ROBUST ADAPTIVE CONTROL FOR A CLASS OF MIMO NONLINEAR SYSTEMS WITH GUARANTEED ERROR BOUNDS*

Haojian Xu and Petros A. Ioannou

Department of Electrical Engineering
University of Southern California
Los Angeles, CA 90089
Phone: (213)740-4452
FAX: (213)821-1109
Email: ioannou@usc.edu

Abstract

The design of stabilizing controllers for multi-input multi-output (MIMO) nonlinear plants with unknown nonlinearities is a challenging problem. The high dimensionality coupled with the inability to identify the nonlinearities on-line or off-line accurately, motivates the design of stabilizing controllers based on approximations or on approximate estimates of the plant nonlinearities that are simple enough to be generated in real time. 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 MIMO nonlinear systems with unknown nonlinearities. The controller is continuous and guarantees closed loop semi-global stability and convergence of the tracking error to a small residual set even in the case where the estimated plant becomes uncontrollable. The size of the tracking error at steady state can be specified *a priori* and guaranteed by choosing certain design parameters. A procedure for choosing these parameters is presented. The properties of the proposed control algorithm are demonstrated using simulations.

* 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 systems 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 [1]. 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 in the presence of disturbances and other modeling uncertainties. During the recent years, considerable research efforts have been made to deal with the design of stabilizing controllers for classes of nonlinear plants. These efforts are described in detail in a recent survey paper [2], where a very elegant and informative historical perspective of the evolution of nonlinear control design is presented and discussed. Most of the recent efforts (e.g., surveyed in [2], and [3]-[5]) on nonlinear control design assumed that the plant nonlinearities are known. The case where the plant nonlinearities are products of unknown constant parameters with known nonlinearities gave rise to a number of adaptive control techniques [1], [6]-[18]. For a class of single input single output (SISO) nonlinear systems, an adaptive control scheme based on a min-max optimization approach is proposed in [19]-[21] by assuming that the nonlinear functions are convex/concave (as opposed to linear assumed in almost every adaptive system) with respect to the unknown parameters. The results are extended to include general parameterizations for a class of SISO systems with a triangular structure [22].

Neural approximation together with adaptive control techniques have been proposed by a number of investigators [23]-[45] for controlling nonlinear plants with unknown nonlinear functions. 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. Most of these studies are focused on the SISO systems [23]-[35]. The control of MIMO nonlinear systems with unknown nonlinear functions introduces additional complexities and is considered by several investigators [36]-[44]. The problems that arise in the control of MIMO nonlinear systems and up to date contributions are described below:

Consider the class of MIMO systems described by the following differential equation

$$\boldsymbol{y}^{(r)} = \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{B}(\boldsymbol{x})\boldsymbol{u} \tag{1}$$

where $\boldsymbol{y} = [y_1, \cdots, y_m]^T \in \Re^m$ is the output vector; $\boldsymbol{y}^{(r)} = [y_1^{(r_1)}, \cdots, y_m^{(r_m)}]^T \in \Re^m$; $y_i^{(r_i)} := d^{r_i} y_i / dt^{r_i}$; $\boldsymbol{x} := [y_1, \cdots, y_1^{(r_1-1)}, \cdots, y_m, \cdots, y_m^{(r_m-1)}]^T \in \Re^r$ with $r = r_1 + \cdots + r_m$ is the state vector available for measurement; $\boldsymbol{u} \in \Re^m$ is the control vector; and $\boldsymbol{f}(\boldsymbol{x}) \in \Re^m$, $\boldsymbol{B}(\boldsymbol{x}) \in \Re^{m\times m}$ are smooth unknown functions of the state $\boldsymbol{x}$. A sufficient condition for controllability is that $\underline{\sigma}(\boldsymbol{B}) > 0$, $\forall$ $\boldsymbol{x}$, where $\underline{\sigma}(\boldsymbol{B})$ denotes the minimum singular value of the matrix $\boldsymbol{B}(\boldsymbol{x})$. For control design purposes $\boldsymbol{u}$ is calculated based on the estimated plant which has the same form as the unknown plant, i.e.,

$$\hat{\boldsymbol{y}}^{(r)} = \hat{\boldsymbol{f}}(\boldsymbol{x},t) + \hat{\boldsymbol{B}}(\boldsymbol{x},t)\boldsymbol{u} \tag{2}$$

where $\hat{\boldsymbol{f}}(\boldsymbol{x},t)$, $\hat{\boldsymbol{B}}(\boldsymbol{x},t)$, are the on-line estimate of $\boldsymbol{f}(\boldsymbol{x})$,$\boldsymbol{B}(\boldsymbol{x})$, respectively. Since $\boldsymbol{u}$ is calculated based on the estimated plant, for the control law $\boldsymbol{u}$ to exist we require $\underline{\sigma}(\hat{\boldsymbol{B}}) > 0$, $\forall\boldsymbol{x},t$. The on-line estimators that generate $\hat{\boldsymbol{f}}$, $\hat{\boldsymbol{B}}$ have to guarantee that $\underline{\sigma}(\hat{\boldsymbol{B}}) > 0$, $\forall\boldsymbol{x},t$. In fact for computational purposes and uniform boundedness it is required that $\underline{\sigma}(\hat{\boldsymbol{B}}) > \varepsilon > 0$, $\forall\boldsymbol{x},t$, where ε is a small constant. This requirement is not guaranteed by the usual estimators, for each time t, unless special modifications are introduced and in some cases additional assumptions are made. Several attempts have been made to deal with this so called "stabilizability problem" in the linear as well as nonlinear case [1]. For example in [36], a neural-net controller is analyzed for a class of *mn*th-order MIMO system where $\boldsymbol{B}(\boldsymbol{x})$ is a constant identity matrix. Several other investigators assume that $\boldsymbol{B}(\boldsymbol{x})$ is a known matrix [37]. Most of the previous studies [18], [38]-[40] for the MIMO case are dealing with the control of robots, whose dynamics is a special case of (1). By using some special properties, e.g., skew-symmetric property, the stabilizability problem is bypassed. These control schemes, however, cannot extend to the general class of MIMO nonlinear systems described by (1). In [15], the controllability of the estimated plant (2) is assumed to be true at each time t even though no guarantees are provided that this is the case. This assumption was relaxed using a robust discontinuous control law by assuming that $\boldsymbol{B}(\boldsymbol{x})$ is positive for all $\boldsymbol{x}$, a lower bound of the norm of $\boldsymbol{B}(\boldsymbol{x})$, and an upper bound of the norm of $\boldsymbol{f}(\boldsymbol{x})$ are known. However, the discontinuous control leads to a nonlinear system with discontinuities and the existence and uniqueness of solutions of the closed loop system cannot be guaranteed [45]. Furthermore, the discontinuous control law may cause chattering at certain boundaries with adverse effects on performance. In [16]-[17], $\boldsymbol{B}(\boldsymbol{x})$ is assumed to contain some unknown constant vector Θ as $\boldsymbol{B}(\boldsymbol{x},\Theta)$ where $\boldsymbol{B}(\boldsymbol{x},\Theta)$ is linear with respective to Θ. In this case, no modeling errors are included in the model for $\boldsymbol{B}(\boldsymbol{x})$. It is assumed that a convex set Ω is known *a priori* such that $\Theta \in \Omega$ implies $\underline{\sigma}(\boldsymbol{B}) > 0$. Projection is used to guarantee $\hat{\Theta}(t) \in \Omega$, $\forall\, t \geq 0$. However, in general it is difficult if at all possible to establish a convex set Ω with $\Theta \in \Omega$ even in the linear case [1]. In [41], the stabilizability problem was discussed without any stability analysis. In [42]-[43] each element of $\boldsymbol{B}(\boldsymbol{x})$ is approximated by neural networks of the form $b_{ij}(\boldsymbol{x}) = b_{ij}^a(\boldsymbol{x},\Theta_{ij}) + d_{ij}(\boldsymbol{x})$ where b_{ij}^a is an approximated function constructed by neural networks, Θ_{ij}, $i,j=1,2,\ldots,m$, is a constant parameter vector corresponding to some unknown weights of the neural network, and d_{ij} denotes the approximation error. The weights Θ_{ij} are estimated on line generating $\hat{\Theta}_{ij}(t)$, the estimate of Θ_{ij} at each time *t*, which in turn is used to generate $b_{ij}^{a}(\boldsymbol{x},\hat{\Theta}_{ij})$ and therefore $\hat{\boldsymbol{B}}^{a}$, the estimate of $\boldsymbol{B}^{a}=\left[b_{ij}^{a}(\boldsymbol{x},\Theta_{ij})\right]_{m\times m}$ at each time *t*. It is assumed that $\hat{\Theta}_{ij}(0)$ is close to the actual value Θ_{ij} and $\hat{\Theta}_{ij}(t)$ is updated slowly by choosing small adaptive gains. Based on this condition it is established that $\hat{\Theta}_{ij}(t)$ stays inside an invariant set where $\underline{\sigma}(\hat{\boldsymbol{B}}^{a})>0$. However, since Θ_{ij} is unknown and does not have any physical meaning, the assumption that $\hat{\Theta}_{ij}(0)$ is close to the desired values is difficult if at all possible to guarantee. In [44], an indirect adaptive control algorithm based on fuzzy-neural systems is proposed for the class of systems (1) provided convex sets Ω_{ij} can be constructed such that $\underline{\sigma}(\hat{\boldsymbol{B}}^{a})\geq\sigma_{\min}$ and $\overline{\sigma}(\hat{\boldsymbol{B}}^{a})\geq\sigma_{\max}$ for all $\hat{\Theta}_{ij}(t)\in\Omega_{ij}$, i,j=1,2,…,m, where $\overline{\sigma}(\hat{\boldsymbol{B}}^{a})$ denotes the maximum and $\underline{\sigma}(\hat{\boldsymbol{B}}^{a})$ the minimum singular value of $\hat{\boldsymbol{B}}^{a}$ respectively and $\sigma_{\min}$, $\sigma_{\max}$ are two positive constants. The estimated weights of the fuzzy/neural systems are then guaranteed to remain inside Ω_{ij} by using projection. However, such sets Ω_{ij} are difficult to construct in general even if the weights were known let alone to know them *a priori* where such weights are completely unknown. If sets Ω_{ij} are constructed to be convex in the parameter space, there is no guarantee that the unknown Θ_{ij} that corresponds to the "optimal" approximation belongs to these sets. If $\Theta_{ij}\notin\Omega_{ij}$ the possibility of instability or bad performance cannot be excluded. In [44], it was shown that if $\boldsymbol{B}(\boldsymbol{x})$ is strictly diagonally dominant with known lower and upper bounds for the main diagonal entries, and if the first derivatives of the main diagonal entries in $\boldsymbol{B}(\boldsymbol{x})$ are upper bounded by known functions and the upper bounds of all off-diagonal elements $b_{ij}(\boldsymbol{x})$, i,j=1,…,m, $i\neq j$, are known, the MIMO system can be decoupled into a set of SISO subsystems. Then a controller of each subsystem can be modeled by a fuzzy system plus a robust term. A direct fuzzy adaptive control scheme is designed for each SISO subsystem provided the fuzzy model is upper bounded by a known continuous function. All off diagonal entries are treated as modeling errors.

