along the direction of the actual $b(\boldsymbol{x})$. Thus the special σ-modification in (34b) could be viewed as a soft projection algorithm used to prevent $b^a(\hat{\Theta}^b,\boldsymbol{x})$ from going to zero, whereas the classical σ-modification, sometimes called also soft projection, is to avoid the estimate of the parameters to drift to infinity (Ioannou and Sun, 1996).

4. SIMULATIONS

In this section, we demonstrate the properties of the proposed adaptive control law using two examples.

Example 1: Consider the following second order nonlinear system

$$\ddot{x}=-4\left(\frac{\sin(4\pi x)}{\pi x}\right)\left(\frac{\sin(\pi\dot{x})}{\pi\dot{x}}\right)^2+\left(2+\sin(2\pi x)+0.1\sin(100t)\right)u \tag{60}$$

$$y=x$$

where the nonlinear functions and parameters are unknown. The output $y=x$ and $\dot{x}$ are assumed to be available for measurement. The output $y(t)$ is required to track a desired trajectory defined by $y_d=\sin(\pi t)$. The magnitude of the tracking error $e=y-y_d$ at steady state is required to be less than 0.05. A one hidden layer radial Gaussian network with a basis function $\zeta_i(\boldsymbol{x})=\exp[-\pi\sigma^2(\boldsymbol{x}-\xi_i)^T(\boldsymbol{x}-\xi_i)]$ is used to approximate $f(x,\dot{x})$ and $b(x)$, which in this case are the ideal bandlimited smooth functions, on a compact set $\Omega=\Omega_x\times\Omega_{\dot{x}}$, where $\Omega_x=\{x\,|\,x\in(-3,3)\}$, $\Omega_{\dot{x}}=\{\dot{x}\,|\,\dot{x}\in(-5,5)\}$,. The mean ξ_i is the center of the radial Gaussian representing the sampling grid, and σ^2 is the variance representing a measure of the width of the radial Gaussian. Note that 0.1sin(100t) is considered as a high-frequency disturbance term and is not estimated. By choosing a sampling grid with mesh size 0.125 and variance 4π, a small uniform approximation bound of $\psi_f\le 0.05$, $\psi_b\le 0.12$, can be achieved (Sanner and Slotine, 1992). Furthermore $(\bar{b}=0.9)>(3\psi_b=0.36)$. The values of

δ_b, Δ, are chosen to be 0.01, 0.04 respectively such that $(\sqrt{\delta_b}+\Delta+3\psi_b=0.5)<(\bar{b}=0.9)$ is satisfied. Using (45a-c), we can calculate $\delta_1=0.18$, $\delta_2=0.026$, $\delta_3=0.154$. By selecting $\lambda=1$ and $\Phi=0.05$, the tracking error is bounded from above by 0.05 at steady state. The design parameters $k_S=2$, $\sigma_v=4.4$, $\sigma_f=4.4$ and $\sigma_b=0.16$ are chosen such that conditions (55a-d) are satisfied. All initial values for $\hat{\Theta}^f(0)$, $\hat{\Theta}^b(0)$ are taken zero. Figures 3 and 4 show the simulation results for the tracking error, and continuous switching function $\rho(t)$.

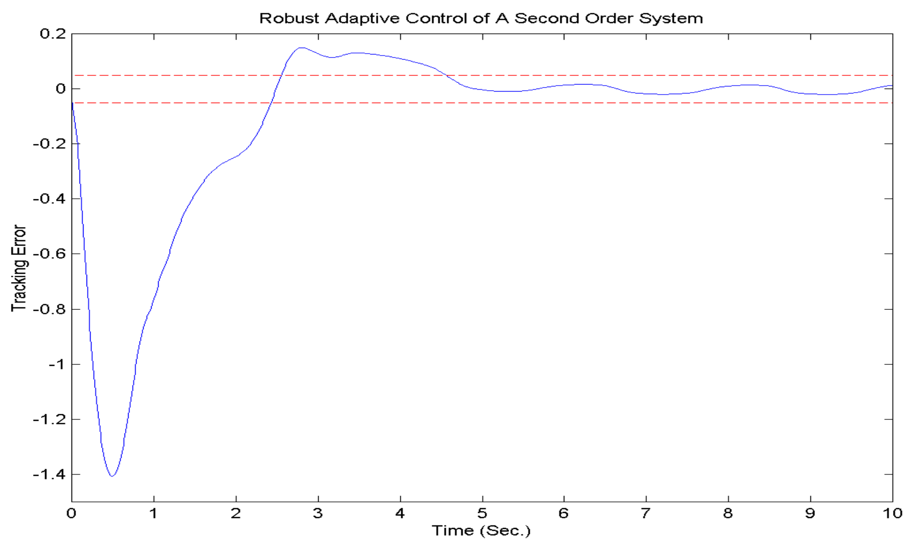

Fig 3. Tracking error during the first 10 seconds
The dashed lines indicate the required error bound

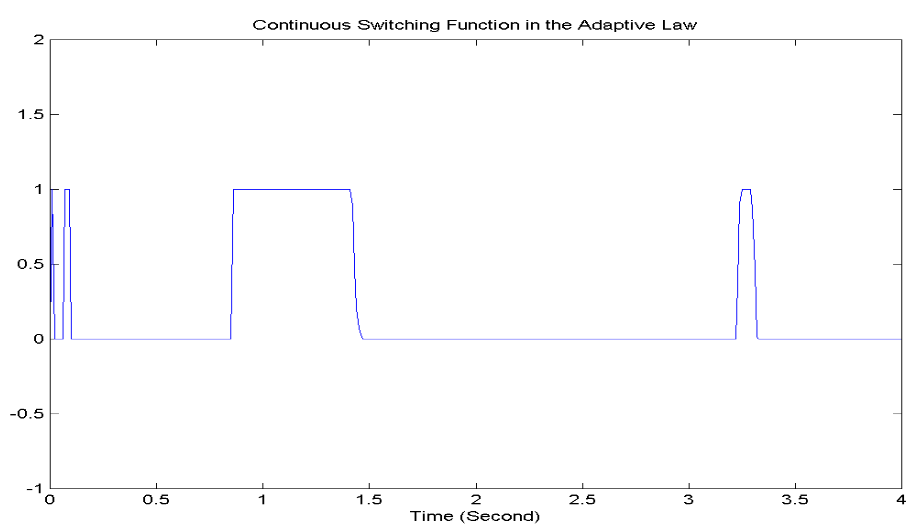

Fig. 4. The continuous switching function $\rho(t)$ in the adaptive law during the first 4 seconds

Example 2: Consider the following nonlinear system

$$\ddot{x} = 3\dot{x}^2 + \{2\exp(x^2/2) + 1.5(\cos x)^2\}u \tag{61}$$

$$y = x$$

where all nonlinear terms and parameters are unknown. The system is initially at $x(0)=1$, $\dot{x}(0)=0.8$. We like to regulate the output $y(t)$ close to zero less than 0.005 say. In this example, one hidden layer radial Gaussian neural network is used to approximate the unknown nonlinearities f, b on a compact set $\Omega = \Omega_x \times \Omega_{\dot{x}}$, where $\Omega_x = \{x \mid x \in (-8,8)\}$, $\Omega_{\dot{x}} = \{\dot{x} \mid \dot{x} \in (-8,8)\}$. The upper bounds of the approximation errors $\psi_f = 0.1$, $\psi_b = 0.1$ are obtained by off-line simulations and training. In this example, $(\bar{b} = 2) > (3\psi_b = 0.3)$. The values of δ_b, Δ, are chosen to be 0.1, 0.1 respectively such that $(\sqrt{\delta_b} + \Delta + 3\psi_b = 0.72) < (\bar{b} = 2)$ is satisfied. Using (45a-c), we obtain $\delta_1 = 0.0855$, $\delta_2 = 0.1053$, $\delta_3 = 0.0526$. By selecting $\lambda = 4.1$, $\Phi = 0.02$, the magnitude of the regulation error is guaranteed to be bounded from above by 0.005 at steady state. The constants $\sigma_v = 4.8$, $\sigma_f = 4.8$, $\sigma_b = 0.13$ and $k_S = 100$ are chosen such that conditions (55a-d) are satisfied. All initial values for $\hat{\theta}_i^f(0)$, $\hat{\theta}_i^b(0)$ are taken zero. Figures 5 and 6 show the simulation results of the regulation error and continuous switching function $\rho(t)$ respectively.