Another important issue in adaptive nonlinear control with unknown nonlinearities is that of tracking error 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 shown to be of the order of the bound on the modeling or approximation error [14],[17],[27]-[29],[41]. In some cases the approximation error is assumed to be upper bounded by some known nonlinearities [14],[28] leading to an upper bound for the tracking error at steady state that is a function of some design parameters. In [30]-[31],[42]-[43] the tracking error is shown to converge inside a small residual set that depends on bounds of unknown signals and design parameters. In [15],[35]-[40], the tracking error may be made smaller by increasing the control gain. In all these cases, it is no clear, from the analysis how changes in the design parameters to improve tracking performance will affect the region of attraction for signal boundedness. Furthermore, the upper bound for the tracking error cannot be easily computed and therefore the controller gain cannot be designed *a priori* to achieve a desired tracking error bound at steady state.

In this paper, we develop a robust adaptive control scheme for the class of MIMO nonlinear linearizable systems with unknown nonlinearities described by (1). The only assumptions made are that the unknown nonlinear functions are smooth, and a sufficient condition for controllability is satisfied. We propose a control law that bypasses the stabilizability problem for the MIMO system. The estimate of the unknown matrix $\boldsymbol{B}(\boldsymbol{x})$ is replaced by the estimate of a scalar function. The adaptive laws use a continuous switching function instead of complicated projection techniques that introduce discontinuities. The proposed scheme guarantees semi-global stability and convergence of the tracking error to a residual set whose size depends on design parameters that can be chosen *a priori*. For any given desired upper bound for the tracking error at steady state, our approach provides a procedure for choosing the design parameters to meet the tracking error bound. Compared to our results for the SISO system [34], the proposed control scheme uses a new unit vector of the tracking errors to assign the control energy to different channels. A new dead zone technique for the MIMO case is incorporated in the adaptive law in order to guarantee closed loop stability and robustness with respect to modeling errors. The adaptive law uses a new continuous switching function to guarantee closed loop system stability and convergence of the tracking error even in the case where the estimated plant loses controllability. The tracking error in each channel converges to a small residual set which can be guaranteed by choosing design parameters appropriately following a design procedure presented in the paper.

This paper is organized as follows: In section 2 the problem statement and preliminaries are presented. In section 3 we present the proposed adaptive control scheme is discussed. In section 4, the theoretical results are applied to the control of a two-link robot. Finally, section 5 includes the conclusions. Throughout this paper, $|\cdot|$ indicates the absolute value, and $\|\cdot\|$ indicates the Euclidean vector norm.

2. PROBLEM STATEMENT AND PRELIMINARIES

Consider the MIMO nonlinear system described by (1), i.e.,

$$\begin{aligned} x_1^{(r_1)} &= f_1(\boldsymbol{x}) + b_{11}(\boldsymbol{x})u_1 + \cdots + b_{1m}(\boldsymbol{x})u_m \\ &\vdots \\ x_m^{(r_m)} &= f_m(\boldsymbol{x}) + b_{m1}(\boldsymbol{x})u_1 + \cdots + b_{mm}(\boldsymbol{x})u_m \end{aligned} \tag{3}$$

$$y_1 = x_1, \cdots, y_m = x_m$$

where, $x_i^{(r_i)} := d^{r_i} x_i / dt^{r_i}$, $\boldsymbol{x} := [x_1 \ \cdots \ x_1^{(r_1-1)}, \ \cdots, \ x_m \ \cdots \ x_m^{(r_m-1)}]^T \in \Re^r$ with $r = r_1 + \cdots + r_m$, is the overall state vector, $u_i \in \Re$, $i = 1,2,\cdots,m$, are the inputs and $y_i \in \Re$, $i = 1,2,\cdots,m$, are the outputs of the system. The nonlinear functions $f_1,\cdots,f_m$ and $b_{ij}, i, j = 1,\cdots,m$ are assumed to be smooth functions. The problem is to design the control input $\boldsymbol{u} = [u_1,\cdots,u_m]^T \in \Re^m$ such that the outputs of the system $y_1,\cdots,y_m$ track the desired trajectories $y_{d_1}(t),\cdots,y_{d_m}(t)$ respectively as close as possible.

The system (3) can also be written in the compact form

$$[x_1^{(r_1)} \quad x_2^{(r_2)} \quad \cdots \quad x_m^{(r_m)}]^T = \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{B}(\boldsymbol{x})\boldsymbol{u} \tag{4}$$

where,

$$\boldsymbol{f}(\boldsymbol{x}) = [f_1(\boldsymbol{x}) \quad \cdots \quad f_m(\boldsymbol{x})]^T \in \Re^m \tag{5}$$

$$\boldsymbol{B}(\boldsymbol{x}) = \begin{bmatrix} b_{11}(\boldsymbol{x}) & \cdots & b_{1m}(\boldsymbol{x}) \\ \vdots & \ddots & \vdots \\ b_{m1}(\boldsymbol{x}) & \cdots & b_{mm}(\boldsymbol{x}) \end{bmatrix} \in \Re^{m\times m} \tag{6}$$

We make the following assumptions:

***Assumption* 1**: The matrix $\frac{1}{2}\left(\boldsymbol{B}(\boldsymbol{x}) + \boldsymbol{B}^T(\boldsymbol{x})\right)$ is known to be either uniformly positive definite or uniformly negative definite for all $\boldsymbol{x}\in\Omega$ where $\Omega \subset \Re^r$ is a compact set, i.e.,

$$\underline{\sigma}\left(\frac{\boldsymbol{B}(\boldsymbol{x}) + \boldsymbol{B}^T(\boldsymbol{x})}{2}\right) \geq \underline{b} > 0, \quad \forall\, \boldsymbol{x} \in \Omega \tag{7}$$

where $\underline{\sigma}(\cdot)$ represents the smallest singular value of the matrix inside the bracket and $\underline{b}$ is its lower bound.

Assumption 1 guarantees that the nonlinear system (3) is uniformly strongly controllable.

***Assumption* 2**: The desired trajectories $y_{d_i}(t)$, i=1,2,…,m, are known bounded functions of time with bounded known derivatives and $\boldsymbol{y}_d := [y_{d_1} \quad \cdots \quad y_{d_1}^{(r_1-1)}, \quad \cdots, \quad y_{d_m} \quad \cdots \quad y_{d_m}^{(r_m-1)}]^T$, $\boldsymbol{y}_d \in \Omega_d \subset \Re^r$, where Ω_d is a known compact set.

***Assumption* 3**: The state $\boldsymbol{x}$ of the system is available for measurement.

***Assumption* 4**: The functions $f_i(\boldsymbol{x})$ and $b_{ij}(\boldsymbol{x})$, $i, j = 1,2,\cdots,m$, are smooth functions but otherwise completely unknown.

Let us first consider the case where $f_i(\boldsymbol{x})$ and $b_{ij}(\boldsymbol{x})$, $i, j = 1,2,\cdots,m$, are completely known and examine whether we can meet the control objective. This is a reasonable step to take since if we cannot meet the control objective in the case of known nonlinearities, it is unlikely that we will do so in the case of unknown nonlinearities.

We define the following error metric, S_i, that describes the desired dynamics of the error system:

$$\begin{aligned} &S_1(t) = (d/dt + \lambda_1)^{r_1-1} e_1(t), \qquad e_1(t) = y_1 - y_{d_1} \\ &\vdots \\ &S_m(t) = (d/dt + \lambda_m)^{r_m-1} e_m(t), \quad e_m(t) = y_m - y_{d_m} \end{aligned} \tag{8}$$

where $\lambda_1,\cdots,\lambda_m$ are positive constants to be selected. It follows from (8) that for $S_i(t) = 0,\ i = 1,2,\cdots,m$, we have a set of linear differential equations whose solutions $e_i(t)$, i=1,2,…,m converges to zero with time constant $(r_i - 1)/\lambda_i$. In addition all the derivatives of $e_i(t)$ up to $r_i - 1$ also converge to zero [3]-[4]. It follows from (8) that

$$\begin{bmatrix} \dot{S}_1 \\ \vdots \\ \dot{S}_m \end{bmatrix} = \begin{bmatrix} f_1(\boldsymbol{x}) \\ \vdots \\ f_m(\boldsymbol{x}) \end{bmatrix} + \begin{bmatrix} v_1(t) \\ \vdots \\ v_m(t) \end{bmatrix} + \begin{bmatrix} b_{11}(\boldsymbol{x}) & \cdots & b_{1m}(\boldsymbol{x}) \\ \vdots & \ddots & \vdots \\ b_{m1}(\boldsymbol{x}) & \cdots & b_{mm}(\boldsymbol{x}) \end{bmatrix} \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix} \tag{9}$$

where,

$$\begin{aligned} &v_1(t) = -y_{d_1}^{(r_1)}(t) + \alpha_{1,r_1-1} e_1^{(r_1-1)}(t) + \cdots + \alpha_{1,1}\dot{e}_1(t) \\ &\vdots \\ &v_m(t) = -y_{d_m}^{(r_m)}(t) + \alpha_{m,r_m-1} e_m^{(r_m-1)}(t) + \cdots + \alpha_{m,1}\dot{e}_m(t) \end{aligned} \tag{10}$$

and $\alpha_{i,r_i-1},\cdots,\alpha_{i,1}$, i=1,2,…,m, are coefficients of the binomial expansion of the corresponding error metric in (8).

Equation (9) can be written in the compact form:

$$\dot{\boldsymbol{S}} = \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{v}(t) + \boldsymbol{B}(\boldsymbol{x})\boldsymbol{u} \tag{11}$$

where,

$$\boldsymbol{S}(t) = [S_1(t) \quad \cdots \quad S_m(t)]^T \in \Re^m \tag{12}$$

$$\boldsymbol{v}(t) = [v_1(t) \quad \cdots \quad v_m(t)]^T \in \Re^m \tag{13}$$

Let us consider the Lyapunov-like function

$$V(t) = \frac{1}{2}\boldsymbol{S}^T\boldsymbol{S} \tag{14}$$

Then the time derivative of $V(t)$ is given by

$$\dot{V} = \boldsymbol{S}^T\dot{\boldsymbol{S}} = \boldsymbol{S}^T\boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{S}^T\boldsymbol{v}(t) + \boldsymbol{S}^T\boldsymbol{B}(\boldsymbol{x})\boldsymbol{u} \tag{15}$$

If we now choose the control law $\boldsymbol{u}$ so that

$$\dot{V} = \boldsymbol{S}^T\dot{\boldsymbol{S}} < 0 \tag{16}$$

then it can be shown that $\lim_{t\to\infty}\|\boldsymbol{S}\| = 0$.

If $\boldsymbol{B}(\boldsymbol{x})$ is invertible, then the control law [3]

$$\boldsymbol{u} = \boldsymbol{B}^{-1}(\boldsymbol{x})\{-\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{v}(t) - \boldsymbol{KS}\} \tag{17}$$

where $\boldsymbol{K} = \mathrm{diag}(k_1, \cdots, k_m)$, and $k_i > 0$, i=1,2,…,m, guarantees

$$\dot{V} = -\sum_{i=1}^{m} k_i S_i^2 \le 0 \tag{18}$$

which implies that $V(t)$ and therefore $S_i(t)$ converge to zero exponentially fast.

The calculation of the inverse of $\boldsymbol{B}(\boldsymbol{x})$ in the control law (17) can be avoided if we use assumption 1 and the following lemma to modify the control law (17).

***Lemma* 1**: Let us define $\boldsymbol{X} := [\boldsymbol{x}, \boldsymbol{y}_d]^T$. For any square matrix $\boldsymbol{B}(\boldsymbol{x})$ satisfying *Assumption* 1, there exists a smooth scalar function $\mu(\boldsymbol{X})$ such that for any vector $\boldsymbol{S}(\boldsymbol{X})$, we have $\boldsymbol{S}^T(\boldsymbol{X})\boldsymbol{B}(\boldsymbol{x})\boldsymbol{S}(\boldsymbol{X}) = \mu(\boldsymbol{X})\|\boldsymbol{S}(\boldsymbol{X})\|^2$, where $|\mu(\boldsymbol{X})| > \underline{b}$.

Proof: The proof is presented in Appendix A.

Instead of (17) let us now consider the control law:

$$\boldsymbol{u} = \frac{\boldsymbol{S}}{\boldsymbol{S}^T\boldsymbol{B}(\boldsymbol{x})\boldsymbol{S}}\left[-k_S\|\boldsymbol{S}\|^2 - \boldsymbol{S}^T\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{S}^T\boldsymbol{v}(t)\right] \tag{19}$$

where $k_S > 0$ is a design constant. Using lemma 1 we can rewrite (19) as

$$\boldsymbol{u} = \frac{\boldsymbol{S}_e}{\mu(\boldsymbol{X})}\left[-k_S\|\boldsymbol{S}\| - \boldsymbol{S}_e^T \boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{S}_e^T \boldsymbol{v}(t)\right] \tag{20}$$

where $\boldsymbol{S}_e = \boldsymbol{S} / \|\boldsymbol{S}\|$ is the unit vector. The control law (20) guarantees that

$$\dot{V} = \boldsymbol{S}^T \dot{\boldsymbol{S}} = -k_S \|\boldsymbol{S}\|^2 \le 0 \tag{21}$$

which in turn implies that $\lim_{t\to\infty} S_i(t) = 0$ exponentially fast and therefore $e_i(t)$ and all its derivatives up to $r_i - 1$ converge to zero exponentially fast.

In the case where $\boldsymbol{f}(\boldsymbol{x})$ and $\boldsymbol{B}(\boldsymbol{x})$ are unknown nonlinear functions, the control law (20) can no longer be used. We assume that the nonlinear functions $\boldsymbol{f}(\boldsymbol{x})$ and $\mu(\boldsymbol{X})$ can be approximated by a general one layer neural network [25]-[29],[33], [46]-[48] on compact sets $\boldsymbol{x} \in \Omega \subset \Re^r$ and $\boldsymbol{X} \in \Omega \times \Omega_d \subset \Re^{2r}$ respectively as

$$f_i(\boldsymbol{x}) \approx f_i^a(\boldsymbol{x}) = \sum_{j=1}^{l_i} \theta_{ij}\zeta_{ij}(\boldsymbol{x}),\ i = 1,2,\cdots,m \tag{22a}$$

$$\mu(\boldsymbol{X}) \approx \mu^a(\boldsymbol{X}) = \sum_{j=1}^{l_0} \theta_{0j}\zeta_{0j}(\boldsymbol{X}) \tag{22b}$$

where $\zeta_{0j}(\boldsymbol{X})$, $\zeta_{ij}(\boldsymbol{x})$, i=1,2,…,m, j=1,2,…, l_i, are selected basis functions, θ_{0j}, θ_{ij} are unknown constant parameters, and l_0, l_i are the number of the nodes respectively.

The neural network approximation errors $d_{f_i}(\boldsymbol{x})$, $d_\mu(\boldsymbol{X})$ are given by

$$d_{f_i}(\boldsymbol{x}) := f_i(\boldsymbol{x}) - f_i^a(\boldsymbol{x}) \tag{23a}$$

$$d_\mu(\boldsymbol{X}) := \mu(\boldsymbol{X}) - \mu^a(\boldsymbol{X}) \tag{23b}$$

Let $\Theta_0 = [\theta_{01}, \theta_{02} \cdots, \theta_{0l_0}]^T$, $\Theta_i = [\theta_{i1}, \theta_{i2} \cdots, \theta_{il_i}]^T$ where i=1,2,…,m. By "optimal" approximation we mean the weights $\Theta_0 \in \Re^{l_0}$, $\Theta_i \in \Re^{l_i}$ are chosen to minimize $d_{f_i}(\boldsymbol{x})$, $d_\mu(\boldsymbol{X})$ for all $\boldsymbol{x} \in \Omega \subset \Re^r$ and $\boldsymbol{X} \in \Omega \times \Omega_d \subset \Re^{2r}$ respectively, i.e.,

$$\Theta_0 := \arg\min_{\Theta \in \Re^{l_0}} \{ \sup_{\boldsymbol{X} \in \Omega \times \Omega_d} |\mu(\boldsymbol{X}) - \mu^a(\boldsymbol{X})| \} \tag{24a}$$

$$\Theta_i := \arg\min_{\Theta \in \Re^{l_i}} \{ \sup_{\boldsymbol{x} \in \Omega} |f_i(\boldsymbol{x}) - f_i^a(\boldsymbol{x})| \},\ i=1,2,\ldots,m \tag{24b}$$

***Assumption* 5**: There exists a set of constant parameters Θ_i, i=0,1,…,m, referred to as optimal output weights such that the approximation error $d_{f_i}(\boldsymbol{x})$ is upper bounded by a known constant $\psi_f > 0$ over the compact set $\Omega \subset \Re^r$ and the approximation error $d_\mu(\boldsymbol{X})$ is upper bounded by a known constant $\psi_\mu > 0$ over the compact set $\Omega \times \Omega_d \subset \Re^{2r}$, i.e.,

$$\sup_{\boldsymbol{x}\in\Omega} \left| d_{f_i}(\boldsymbol{x}) \right| \le \psi_f, \quad i=1,\ldots,m, \quad \forall\, \boldsymbol{x} \in \Omega \tag{25a}$$

$$\sup_{\boldsymbol{X}\in\Omega\times\Omega_d} \left| d_\mu(\boldsymbol{X}) \right| \le \psi_\mu, \quad \forall\, \boldsymbol{X} \in \Omega \times \Omega_d \tag{25b}$$

We should note that we don't require the approximation errors to arbitrarily small. It is also assumed that the basis functions, number of nodes l_i, i=0,1,…,m, are specified by the designer, and the only unknowns are the output weights. As shown in [27], [33], [46]-[48] and the references therein, different basis functions can be used to satisfy assumption 5.

Remark 1: We should note that we approximate the unknown scalar function $\mu(\boldsymbol{X})$ instead of the unknown matrix $\boldsymbol{B}(\boldsymbol{x})$. The price paid in this case is that the dimension of $\boldsymbol{X} \in \Re^{2r}$ is doubled that of the state $\boldsymbol{x} \in \Re^r$. However, for the regulation problem, $\mu(\boldsymbol{X})$ can be replaced by $\mu(\boldsymbol{x})$.

Let us now assume that the approximation functions $f_i^a(\boldsymbol{x})$, $\mu^a(\boldsymbol{X})$ instead of the actual ones $f_i(\boldsymbol{x})$ and $\mu(\boldsymbol{X})$ are known. In this case replacing $f_i(\boldsymbol{x})$ and $\mu(\boldsymbol{X})$ with $f_i^a(\boldsymbol{x})$, $\mu^a(\boldsymbol{X})$ in the control law (20) will be a straightforward approach. The stability and performance analysis of the closed loop system, however, is difficult due to the approximation errors $d_{f_i}(\boldsymbol{x})$, $d_\mu(\boldsymbol{X})$ being different than zero. In order to deal with the approximation error, we modify (20) and propose the control law

$$\boldsymbol{u} = \frac{\boldsymbol{S}_1}{\mu^a(\boldsymbol{X})}\left[-k_S\left\|\boldsymbol{S}\right\| - \boldsymbol{S}_1^T \boldsymbol{f}^a(\boldsymbol{x}) - \boldsymbol{S}_1^T \boldsymbol{v}(t) - \sigma_v \left\|\boldsymbol{v}(t)\right\| - \sigma_f \left\|\boldsymbol{f}^a(\boldsymbol{x})\right\|\right] \tag{26}$$

$$\boldsymbol{S}_1 = \begin{cases} \boldsymbol{S}/\left\|\boldsymbol{S}\right\|, & \text{if } \left\|\boldsymbol{S}\right\| > \Phi \\ \boldsymbol{S}/\Phi, & \text{if } \left\|\boldsymbol{S}\right\| \le \Phi \end{cases} \tag{27}$$

where $\boldsymbol{f}^a(\boldsymbol{x}) := [f_1^a(\boldsymbol{x}), \cdots, f_m^a(\boldsymbol{x})]$; $\sigma_v > 0$, $\sigma_f > 0$, $\Phi > 0$ are small design constants, and $k_S > 0$ is a constant chosen by the designer. The small linear boundary layer characterized by Φ is used to smooth out the control discontinuity and avoid possible singularities in calculating $\boldsymbol{S}_1$.