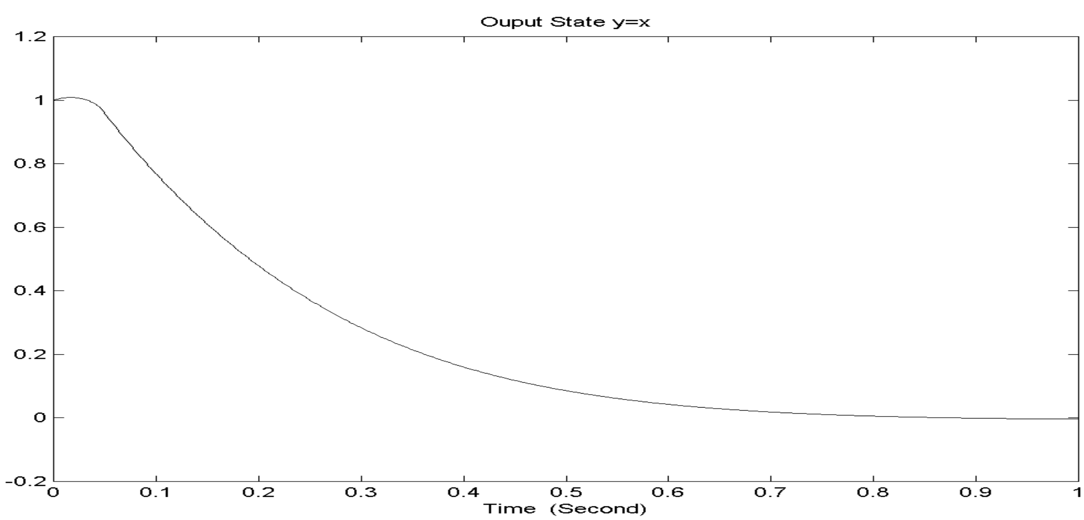

Fig. 5. The regulation of output during the first 1second

Figure 6 demonstrates that switching stops after an initial learning stage and $\rho(t)$ converges to a constant with no further switching.