The design parameter Φ can be incorporated into $\left\|\boldsymbol{S}\right\|$ as a deadzone width by defining a new function S_Φ as

$$S_\Phi := \left\|\boldsymbol{S}\right\| - \Phi\,\mathrm{sat}(\left\|\boldsymbol{S}\right\|/\Phi) \tag{28}$$

$$\mathrm{sat}(\|\boldsymbol{S}\|/\Phi)=\begin{cases}1, & \text{if } \|\boldsymbol{S}\|>\Phi \\ \|\boldsymbol{S}\|/\Phi, & \text{if } \|\boldsymbol{S}\|\le\Phi\end{cases} \tag{29}$$

The properties of S_Φ are described by the following lemma.

***Lemma* 2**: The function S_Φ defined in (28) has the following properties:

$$S_\Phi=0,\ \dot{S}_\Phi=0,\ \text{if } \|\boldsymbol{S}\|\le\Phi \tag{30a}$$

$$\dot{S}_\Phi=\boldsymbol{S}_1^T\dot{\boldsymbol{S}},\ \text{if } \|\boldsymbol{S}\|>\Phi \tag{30b}$$

Proof: Form (28),(29), we have $S_\Phi=0$ when $\|\boldsymbol{S}\|\le\Phi$ which also implies that $\dot{S}_\Phi=0$ for $\|\boldsymbol{S}\|\le\Phi$. If $\|\boldsymbol{S}\|>\Phi$, $S_\Phi=\|\boldsymbol{S}\|-\Phi$, then $\dot{S}_\Phi=d(\|\boldsymbol{S}\|)/dt=\boldsymbol{S}^T\dot{\boldsymbol{S}}/\|\boldsymbol{S}\|=\boldsymbol{S}_1^T\dot{\boldsymbol{S}}$. ■

The following theorem establishes the stability and performance properties of the closed loop system with (26) as the control law.

***Theorem* 1**: Consider the system (3) and the control law (26). Assume that assumptions 1-5 hold. Let ψ_f,ψ_μ, be some positive constants that satisfy (25a-b) over the compact sets $\Omega\subset\Re^r$ and $\Omega\times\Omega_d\subset\Re^{2r}$ for some $\Theta_0\in\Re^{l_0}$, $\Theta_i\in\Re^{l_i}$, i=1,2,…,m. Moreover the lower bound $\underline{b}$ of $\mu(\boldsymbol{X})$ satisfies the condition $\underline{b}>2\psi_\mu$. Then for any arbitrary small positive constant Φ, there exists a positive constant $\delta_0<1$, such that if $k_S>\dfrac{\sqrt{m}\psi_f}{(1-\delta_0)\Phi}$, $\sigma_v\ge\dfrac{\delta_0}{1-\delta_0}$, $\sigma_f\ge\dfrac{\delta_0}{1-\delta_0}$, for all $\boldsymbol{x}(0)\in\Omega_{\boldsymbol{x}}\subset\Omega$, all signals in the closed-loop system (3) are bounded and the tracking errors $e_i(t)$ converge to the residual set $R_{e_i}=\{e_i\big|\ |e_i(t)|\le\lambda_i^{-r_i+1}\Phi\}$, i=1,2…,m, as $t\to\infty$ exponentially fast.

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

$$V(t)=\frac{1}{2}S_\Phi^2 \tag{31}$$

Note that $V(t)$=0 implies that $S_\Phi=0$ and $\|\boldsymbol{S}\|\le\Phi$. The time derivative $V(t)$ is given by

$$\dot{V}(t)=0,\ \text{if } \|\boldsymbol{S}\|\le\Phi \tag{32a}$$

$$\dot{V}(t)=S_\Phi\dot{S}_\Phi,\ \text{if } \|\boldsymbol{S}\|>\Phi \tag{32b}$$

In the following proof, we only consider the region $\|\boldsymbol{S}\|>\Phi$. Rewrite $\boldsymbol{u}$ in (26) as:

$$\boldsymbol{u} = \boldsymbol{S}_1 u_0 \tag{33}$$

$$u_0 = \frac{1}{\mu^a(\boldsymbol{X})}\left[-k_S\|\boldsymbol{S}\| - \boldsymbol{S}_1^T \boldsymbol{f}^a(\boldsymbol{x}) - \boldsymbol{S}_1^T \boldsymbol{v}(t) - \sigma_v\|\boldsymbol{v}(t)\| - \sigma_f\left\|\boldsymbol{f}^a(\boldsymbol{x})\right\|\right] \tag{34}$$

Using Lemma 2 and $\|\boldsymbol{S}\| = S_\Phi + \Phi$, we have

$$\begin{aligned}\dot{S}_\Phi &= \boldsymbol{S}_1^T\dot{\boldsymbol{S}} = \boldsymbol{S}_1^T\boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{S}_1^T\boldsymbol{v}(t) + \boldsymbol{S}_1^T\boldsymbol{B}(\boldsymbol{x})\boldsymbol{u} \\ &= \boldsymbol{S}_1^T\boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{S}_1^T\boldsymbol{v}(t) + \mu(\boldsymbol{X})u_0 \\ &= -k_S\|\boldsymbol{S}\| - \sigma_v\|\boldsymbol{v}(t)\| - \sigma_f\left\|\boldsymbol{f}^a(\boldsymbol{x})\right\| + \boldsymbol{S}_1^T[\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}^a(\boldsymbol{x})] + [\mu(\boldsymbol{X}) - \mu^a(\boldsymbol{X})]u_0 \\ &= -k_S S_\Phi - k_S\Phi - \sigma_v\|\boldsymbol{v}(t)\| - \sigma_f\left\|\boldsymbol{f}^a(\boldsymbol{x})\right\| + \boldsymbol{S}_1^T\boldsymbol{d}_f(\boldsymbol{x}) + d_\mu(\boldsymbol{X})u_0\end{aligned} \tag{35}$$

where $\boldsymbol{d}_f(\boldsymbol{x}) := [d_{f_1},\cdots,d_{f_m}]^T$. The term $S_1^T\boldsymbol{d}_f(\boldsymbol{x}) + d_\mu(\boldsymbol{X})u_0$ in (35) is a modeling error term representing the effect of the approximation errors $\boldsymbol{d}_f(\boldsymbol{x})$, $d_\mu(\boldsymbol{X})$. Using the inequality $\left\|\boldsymbol{d}_f\right\| \le \sqrt{m}\psi_f$, it follows that

$$\left|\boldsymbol{S}_1^T\boldsymbol{d}_f(\boldsymbol{x}) + d_\mu(\boldsymbol{X})u_0\right| \le \sqrt{m}\psi_f + \psi_\mu|u_0| \tag{36}$$

Using $\left|\mu^a(\boldsymbol{X})\right| \ge \underline{b} - \psi_\mu$, we have,

$$\begin{aligned}\left|\boldsymbol{S}_1^T\boldsymbol{d}_f(\boldsymbol{x}) + d_\mu(\boldsymbol{X})u_0\right| &\le \sqrt{m}\psi_f + \delta_0 k_S\|\boldsymbol{S}\| + \delta_0(1+\sigma_v)\|\boldsymbol{v}(t)\| + \delta_0(1+\sigma_f)\left\|\boldsymbol{f}^a(\boldsymbol{x})\right\| \\ &= \sqrt{m}\psi_f + \delta_0 k_S\Phi + \delta_0 k_S S_\Phi + \delta_0(1+\sigma_v)\|\boldsymbol{v}(t)\| + \delta_0(1+\sigma_f)\left\|\boldsymbol{f}^a(\boldsymbol{x})\right\|\end{aligned} \tag{37}$$

Where δ_0 is defined as:

$$\delta_0 := \frac{\psi_\mu}{\underline{b} - \psi_\mu} \tag{38}$$

Using (35) and (37), $\dot{V}$ becomes

$$\begin{aligned}\dot{V} &= -k_S S_\Phi^2 - k_S S_\Phi\Phi - \sigma_v\|\boldsymbol{v}(t)\|S_\Phi - \sigma_f\left\|\boldsymbol{f}^a(\boldsymbol{x})\right\|S_\Phi + [\boldsymbol{S}_1^T\boldsymbol{d}_f(\boldsymbol{x}) + d_\mu(\boldsymbol{X})u_0]S_\Phi \\ &\le -(1-\delta_0)k_S S_\Phi^2 - \{(1-\delta_0)k_S\Phi - \sqrt{m}\psi_f\}S_\Phi - \{(1-\delta_0)\sigma_v - \delta_0\}\|\boldsymbol{v}(t)\|S_\Phi - \{(1-\delta_0)\sigma_f - \delta_0\}\left\|\boldsymbol{f}^a(\boldsymbol{x})\right\|S_\Phi\end{aligned} \tag{39}$$

Let us now choose the design constants k_S, Φ, σ_v, σ_f so that the following inequalities are satisfied.

$$\delta_0 < 1 \tag{40a}$$

$$(1-\delta_0)k_S\Phi \ge \sqrt{m}\psi_f \tag{40b}$$

$$(1-\delta_0)\sigma_v \ge \delta_0 \tag{40c}$$

$$(1-\delta_0)\sigma_f \ge \delta_0 \tag{40d}$$

Then it follows that

$$\dot{V} \le -(1-\delta_0)k_S S_\Phi^2 < 0 \tag{41}$$

We now need to establish that such design constants exist to satisfy the inequalities (40a-d). In (40a)

$$\delta_0 = \frac{\psi_\mu}{\underline{b}-\psi_\mu} < 1 \tag{42}$$

Assume that

$$\underline{b} > 2\psi_\mu \tag{43}$$

then (40a) is satisfied. Given (43) we design k_S, Φ, σ_v, σ_f to satisfy

$$k_S\Phi \ge \frac{\sqrt{m}\psi_f}{1-\delta_0} \tag{44a}$$

$$\sigma_v \ge \frac{\delta_0}{1-\delta_0} \tag{44b}$$

$$\sigma_f \ge \frac{\delta_0}{1-\delta_0} \tag{44c}$$

Given that the above inequalities are satisfied we have

$$\dot{V} = 0 \text{, if } \|\boldsymbol{S}\| \le \Phi \tag{45a}$$

$$\dot{V} \le -(1-\delta_0)k_S V < 0 \text{, if } \|\boldsymbol{S}\| > \Phi \tag{45b}$$