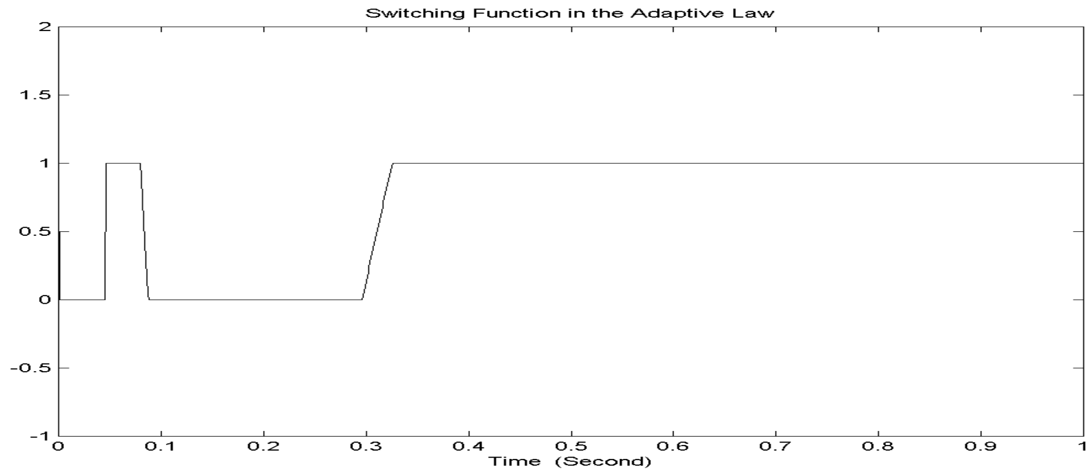

Fig. 6. The switching function $\rho(t)$ during the first 1 second

5. CONCLUSIONS

In this paper, we consider the control problem of a single input feedback linearizable nonlinear system with unknown nonlinearities. The unknown nonlinearities are approximated and estimated on-line using a single layer neural network. A robust adaptive controller scheme is designed that uses the estimated nonlinear functions and employs a number of robust modifications in order to compensate for uncertainties in the estimation. The control scheme guarantees semiglobal stability and convergence of the tracking error to a small residual set even in the case where the estimated plant loses controllability. Semiglobal stability is characterized by a region of attraction for stability whose size depends on the compact set used to approximate the nonlinear functions of the plant. Our results present a methodology for choosing the various design parameters so that the tracking error is guaranteed to converge and remain within a desirable bound at steady state. The extension of these results to a wider class of nonlinear system is currently under investigation.

Acknowledgements: The authors would like to thank Professor Elias Kosmatopoulos of the Technical University of Crete, Professors Vsevolod Kuntsevich, Vyacheslav Gubarev and Drs Leonid Zhiteckij and Nikolay Aksinov of the Space Research Institute of Ukranian National Academy for many useful discussions on adaptive control and results of this paper that took place as a result of a joint collaborative project supported by NATO.

REFERENCES

Aloliwi, B. and H. K. Khalil (1997). Robust Adaptive Output Feedback Control of Nonlinear Systems Without Persistence of Excitation, *Automatica*, **33**, 2025-2032.

Chen, F.-C. and C.-C. Liu (1994). Adaptively Controlling Nonlinear Continuous-Time Systems Using Multilayer Neural Networks, *IEEE Transaction on Automat. Contr.*, **39**, 1306-1310.

Chen, F.-C. and H. K. Khalil (1992). Adaptive Control of Nonlinear Systems Using Neural Networks, *Int. J. Control*, **55**, 1299-1317.

Ioannou, P. A. and J. Sun (1996). *Robust Adaptive Control*, Prentice Hall, Upper Saddle River, NJ.

Kanellakopoulos, I., P. Kokotovic, and A. S. Morse (1991). Systematic Design of Adaptive Controllers for Feedback Linearizable Systems, *IEEE Trans. Automat. Contr.*, **36**, 1241-1253.

Khalil, H. K. (1996). Adaptive Output Feedback Control of Nonlinear Systems Represented by Input-Output Models, *IEEE Trans. Automat. Contr.*, **41**, 177-188.

Kokotovic, P. and M. Arcak (2001). Constructive Nonlinear Control: A Historical Perspective, *Automatica*, **37**, 637-662.

Kosmatopoulos, E. B., M. M. Polycarpou, M. A. Christodoulou, and P. A. Ioannou (1995). High-Order Neural Network Structures for Identification of Dynamical Systems, *IEEE Trans. Neural Networks*, **6**, 422-431.

Kosmatopoulos, E. B. and P. A. Ioannou (1999). A Switching Adaptive Controller for Feedback Linearizable Systems, *IEEE Trans. Automat. Contr.*, **44**, 742-750.

Krstic, M., I. Kanellakopoulos, and P. Kokotovic (1995). *Nonlinear and Adaptive Control Design*, New York: Wiley.

Jankovic, M. (1996). Adaptive output feedback control of nonlinear feedback linearizable systems, *International Journal of Adaptive Control and Signal Processing*, **10**, 1-18.

Lewis, F. L., A. Yesildirek, and K. Liu (1996). Multilayer neural-net robot controller with guaranteed tracking performance, *IEEE Trans. Neural Networks*, **7**, 388-399.

Liu, C.-C. and F.-C. Chen (1993). Adaptive control of nonlinear continuous-time systems using neural networks – general relative degree and MIMO cases, *Int. J. Control*, **58**, 317-335.

Ordonez, R. and K. M. Passino (1999). Stable multi-input multi-output adaptive fuzzy/neural control, *IEEE Trans. Fuzzy Systems*, **7**, 345-353.

Polycarpou, M. M. and M. J. Mears (1998). Stable adaptive tracking of uncertain systems using nonlinearly parametrized on-line approximators, *Int. J. Control*, **70**, 363-384.

Polycarpou, M. M. (1996). Stable Adaptive Neural Control Scheme for Nonlinear Systems, *IEEE Trans. Automat. Contr.*, **41**, 447-451.

Polycarpou, M. M. and P. A. Ioannou (1993). On the Existence and Uniqueness of Solutions in Adaptive Control Systems, *IEEE Trans. Automat. Contr.*, **38**, 474-479.

Polycarpou, M. M. and P. A. Ioannou (1992). Modeling, identification and stable adaptive control of continuous-time nonlinear dynamical systems using neural networks, *Proceedings of ACC'92*, 36-40.

Rovithakis, G. A. (2001). Stable adaptive neuro-control design via Lyapunov function derivative estimation, *Automatica*, **37**, 1213-1221.

Sanner, R. and J. E. Slotine (1992). Gaussian networks for Direct Adaptive Control, *IEEE Trans. Neural Networks*, **3**, 837-863.

Sastry, S. S. and A. Isidori (1989). Adaptive Control of Linearizable Systems, *IEEE Trans. Automat. Contr.*, **34**, 1123-1131.

Seshagiri, S. and H. K. Khalil (2000). Output feedback control of nonlinear systems using RBF neural networks, *IEEE Trans. Neural Networks*, **11**, 69-79.

Seto, D., A. M. Annaswamy, and J. Baillieul (1994). Adaptive control of nonlinear systems with a triangular structure, *IEEE Trans. Automat. Contr.*, **39**, 1411-1428.

Slotine, J. E. and J. A. Coetsee (1986). Adaptive Sliding Controller Synthesis for non-linear Systems, *Int. J. Control*, **43**, 1631-1651.

Slotine, J.E. and W. Li (1991). *Applied nonlinear control*, Prentice Hall, Englewood Cliffs, NJ.

Spooner, J. T. and K. M. Passino (1996). Stable adaptive control using fuzzy systems and neural networks, *IEEE Trans. Fuzzy Systems*, **4**, 339-359.

Taylor, D., P. V. Kokotovic, R. Marino and I. Kanellakopoulos (1989). Adaptive Regulation of Nonlinear Systems with Unmodeled Dynamics, *IEEE Trans. Automat. Contr.*, **34**, 405-412.

Yesildirek, A. and F. L. Lewis (1995). Feedback linearization using neural networks, *Automatica*, **31**, 1659-1664.

APPENDIX-PROOFS OF INEQUALITY (46)

In this appendix, we prove inequality (46) used in the proof of *theorem*. Let us start with the equality

$$d_b(\boldsymbol{x})u - \delta_b u' = \rho\{d_b(\boldsymbol{x})u - \delta_b u'\} + (1-\rho)\{d_b(\boldsymbol{x})u - \delta_b u'\} \tag{A.1}$$

From (35), (40), the function u' can be written as:

$$u' = \frac{1}{(b^a(\hat{\Theta}^b, \boldsymbol{x}))^2 + \delta_b}\bar{u} \tag{A.2}$$

Then,

$$(b^a(\hat{\Theta}^b, \boldsymbol{x}))^2 u' + \delta_b u' = \bar{u} \tag{A.3}$$

Since $b^a(\hat{\Theta}^b, \boldsymbol{x})u' = u$, (A.3) can be written as

$$b^a(\hat{\Theta}^b, \boldsymbol{x})u = \bar{u} - \delta_b u' \tag{A.4}$$

Substituting $b^a(\hat{\Theta}^b, \boldsymbol{x}) = b^a(\Theta^b, \boldsymbol{x}) + b^a(\tilde{\Theta}^b, \boldsymbol{x})$ into (A.4), we obtain

$$b^a(\Theta^b, \boldsymbol{x})u = \bar{u} - \delta_b u' - b^a(\tilde{\Theta}^b, \boldsymbol{x})u \tag{A.5}$$

Using (A.2), it follows that

$$\begin{aligned} b^a(\Theta^b, \boldsymbol{x})u &= \left\{1 - \frac{\delta_b}{(b^a(\hat{\Theta}^b, \boldsymbol{x}))^2 + \delta_b}\right\}\bar{u} - b^a(\tilde{\Theta}^b, \boldsymbol{x})u \\ &= \frac{(b^a(\hat{\Theta}^b, \boldsymbol{x}))^2}{(b^a(\hat{\Theta}^b, \boldsymbol{x}))^2 + \delta_b}\bar{u} - b^a(\tilde{\Theta}^b, \boldsymbol{x})u \end{aligned} \tag{A.6}$$