The results (45a-b) hold under assumptions 1 and 5, i.e., for $\boldsymbol{x}\in\Omega$. We therefore need to establish that the proposed control law and initial conditions do not force $\boldsymbol{x}$ to get out of the set Ω at any point in time $t \ge 0$. Define the sets

$$\Omega_e := \{\boldsymbol{e} = \boldsymbol{x} - \boldsymbol{y}_d \mid \boldsymbol{x} \in \Omega, \boldsymbol{y}_d \in \Omega_d\} \tag{46a}$$

$$M := \{\boldsymbol{e} \mid V(t) \le V_0\} \tag{46b}$$

where $\Omega \supset \Omega_d$ and $V_0 > \Phi$ is chosen as the constant so that $M = \Omega_e$. Then for all $\boldsymbol{x}(0) \in \Omega_x$ where $\Omega_x \subset \Omega$, it follows from (46a) that $\boldsymbol{e}(0) \in \Omega_e$ which implies that $V(t)$ is bounded from above by V_0 for all $t \ge 0$, which in turn implies that $\boldsymbol{e}(t) \in \Omega_e$, $\forall t \ge 0$. $\boldsymbol{e}(t) \in \Omega_e$ implies $\boldsymbol{x}\in\Omega$, $\forall t \ge 0$ and therefore $\boldsymbol{x}$ cannot leave Ω at any time $t \ge 0$. From (31),(45b), it follows that S_Φ converges to zero exponentially fast which in turn implies that $\|\boldsymbol{S}\|$ converges to the residual set $R_S = \{\boldsymbol{S} \mid \|\boldsymbol{S}\| \le \Phi\}$ exponentially fast [1]. The boundedness of $\boldsymbol{u}$ and all signals in the closed loop follows from the boundedness of $\boldsymbol{e}$,$\boldsymbol{x}$,$\boldsymbol{S}$. Inside the residual set we also have $|S_i| \le \Phi$ which in turn implies that $|e_i(t)| \le \lambda_i^{-r_i+1}\Phi$, $i = 1,\cdots,m$ [4], [10]. ■

In the case where $f_i^a(\boldsymbol{x})$, i=1,2,…,m and $\mu^a(\boldsymbol{X})$ are unknown, the control law (26) cannot be used. In this case we follow the Certainty Equivalence (CE) principle and replace the unknown functions $f_i^a(\boldsymbol{x})$, $\mu^a(\boldsymbol{X})$ in (26) with their estimates. Let the estimates of the unknown functions $f_i^a(\boldsymbol{x})$, $\mu^a(\boldsymbol{X})$ at time t be formed as

$$\hat{f}_i^a(\boldsymbol{x},t) = \sum_{j=1}^{l_i} \hat{\theta}_{ij}(t)\zeta_{ij}(\boldsymbol{x}), \quad i=1,\ldots,m \tag{47a}$$

$$\hat{\mu}^a(\boldsymbol{X},t) = \sum_{j=1}^{l_0} \hat{\theta}_{0j}(t)\zeta_{0j}(\boldsymbol{X}) \tag{47b}$$

where $\hat{\theta}_{ij}(t)$ and $\hat{\theta}_{0j}(t)$ are the estimates of θ_{ij} and θ_{0j} at time t respectively to be generated by some adaptive law. The difference between the estimated and actual parameter values results in the estimation errors

$$\tilde{f}_i^a(\boldsymbol{x},t) = \sum_{j=1}^{l_i} \tilde{\theta}_{ij}(t)\zeta_{ij}(\boldsymbol{x}) \tag{48a}$$

$$\tilde{\mu}^a(\boldsymbol{X},t) = \sum_{j=1}^{l_0} \tilde{\theta}_{0j}(t)\zeta_{0j}(\boldsymbol{X}) \tag{48b}$$

where,

$$\tilde{\theta}_{ij}(t) = \hat{\theta}_{ij}(t) - \theta_{ij}, \quad i=0,1,\ldots,m \tag{49}$$

are the parameter errors.

Given the estimates $\hat{f}_i^a(\boldsymbol{x},t)$ and $\hat{\mu}^a(\boldsymbol{X},t)$ we can use the CE approach [1] to come up with an initial guess for the adaptive control law

$$\boldsymbol{u} = \frac{\boldsymbol{S}_1}{\hat{\mu}^a(\boldsymbol{X},t)}\left[-k_S\|\boldsymbol{S}\| - \boldsymbol{S}_1^T\hat{\boldsymbol{f}}^a(\boldsymbol{x},t) - \boldsymbol{S}_1^T\boldsymbol{v}(t) - \sigma_v\|\boldsymbol{v}(t)\| - \sigma_f\left\|\hat{\boldsymbol{f}}^a(\boldsymbol{x},t)\right\|\right] \tag{50}$$

where $\hat{\boldsymbol{f}}^a(\boldsymbol{x},t) := [\hat{f}_1^a(\boldsymbol{x},t) \quad \cdots \quad \hat{f}_m^a(\boldsymbol{x},t)]^T$, and design an adaptive law for generating the parameter estimates $\hat{\theta}_{ij}(t)$ and therefore $\hat{f}_i^a(\boldsymbol{x},t)$ and $\hat{\mu}^a(\boldsymbol{X},t)$ so that the overall system is stable and the tracking error converges to a small residual set with time. However, it is well known in adaptive control that the estimate $\hat{\mu}^a(\boldsymbol{X},t)$ cannot be guaranteed to be away from zero for any given time t. This implies that $\boldsymbol{u}$ cannot be guaranteed to be bounded uniformly with time. Therefore the CE control law (50) cannot be used to stabilize the closed loop system for the case where the estimated plant loses its controllability, i.e., $\hat{\mu}^a(\boldsymbol{X},t) \approx 0$ at some time t.

In the following section we modify the control law (50) to bypass this stabilizability problem [1] and guarantee stability and performance for the case $\boldsymbol{f}(\boldsymbol{x})$ and $\boldsymbol{B}(\boldsymbol{x})$ are unknown.

3. ROBUST ADAPTIVE CONTROL SCHEME

Let us consider the system (3) and control problem solved in section 2 for the case of known nonlinearities. In this section we assume that the nonlinear functions in (3) are unknown and design a control law to meet the control objective. In order to take care the case where the estimated plant becomes uncontrollable at some points in time, we modify the CE control law (50) as:

$$\boldsymbol{u} = \frac{\hat{\mu}^a(\boldsymbol{X},t)\boldsymbol{S}_1}{(\hat{\mu}^a(\boldsymbol{X},t))^2 + \delta_\mu}\left[-k_S\|\boldsymbol{S}\| - \boldsymbol{S}_1^T\hat{\boldsymbol{f}}^a(\boldsymbol{x},t) - \boldsymbol{S}_1^T\boldsymbol{v}(t) - \sigma_v\|\boldsymbol{v}(t)\| - \sigma_f\left\|\hat{\boldsymbol{f}}^a(\boldsymbol{x},t)\right\|\right] \tag{51}$$

where $\delta_\mu > 0$ is a small design constant. By design, the control law in (51) is free of singularities since $(\hat{\mu}^a(\boldsymbol{X},t))^2 + \delta_\mu \geq \delta_\mu > 0$, $\forall$ $\boldsymbol{X}$,t. Therefore, the proposed controller overcomes the difficulty encountered in many adaptive control laws where the identified model becomes uncontrollable at some points in time. It is also interesting to note that $\boldsymbol{u} \to 0$ with the same speed as $\hat{\mu}^a(\boldsymbol{X},t) \to 0$. Thus, when the estimate $\hat{\mu}^a(\boldsymbol{X},t)$ 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. This design is critical since the potential loss of controllability has been the main drawback of many nonlinear adaptive laws that are based on inverse dynamics. The nonlinear terms $\sigma_v\|\boldsymbol{v}(t)\|$, $\sigma_f\left\|\hat{\boldsymbol{f}}^a(\boldsymbol{x},t)\right\|$ are used to compensate for the effect of the approximation errors $\boldsymbol{d}_f(\boldsymbol{x})$, $d_\mu(\boldsymbol{X})$ and effect of the design constant δ_μ.

The control law (51) can be rewritten in the compact form

$$\boldsymbol{u} = \boldsymbol{S}_1 u_0 \tag{52}$$

where

$$u_0 = \frac{\hat{\mu}^a(\boldsymbol{X},t)}{(\hat{\mu}^a(\boldsymbol{X},t))^2 + \delta_\mu}\left[-k_S\|\boldsymbol{S}\| - \boldsymbol{S}_1^T\hat{\boldsymbol{f}}^a(\boldsymbol{x},t) - \boldsymbol{S}_1^T\boldsymbol{v}(t) - \sigma_v\|\boldsymbol{v}(t)\| - \sigma_f\left\|\hat{\boldsymbol{f}}^a(\boldsymbol{x},t)\right\|\right] \tag{53}$$

provides some insight into the action of the controller. The control vector $\boldsymbol{u}$ is expressed as a normalized directional vector $\boldsymbol{S}_1$ scaled by the control effort u_0. $\boldsymbol{S}_1$ represents the direction of the error metric vector $\boldsymbol{S}$. The control vector $\boldsymbol{u}$ can be viewed as apportioning the total control effort u_0 in different directions. The components, corresponding to large values in

$\boldsymbol{S}_1$, have relatively larger control energy than those components with smaller values. Intuitively, this suggests that the bigger control energy is directed to those $S_i(t)$ with higher values.

The adaptive laws for generating the estimates $\hat{\theta}_{ij}(t)$, i=0,1,…,m, j=1,2,…, l_i are as follows:

$$\dot{\hat{\theta}}_{ij}(t) = k_{f_i} S_i \frac{S_\Phi}{S_\Phi + \Phi} \zeta_{ij}(\boldsymbol{x}), \quad i = 1, \cdots, m \tag{54a}$$

$$\dot{\hat{\theta}}_{0j}(t) = k_\mu S_\Phi u_0 \zeta_{0j}(\boldsymbol{X}) + \rho(t) k_\mu \sigma_\mu S_\Phi \operatorname{sgn}(\mu(\boldsymbol{X}))(|u_0| + |u'|)\zeta_{0j}(\boldsymbol{X}) \tag{54b}$$

where

$$u' = \frac{1}{(\hat{\mu}^a(\boldsymbol{X},t))^2 + \delta_\mu}\left[-k_S\|\boldsymbol{S}\| - \boldsymbol{S}_1^T \hat{\boldsymbol{f}}^a(\boldsymbol{x},t) - \boldsymbol{S}_1^T \boldsymbol{v}(t) - \sigma_v\|\boldsymbol{v}(t)\| - \sigma_f\left\|\hat{\boldsymbol{f}}^a(\boldsymbol{x},t)\right\|\right] \tag{55}$$

$k_{f_i} > 0,\ i = 1, \cdots, m$ and $k_\mu > 0$ are adaptive gains chosen by the designer, u_0 is defined in (53), sgn(·) is the sign function (sgn(x)=1, if x≥0 and sgn(x)=−1, otherwise). Then sgn(μ)=1 if $\frac{1}{2}(\boldsymbol{B}(\boldsymbol{x}) + \boldsymbol{B}^T(\boldsymbol{x}))$ is positive definite and sgn(μ)=−1 if $\frac{1}{2}(\boldsymbol{B}(\boldsymbol{x}) + \boldsymbol{B}^T(\boldsymbol{x}))$ is negative definite. $\rho(t)$ is a switching function defined as:

$$\rho(t) = \begin{cases} 1, & \text{if } |\hat{\mu}^a| \le \underline{b} - \psi_\mu - \Delta \\ (\underline{b} - \psi_\mu - |\hat{\mu}^a|)/\Delta, & \text{if } (\underline{b} - \psi_\mu - \Delta) < |\hat{\mu}^a| < (\underline{b} - \psi_\mu) \\ 0, & \text{if } |\hat{\mu}^a| \ge \underline{b} - \psi_\mu \end{cases} \tag{56}$$

where $\Delta > 0$ is a design parameter used to avoid discontinuity in $\rho(t)$. A continuous switching function $\rho(t)$ as shown in Figure 1, instead of a discontinuous one, is used to guarantee that the resulting differential equation representing the closed loop system satisfies the conditions for existence and uniqueness of solutions [45]. The constant ψ_μ is defined in (25b) and represents the upper bound in the approximation of $\mu(\boldsymbol{X})$ with $\mu^a(\boldsymbol{X})$.