From (A.6), the control law u can be expressed as

$$u = \left\{\frac{(b^a(\hat{\Theta}^b, \boldsymbol{x}))^2}{(b^a(\hat{\Theta}^b, \boldsymbol{x}))^2 + \delta_b}\right\}\frac{1}{b^a(\Theta^b, \boldsymbol{x})}\bar{u} - \frac{1}{b^a(\Theta^b, \boldsymbol{x})}b^a(\tilde{\Theta}^b, \boldsymbol{x})u \tag{A.7}$$

Using the fact $\left|b^a(\Theta^b, \boldsymbol{x})\right| = \left|b(\boldsymbol{x}) - d_b(\boldsymbol{x})\right| \geq \bar{b} - \psi_b$, we have

$$\begin{aligned} |u| &\leq \frac{1}{\left|b^a(\Theta^b, \boldsymbol{x})\right|}|\bar{u}| + \frac{1}{\left|b^a(\Theta^b, \boldsymbol{x})\right|}\left|b^a(\tilde{\Theta}^b, \boldsymbol{x})\right||u| \\ &\leq \frac{1}{\bar{b} - \psi_b}|\bar{u}| + \frac{1}{\bar{b} - \psi_b}\left|b^a(\tilde{\Theta}^b, \boldsymbol{x})\right||u| \end{aligned} \tag{A.8}$$

Since (A.3) can also be written as

$$\{[b^a(\Theta^b, \boldsymbol{x}) + b^a(\tilde{\Theta}^b, \boldsymbol{x})]^2 + \delta_b\}u' = \bar{u} \tag{A.9}$$

we have

$$\{(b^a(\Theta^b, \boldsymbol{x}))^2 + (b^a(\tilde{\Theta}^b, \boldsymbol{x}))^2 + \delta_b\}u' = \bar{u} - 2b^a(\Theta^b, \boldsymbol{x})b^a(\tilde{\Theta}^b, \boldsymbol{x})u' \tag{A.10}$$

Therefore u' can be expressed as

$$u' = \frac{1}{(b^a(\Theta^b, \boldsymbol{x}))^2 + (b^a(\tilde{\Theta}^b, \boldsymbol{x}))^2 + \delta_b}\bar{u} - \frac{2b^a(\Theta^b, \boldsymbol{x})}{(b^a(\Theta^b, \boldsymbol{x}))^2 + (b^a(\tilde{\Theta}^b, \boldsymbol{x}))^2 + \delta_b}b^a(\tilde{\Theta}^b, \boldsymbol{x})u' \tag{A.11}$$

and,

$$
\begin{aligned}
|u'| &\le \frac{1}{(b^a(\Theta^b,\boldsymbol{x}))^2 + (b^a(\widetilde{\Theta}^b,\boldsymbol{x}))^2 + \delta_b}|\overline{u}| + \frac{2|b^a(\Theta^b,\boldsymbol{x})|}{(b^a(\Theta^b,\boldsymbol{x}))^2 + (b^a(\widetilde{\Theta}^b,\boldsymbol{x}))^2 + \delta_b}|b^a(\widetilde{\Theta}^b,\boldsymbol{x})||u'| \\
&\le \frac{1}{(b^a(\Theta^b,\boldsymbol{x}))^2 + \delta_b}|\overline{u}| + \frac{2|b^a(\Theta^b,\boldsymbol{x})|}{(b^a(\Theta^b,\boldsymbol{x}))^2 + \delta_b}|b^a(\widetilde{\Theta}^b,\boldsymbol{x})||u'| \\
&\le \frac{1}{(b^a(\Theta^b,\boldsymbol{x}))^2 + \delta_b}|\overline{u}| + \frac{2}{|b^a(\Theta^b,\boldsymbol{x})|}|b^a(\widetilde{\Theta}^b,\boldsymbol{x})||u'| \\
&\le \frac{1}{(\overline{b}-\psi_b)^2 + \delta_b}|\overline{u}| + \frac{2}{\overline{b}-\psi_b}|b^a(\widetilde{\Theta}^b,\boldsymbol{x})||u'|
\end{aligned}
\tag{A.12}
$$

Using (A.8) and (A.12), the absolute value of the first term in (A.1) can be written in the following form:

$$
\begin{aligned}
|\rho\{d_b(\boldsymbol{x})u - \delta_b u'\}| &\le \rho|d_b(\boldsymbol{x})||u| + \rho\delta_b|u'| \\
&\le \rho\left\{\frac{\psi_b}{\overline{b}-\psi_b} + \frac{\delta_b}{(\overline{b}-\psi_b)^2+\delta_b}\right\}|\overline{u}| + \rho\frac{2\delta_b}{\overline{b}-\psi_b}|b^a(\widetilde{\Theta}^b,\boldsymbol{x})||u'| + \rho\frac{\psi_b}{\overline{b}-\psi_b}|b^a(\widetilde{\Theta}^b,\boldsymbol{x})||u| \\
&\le \rho\left\{\frac{\psi_b}{\overline{b}-\psi_b-\Delta} + \frac{\delta_b}{(\overline{b}-\psi_b-\Delta)^2+\delta_b}\right\}|\overline{u}| + \rho\frac{2\delta_b}{\overline{b}-\psi_b}|b^a(\widetilde{\Theta}^b,\boldsymbol{x})||u'| + \rho\frac{\psi_b}{\overline{b}-\psi_b}|b^a(\widetilde{\Theta}^b,\boldsymbol{x})||u|
\end{aligned}
\tag{A.13}
$$

The second term in (A.1) can be written as:

$$
(1-\rho)\{d_b(\boldsymbol{x})u - \delta_b u'\} = (1-\rho)\frac{d_b(\boldsymbol{x})b^a(\hat{\Theta}^b,\boldsymbol{x})}{(b^a(\hat{\Theta}^b,\boldsymbol{x}))^2+\delta_b}\overline{u} - (1-\rho)\frac{\delta_b}{(b^a(\hat{\Theta}^b,\boldsymbol{x}))^2+\delta_b}\overline{u}
\tag{A.14}
$$

Since $(1-\rho) \neq 0$ only if $|b^a(\hat{\Theta}^b,\boldsymbol{x})| \ge \overline{b} - \psi_b - \Delta$, we obtain:

$$
\begin{aligned}
|(1-\rho)\{d_b(\boldsymbol{x})u - \delta_b u'\}| &\le (1-\rho)\frac{|d_b(\boldsymbol{x})|}{|b^a(\hat{\Theta}^b,\boldsymbol{x})|}|\overline{u}| + (1-\rho)\frac{\delta_b}{(b^a(\hat{\Theta}^b,\boldsymbol{x}))^2+\delta_b}|\overline{u}| \\
&\le (1-\rho)\left\{\frac{\psi_b}{\overline{b}-\psi_b-\Delta} + \frac{\delta_b}{(\overline{b}-\psi_b-\Delta)^2+\delta_b}\right\}|\overline{u}|
\end{aligned}
\tag{A.15}
$$

From (A.13) and (A.15), (46) follows, i.e.

$$
\begin{aligned}
|\{d_b(\boldsymbol{x})u - \delta_b u'| &\le \left\{\frac{\psi_b}{\overline{b}-\psi_b-\Delta} + \frac{\delta_b}{(\overline{b}-\psi_b-\Delta)^2+\delta_b}\right\}|\overline{u}| + \rho\frac{2\delta_b}{\overline{b}-\psi_b}|b^a(\widetilde{\Theta}^b,\boldsymbol{x})||u'| + \rho\frac{\psi_b}{\overline{b}-\psi_b}|b^a(\widetilde{\Theta}^b,\boldsymbol{x})||u| \\
&= \delta_1|\overline{u}| + \rho\delta_2|b^a(\widetilde{\Theta}^b,\boldsymbol{x})||u'| + \rho\delta_3|b^a(\widetilde{\Theta}^b,\boldsymbol{x})||u|
\end{aligned}
\tag{A.16}
$$

where δ_1, δ_2, δ_3 are as defined in (45a-c).