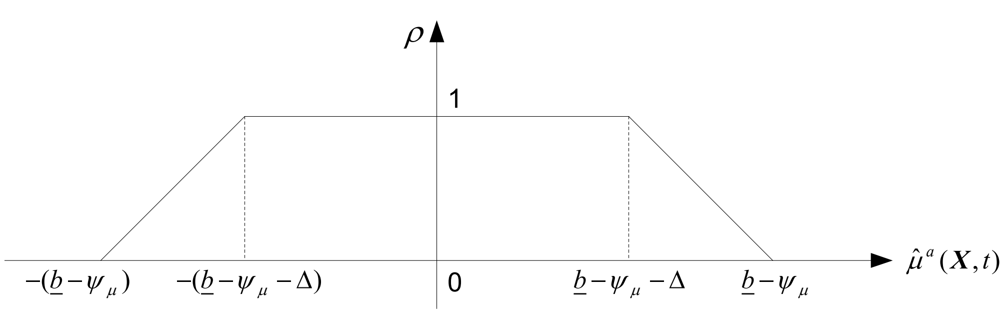

Fig. 1. Continuous Switching Function $\rho(t)$

The properties of the overall control law (51), (54a-b) are described by the following theorem.

***Theorem* 2**: Consider the system (3), the control law (51) and the adaptive laws (54a-b). Assume that assumptions 1-5 hold. Let ψ_f, ψ_μ are some positive constants such that (25a) is satisfied on a compact set $\Omega \subset \Re^r$ and (25b) is satisfied on the compact set $\Omega \times \Omega_d \subset \Re^{2r}$ for some $\Theta := [\Theta_0, \Theta_1, \cdots \Theta_m]^T \in \Re^l$. If the lower bound $\underline{b}$ of $\mu(\boldsymbol{X})$ satisfies the condition $\underline{b} > \sqrt{\delta_\mu} + 3\psi_\mu + \Delta$ for some arbitrary small positive constants δ_μ, Δ, then there exist positive constants $\delta_1 < 1$, $\delta_2 < 1$, $\delta_3 < 1$ such that if $k_S > \dfrac{\sqrt{m}\psi_f}{(1-\delta_1)\Phi}$, $\sigma_v \geq \dfrac{\delta_1}{1-\delta_1}$, $\sigma_f \geq \dfrac{\delta_1}{1-\delta_1}$, and $\sigma_\mu \geq \max(\delta_2, \delta_3)$, for all $\{\boldsymbol{x}(0) + \boldsymbol{y}_d(0), \widetilde{\Theta}(0)\} \in M$, where $M \subset \Re^{l+r}$, all signals in the closed-loop system (3) are bounded and the tracking errors $e_i(t)$ converge to the residual set $R_{e_i} = \{e_i \big| \, |e_i(t)| \leq \lambda_i^{-r_i+1}\Phi\}$, i=1,…,m, as $t \to \infty$. The size of the residue set depends only on the design constants λ_i, Φ and can be specified a priori by the designer.

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

$$V(t) = \frac{1}{2}S_\Phi^2 + \left\{ \frac{1}{2k_{f_1}} \sum_{j=1}^{l_1} \left(\widetilde{\theta}_{1j}(t)\right)^2 + \cdots + \frac{1}{2k_{f_m}} \sum_{j=1}^{l_m} \left(\widetilde{\theta}_{mj}(t)\right)^2 \right\} + \frac{1}{2k_\mu} \sum_{j=1}^{l_0} \left(\widetilde{\theta}_{0j}(t)\right)^2 \tag{57}$$

Using the adaptive laws (54a-b), we can establish that

$$\dot{V}(t) = 0, \text{ if } \|\boldsymbol{S}\| \leq \Phi \tag{58}$$

$$\dot{V}(t) = S_\Phi \dot{S}_\Phi + \left\{ \frac{1}{k_{f_1}} \sum_{j=1}^{l_1} \widetilde{\theta}_{1j} \dot{\hat{\theta}}_{1j}(t) + \cdots + \frac{1}{k_{f_m}} \sum_{j=1}^{l_m} \widetilde{\theta}_{mj} \dot{\hat{\theta}}_{mj}(t) \right\} + \frac{1}{k_\mu} \sum_{j=1}^{l_0} \widetilde{\theta}_{0j} \dot{\hat{\theta}}_{0j}(t), \text{ if } \|\boldsymbol{S}\| > \Phi \tag{59}$$

In the following proof, we only consider the region $\|\boldsymbol{S}\| > \Phi$. Rewrite u_0 in (53) as:

$$u_0 = \frac{\hat{\mu}^a(\boldsymbol{X}, t)}{(\hat{\mu}^a(\boldsymbol{X}, t))^2 + \delta_\mu} \overline{u} \tag{60}$$

$$\overline{u} = -k_S \|\boldsymbol{S}\| - \boldsymbol{S}_1^T \hat{\boldsymbol{f}}^a(\boldsymbol{x}, t) - \boldsymbol{S}_1^T \boldsymbol{v}(t) - \sigma_v \|\boldsymbol{v}(t)\| - \sigma_f \left\|\hat{\boldsymbol{f}}^a(\boldsymbol{x}, t)\right\| \tag{61}$$

Using Lemma 2 and (60)-(61), we obtain

$$
\begin{aligned}
\dot{S}_\Phi &= \boldsymbol{S}_1^T \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{S}_1^T \boldsymbol{v}(t) + \mu(\boldsymbol{X})u_0 \\
&= \boldsymbol{S}_1^T \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{S}_1^T \boldsymbol{v}(t) + \frac{(\hat{\mu}^a(\boldsymbol{X},t))^2}{(\hat{\mu}^a(\boldsymbol{X},t))^2 + \delta_\mu}\overline{u} + \{\mu(\boldsymbol{X}) - \hat{\mu}^a(\boldsymbol{X},t)\}u_0 \\
&= \boldsymbol{S}_1^T \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{S}_1^T \boldsymbol{v}(t) + \overline{u} - \frac{\delta_\mu}{(\hat{\mu}^a(\boldsymbol{X},t))^2 + \delta_\mu}\overline{u} + \{\mu(\boldsymbol{X}) - \hat{\mu}^a(\boldsymbol{X},t)\}u_0 \\
&= -k_S\|\boldsymbol{S}\| - \sigma_v\|\boldsymbol{v}(t)\| - \sigma_f\left\|\hat{\boldsymbol{f}}^a(\boldsymbol{x},t)\right\| + \boldsymbol{S}_1^T\{\boldsymbol{f}(\boldsymbol{x}) - \hat{\boldsymbol{f}}^a(\boldsymbol{x},t)\} + \{\mu(\boldsymbol{X}) - \hat{\mu}^a(\boldsymbol{X},t)\}u_0 - \delta_\mu u'
\end{aligned}
\tag{62}
$$

Using the identities,

$$
\begin{aligned}
\boldsymbol{f}(\boldsymbol{x}) - \hat{\boldsymbol{f}}^a(\boldsymbol{x},t) &= (\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}^a(\boldsymbol{x})) - (\hat{\boldsymbol{f}}^a(\boldsymbol{x},t) - \boldsymbol{f}^a(\boldsymbol{x})) \\
&= \boldsymbol{d}_f(\boldsymbol{x}) - \tilde{\boldsymbol{f}}^a(\boldsymbol{x},t)
\end{aligned}
\tag{63a}
$$

$$
\begin{aligned}
\mu(\boldsymbol{X}) - \hat{\mu}^a(\boldsymbol{X},t) &= [\mu(\boldsymbol{X}) - \mu^a(\boldsymbol{X})] - [\hat{\mu}^a(\boldsymbol{X},t) - \mu^a(\boldsymbol{X})] \\
&= d_\mu(\boldsymbol{X}) - \tilde{\mu}^a(\boldsymbol{X},t)
\end{aligned}
\tag{63b}
$$

where $\tilde{\boldsymbol{f}}^a(\boldsymbol{x},t) := [\tilde{f}_1^a(\boldsymbol{x},t) \quad \cdots \quad \tilde{f}_m^a(\boldsymbol{x},t)]^T$ and $\|\boldsymbol{S}\| = S_\Phi + \Phi$ it follows that

$$
\dot{S}_\Phi = -k_S S_\Phi - k_S\Phi - \sigma_v\|\boldsymbol{v}(t)\| - \sigma_f\left\|\hat{\boldsymbol{f}}^a(\boldsymbol{x},t)\right\| - \boldsymbol{S}_1^T\tilde{\boldsymbol{f}}^a(\boldsymbol{x},t) - \tilde{\mu}^a(\boldsymbol{X},t)u_0 + \boldsymbol{S}_1^T\boldsymbol{d}_f(\boldsymbol{x}) + \{d_\mu(\boldsymbol{X})u_0 - \delta_\mu u'\} \tag{64}
$$

The term $\{d_\mu(\boldsymbol{X})u_0 - \delta_\mu u'\}$ in (64) is a modeling error term representing the effect of the design parameter δ_μ and the approximation error $d_\mu(\boldsymbol{X})$. The price paid to avoid the singularity in control law (51) is that the design parameter δ_μ appears as a disturbance in the closed loop system. The modeling error term $\{d_\mu(\boldsymbol{X})u_0 - \delta_\mu u'\}$ will become dominant when the estimate $\hat{\mu}^a(\boldsymbol{X},t) \to 0$.

Define:

$$
\delta_1 := \frac{\psi_\mu}{\underline{b} - \psi_\mu - \Delta} + \frac{\delta_\mu}{(\underline{b} - \psi_\mu - \Delta)^2 + \delta_\mu} \tag{65a}
$$

$$
\delta_2 := \frac{2\delta_\mu}{\underline{b} - \psi_\mu} \tag{65b}
$$

$$
\delta_3 := \frac{\psi_\mu}{\underline{b} - \psi_\mu} \tag{65c}
$$

As shown in Appendix B, the absolute value of the modeling error in $\dot{S}_\Phi$ can be expressed as:

$$
\left|d_\mu(\boldsymbol{X})u_0 - \delta_\mu u'\right| \le \delta_1|\overline{u}| + \rho(t)\delta_2\left|\tilde{\mu}^a(\boldsymbol{X},t)\right||u'| + \rho(t)\delta_3\left|\tilde{\mu}^a(\boldsymbol{X},t)\right||u_0| \tag{66}
$$

Using

$$
|\overline{u}| \le k_S S_\Phi + k_S\Phi + (1+\sigma_v)\|\boldsymbol{v}(t)\| + (1+\sigma_f)\left\|\hat{\boldsymbol{f}}^a(\boldsymbol{x},t)\right\| \tag{67}
$$

(66) can be rewritten as

$$
\begin{aligned}
\left|d_\mu(\boldsymbol{X})u_0-\delta_\mu u'\right| &\le \delta_1 k_S S_\Phi+\delta_1 k_S\Phi+\delta_1(1+\sigma_v)\|\boldsymbol{v}(t)\|+\delta_1(1+\sigma_f)\left\|\hat{\boldsymbol{f}}^a(\boldsymbol{x},t)\right\| \\
&\quad+\rho(t)\delta_2\left|\tilde{\mu}^a(\boldsymbol{X},t)\right|\left|u'\right|+\rho(t)\delta_3\left|\tilde{\mu}^a(\boldsymbol{X},t)\right|\left|u_0\right|
\end{aligned}
\tag{68}
$$

Using (64), the first term in $\dot{V}$ is,

$$
\begin{aligned}
S_\Phi\dot{S}_\Phi &= -k_S S_\Phi^2-k_S\Phi S_\Phi-\sigma_v\|\boldsymbol{v}(t)\|S_\Phi-\sigma_f\left\|\hat{\boldsymbol{f}}^a(\boldsymbol{x},t)\right\|S_\Phi-\boldsymbol{S}_1^T\tilde{\boldsymbol{f}}^a(\boldsymbol{x},t)S_\Phi \\
&\quad-\tilde{\mu}^a(\boldsymbol{X},t)u_0S_\Phi+\boldsymbol{S}_1^T\boldsymbol{d}_f(\boldsymbol{x})S_\Phi+\{d_\mu(\boldsymbol{X})u_0-\delta_\mu u'\}S_\Phi
\end{aligned}
\tag{69}
$$

Substituting (68) into (69), we obtain:

$$
\begin{aligned}
S_\Phi\dot{S}_\Phi &\le -(1-\delta_1)k_S S_\Phi^2-\{(1-\delta_1)k_S\Phi-\sqrt{m}\psi_f\}S_\Phi-\{\sigma_v-\delta_1(1+\sigma_v)\}\|\boldsymbol{v}(t)\|S_\Phi-\{\sigma_f-\delta_1(1+\sigma_f)\}\left\|\hat{\boldsymbol{f}}^a(\boldsymbol{x},t)\right\|S_\Phi \\
&\quad-\boldsymbol{S}_1^T\tilde{\boldsymbol{f}}^a(\boldsymbol{x},t)S_\Phi-\tilde{\mu}^a(\boldsymbol{X},t)u_0S_\Phi+\rho(t)\delta_2\left|\tilde{\mu}^a(\boldsymbol{X},t)\right|\left|u'\right|S_\Phi+\rho(t)\delta_3\left|\tilde{\mu}^a(\boldsymbol{X},t)\right|\left|u_0\right|S_\Phi
\end{aligned}
\tag{70}
$$

Let us now consider the second term of $\dot{V}$ in (59). Using the adaptive law (54a) we have

$$
\begin{aligned}
\frac{1}{k_{f_1}}\sum_{j=1}^{l_{f_1}}\tilde{\theta}_{1j}\dot{\hat{\theta}}_{1j}+\cdots+\frac{1}{k_{f_m}}\sum_{j=1}^{l_{fm}}\tilde{\theta}_{mj}\dot{\hat{\theta}}_{mj} &= S_1\tilde{f}_1^a(\boldsymbol{x},t)\frac{S_\Phi}{S_\Phi+\Phi}+\cdots+S_m\tilde{f}_m^a(\boldsymbol{x},t)\frac{S_\Phi}{S_\Phi+\Phi} \\
&= \boldsymbol{S}^T\tilde{\boldsymbol{f}}^a(\boldsymbol{x},t)\frac{S_\Phi}{S_\Phi+\Phi} \\
&= \boldsymbol{S}_1^T\tilde{\boldsymbol{f}}^a(\boldsymbol{x},t)S_\Phi
\end{aligned}
\tag{71}
$$

Finally, in view of the adaptive law (54b), the last term of $\dot{V}$ in (59) can be written as:

$$
\begin{aligned}
\frac{1}{k_\mu}\sum_{j=1}^{l_0}\tilde{\theta}_{0j}\dot{\hat{\theta}}_{0j} &= \frac{1}{k_\mu}\sum_{j=1}^{l_0}\tilde{\theta}_{0j}\{k_\mu S_\Phi u_0\zeta_{0j}(\boldsymbol{X})+\rho(t)k_\mu\sigma_\mu S_\Phi\,\mathrm{sgn}(\mu(\boldsymbol{X}))(\left|u_0\right|+\left|u'\right|)\zeta_{0j}(\boldsymbol{X})\} \\
&= \tilde{\mu}^a(\boldsymbol{X},t)S_\Phi u_0-\rho(t)\sigma_\mu\left|\tilde{\mu}^a(\boldsymbol{X},t)\right|\left|u_0\right|S_\Phi-\rho(t)\sigma_\mu\left|\tilde{\mu}^a(\boldsymbol{X},t)\right|\left|u'\right|S_\Phi
\end{aligned}
\tag{72}
$$

Now, $\tilde{\mu}^a(\boldsymbol{X},t)=\hat{\mu}^a(\boldsymbol{X},t)-\mu^a(\boldsymbol{X})=\{\hat{\mu}^a(\boldsymbol{X},t)+d_\mu(\boldsymbol{X})\}-\mu(\boldsymbol{X})$. This together with the fact that $\rho(t)\neq 0$ only for $\left|\hat{\mu}^a(\boldsymbol{X},t)\right|<\underline{b}-\psi_\mu$ and $\left|\hat{\mu}^a(\boldsymbol{X},t)+d_\mu(\boldsymbol{X})\right|\le\left|\hat{\mu}^a(\boldsymbol{X},t)\right|+\left|d_\mu(\boldsymbol{X})\right|\le\underline{b}$ implies that for $\rho(t)\neq 0$, the sign of $\rho(t)\tilde{\mu}^a(\boldsymbol{X},t)$ is always the opposite sign of $\mu(\boldsymbol{X})$, $\forall\ t\ge 0$.

Combining (70), (71), and (72), $\dot{V}$ can be expressed as:

$$
\begin{aligned}
\dot{V} &\le -(1-\delta_1)k_S S_\Phi^2-\{(1-\delta_1)k_S\Phi-\sqrt{m}\psi_f\}S_\Phi-\{(1-\delta_1)\sigma_v-\delta_1\}\|\boldsymbol{v}(t)\|S_\Phi \\
&\quad-\{(1-\delta_1)\sigma_f-\delta_1\}\left\|\hat{\boldsymbol{f}}^a(\boldsymbol{x},t)\right\|S_\Phi-\rho(t)(\sigma_\mu-\delta_2)\left|\tilde{\mu}^a(\boldsymbol{X},t)\right|\left|u'\right|S_\Phi-\rho(t)(\sigma_\mu-\delta_3)\left|\tilde{\mu}^a(\boldsymbol{X},t)\right|\left|u_0\right|S_\Phi
\end{aligned}
\tag{73}
$$

Let us now choose the design constants k_S, Φ, σ_v, σ_f, σ_μ so that the following inequalities are satisfied.

$$
\delta_1<1 \tag{74a}
$$

$$(1-\delta_1)k_S\Phi \ge \sqrt{m}\psi_f \tag{74b}$$

$$(1-\delta_1)\sigma_v \ge \delta_1 \tag{74c}$$

$$(1-\delta_1)\sigma_f \ge \delta_1 \tag{74d}$$

$$\delta_\mu \ge \delta_2 \tag{74e}$$

$$\delta_\mu \ge \delta_3 \tag{74f}$$

It follows that

$$\dot{V} \le -(1-\delta_1)k_S S_\Phi^2 < 0 \tag{75}$$

We now need to establish that such design constants exist to satisfy the inequality (74a-f). In (74a)

$$\delta_1 = \frac{\psi_\mu}{\underline{b}-\psi_\mu-\Delta} + \frac{\delta_\mu}{(\underline{b}-\psi_\mu-\Delta)^2+\delta_\mu} < 1 \tag{76}$$

Assume that

$$\underline{b} > \sqrt{\delta_\mu} + 3\psi_\mu + \Delta \tag{77}$$

then (74a) is satisfied. This inequality depends on the value of $\underline{b}$ which is a characteristic of the nonlinear system. It suggests that the design constant δ_μ has to be chosen relatively small and a sufficient number of nodes in $\mu^a(\boldsymbol{X})$ have to be used in order to make ψ_μ small.

Given (77) we design k_S, Φ, σ_v, σ_f, σ_μ to satisfy

$$k_S\Phi \ge \frac{\sqrt{m}\psi_f}{1-\delta_1} \tag{78a}$$

$$\sigma_v \ge \frac{\delta_1}{1-\delta_1} \tag{78b}$$

$$\sigma_f \ge \frac{\delta_1}{1-\delta_1} \tag{78c}$$

$$\sigma_\mu \ge \max\{\delta_2, \delta_3\} \tag{78d}$$

Given that the above inequalities (78a-d) are satisfied we have

$$\dot{V} = 0 \text{, if } \|\boldsymbol{S}\| \le \Phi \tag{79a}$$

$$\dot{V} \le -(1-\delta_1)k_S S_\Phi^2 < 0 \text{, if } \|\boldsymbol{S}\| > \Phi \tag{79b}$$

The results (79a-b) hold under the assumption $\boldsymbol{x}\in\Omega$ for all $t \geq 0$. We need to establish that for $\boldsymbol{x}(0) \in \Omega_x$ and $\widetilde{\Theta}(0) \in \Omega_\Theta$, where Ω_x, Ω_Θ, need to be specified, we have $\boldsymbol{x}\in\Omega$. We start by considering the sets

$$\Omega_e := \{\boldsymbol{e} = \boldsymbol{x} - \boldsymbol{y}_d \mid \boldsymbol{x} \in \Omega, \boldsymbol{y}_d \in \Omega_d\} \tag{80a}$$

$$M = \{\boldsymbol{e}, \widetilde{\Theta} \mid V(t) \leq V_0\} \tag{80b}$$

where $\Omega \supset \Omega_d$ and $V_0 > \Phi$ is chosen as the largest constant for which $M \subset \Omega_e \times \Omega_\Theta$. Then for all $\{\boldsymbol{x}(0) - \boldsymbol{y}_d(0), \widetilde{\Theta}(0)\} \in M$, it follows from (57) and (79a-b) that $V(t)$ is bounded from above by V_0 for all $t \geq 0$, which implies that $\boldsymbol{e}(t) \in \Omega_e$, $\forall t \geq 0$. From (80a) it follows $\boldsymbol{x}(t)\in\Omega$, $t \geq 0$. Therefore $\boldsymbol{x}$ cannot leave Ω at any time $t \geq 0$. It is noted that by choosing the adaptive gains k_{f_i}, k_μ large Ω_Θ can be made large.

The properties of $V(t)$ together with $\dot{V} \leq 0$ imply that $V(t)$ and therefore $S_\Phi, \widetilde{\theta}_{ij}(t)$ are bounded for all $t \geq 0$, i.e, $S_\Phi, \widetilde{\theta}_{ij}(t) \in L_\infty$. This in turn implies that $\boldsymbol{x}$, $\boldsymbol{u}$ are bounded and $V(t)$ has a limit, i.e. $\lim_{t\to\infty} V(t) = V_\infty$. Using the fact that $S_\Phi = 0$ for $\|\boldsymbol{S}\| \leq \Phi$ and (55b), we have $\lim_{t\to\infty} \int_0^t S_\Phi^2(\tau)d\tau = \int_0^\infty S_\Phi^2(t)dt \leq \dfrac{V(0) - V_\infty}{(1-\delta_1)k_S} < \infty$ which implies that $S_\Phi \in L_2$. From $S_\Phi, \widetilde{\theta}_{ij}(t) \in L_\infty$, it follows that all signals are bounded which implies that $\dot{S}_\Phi \in L_\infty$. From $\dot{S}_\Phi \in L_\infty$ and $S_\Phi \in L_2$ we have $S_\Phi \to 0$ as $t \to \infty$ which implies that $\|\boldsymbol{S}\|$ converges to the residual set $R_S = \{\boldsymbol{S} \mid \|\boldsymbol{S}\| \leq \Phi\}$ [1]. Inside the residual set we also have $|S_i| \leq \Phi$ which in turn implies that $|e_i(t)| \leq \lambda_i^{-r_i+1}\Phi$, $i = 1, \cdots, m$ [4], [10]. ■

<u>Design Parameter Procedure</u>

The design parameters can be chosen to guarantee that the tracking error is within a desired prespecified bound at steady state by using the following procedure provided the lower bound $\underline{b}$ related to the controllability of the plant, and ψ_μ, ψ_f, the upper bounds for the approximation errors are known *a priori*.

(1) Using ψ_μ, the upper bound of the approximation error $d_\mu(\boldsymbol{X})$, check if the lower bound of $\mu(\boldsymbol{X})$ satisfies $\underline{b} > 3\psi_\mu$. If so, choose the design parameters δ_μ, Δ such that $\sqrt{\delta_\mu} + \Delta + 3\psi_\mu < \underline{b}$. If not then the number of nodes l_0 of the neural network for $\mu(\boldsymbol{X})$ has to be increased in order to obtain a smaller approximation error bound ψ_μ.

(2) 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 bound.

(3) Calculate $\delta_1, \delta_2, \delta_3$ using (65a-c) and the knowledge of $\underline{b}, \psi_\mu, \delta_\mu, \Delta$, i.e.,

$$\delta_1 = \frac{\psi_\mu}{\underline{b} - \psi_\mu - \Delta} + \frac{\delta_\mu}{(\underline{b} - \psi_\mu - \Delta)^2 + \delta_\mu}, \quad \delta_2 = \frac{2\delta_\mu}{\underline{b} - \psi_\mu}, \quad \delta_3 = \frac{\psi_\mu}{\underline{b} - \psi_\mu}$$

(4) Choose σ_μ in the adaptive law (54b) such that $\sigma_\mu \geq \max(\delta_2, \delta_3)$.

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

Remark 2: It is worth noting that the ratio, $\psi_\mu / \underline{b} < 1/3$, gives an upper bound for the approximation error $d_\mu(\boldsymbol{X})$ that can be tolerated by the closed loop system. Larger $\underline{b}$ implies that the closed loop system can tolerate larger approximation error $d_\mu(\boldsymbol{X})$. On the other hand, small $\underline{b}$ requires more accurate approximation of $\mu(\boldsymbol{X})$ by $\mu^a(\boldsymbol{X})$, which normally implies more nodes in the neural network for $\mu^a(\boldsymbol{X})$.

Remark 3: The tracking errors at steady state are guaranteed to converge inside the residual set $|e_i(t)| \leq \lambda_i^{-r_i+1}\Phi$, $i = 1, \cdots, m$, whose size depends on the design parameter Φ. The smaller the design parameter Φ is, the smaller the tracking error is guaranteed to be at steady state. The design parameter Φ is inversely proportional to the gain k_S. For a given Φ, the gain k_S depends on the approximation error ψ_f, δ_1. From (78a-d), it is easy to check that δ_1 is a critical parameter used to determine the design parameters k_S, σ_f, σ_v. Let us rewrite δ_1 as $\delta_1 = \frac{\psi_\mu / \underline{b}}{1 - \psi_\mu / \underline{b} - \Delta / \underline{b}} + \frac{(\sqrt{\delta_\mu} / \underline{b})^2}{(1 - \psi_\mu / \underline{b} - \Delta / \underline{b})^2 + (\sqrt{\delta_\mu} / \underline{b})^2}$, which is a function of the 3 ratios $\psi_\mu / \underline{b}$, $\sqrt{\delta_\mu} / \underline{b}$, $\Delta / \underline{b}$. We require $\sigma_\mu \geq \max\{\delta_2, \delta_3\}$. Using the expressions $\delta_2 = \frac{2(\delta_\mu / \underline{b})}{1 - (\psi_\mu / \underline{b})}$, $\delta_3 = \frac{(\psi_\mu / \underline{b})}{1 - (\psi_\mu / \underline{b})}$, it follows that δ_2, δ_3, depend on the ratios $\delta_\mu / \underline{b}$, $\psi_\mu / \underline{b}$. From $\psi_\mu / \underline{b} < 1/3$, we have $\delta_2 < 3\delta_\mu / \underline{b}$, $\delta_3 < 1/2$. Smaller ratios $\psi_\mu / \underline{b}$, $\delta_\mu / \underline{b}$, $\Delta / \underline{b}$ imply smaller values for the design parameters k_S, σ_f, σ_v, σ_μ. However, a smaller value for $\psi_\mu / \underline{b}$ requires a better approximation of the unknown function $\mu(\boldsymbol{X})$, which implies a higher order neural network. A smaller value for $\delta_\mu / \underline{b}$ may imply a larger control input when $\hat{\mu}^a(\boldsymbol{X}, t)$ becomes smaller. A smaller value for $\Delta / \underline{b}$ may imply a faster switching action. Therefore, there is a tradeoff between the design parameters, k_S, σ_f, σ_v, σ_μ, and the ratios, $\psi_\mu / \underline{b}$, $\delta_\mu / \underline{b}$,

$\Delta/\underline{b}$. 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 4: The σ-modification term $\rho(t)k_\mu\sigma_\mu S_\Phi \,\mathrm{sgn}(\mu(\boldsymbol{X}))(|u'|+|u|)\zeta_{0j}(\boldsymbol{X})$ has been incorporated in the adaptive law (54b) to ensure stability and robustness. In the expression for $\dot{S}_\Phi$, the modeling error term $\{d_\mu(\boldsymbol{X})u_0-\delta_\mu u'\}$ appears because of the design parameter, δ_μ, and the approximation error $d_\mu(\boldsymbol{X})$. This term may become dominant in the case where $\hat{\mu}^a$ approaches zero. This implies the signals in the closed loop system may become unbounded when $\hat{\mu}^a\to 0$. The σ-modification is activated to guarantee stability when the estimate $\hat{\mu}^a$ becomes smaller than the lower bound of $\mu^a(\boldsymbol{X})$. In fact, it guarantees the convergence of the tracking error even if $\hat{\mu}^a$ approaches zero since it appears as a negative term in the derivative of the Lyapunov-like function to cancel the effect of the modeling error $\{d_\mu(\boldsymbol{X})u_0-\delta_\mu u'\}$. Therefore this special σ-modification ensures robustness whereas the classical σ-modification is used to avoid the estimate of the parameters to drift to infinity [1]. From the adaptive learning point of view, when the σ-modification is activated, $\hat{\theta}_{0j}(t)$, j=1,2,…, l_0 , will increase along the direction of the actual $\mu(\boldsymbol{X})$. Thus the special σ-modification could be viewed as a soft projection algorithm used to prevent $\hat{\mu}^a(\boldsymbol{X},t)$ from going to zero.

Remark 5: For the case where each element of $\boldsymbol{B}(\boldsymbol{x})$ is in the linear-in-the-parameters form, i.e.,

$$b_{ij}(\boldsymbol{x})=\Theta_{ij}^T\boldsymbol{\zeta}_{ij}(\boldsymbol{x})+d_{ij}(\boldsymbol{x}),\ \ i,j=1,2,\ldots,m \tag{81}$$

where $\Theta_{ij}\in\Re^{l_{ij}}$ are unknown parameter vectors, $\boldsymbol{\zeta}_{ij}(\boldsymbol{x})\in\Re^{l_{ij}}$ are known function vectors, and $d_{ij}(\boldsymbol{x})\in\Re$ are bounded modeling errors, the results of Theorem 2 are global in the sense $\Omega_1\equiv\Re^r$. In this case,

$$\mu(\boldsymbol{X})=\Theta_0^T\boldsymbol{\zeta}(\boldsymbol{X})+d_\mu(\boldsymbol{X}) \tag{82}$$

where $\Theta_0=[\Theta_{11}^T,\cdots\Theta_{1m}^T,\cdots,\Theta_{m1}^T,\cdots,\Theta_{mm}^T]^T\in\Re^{l_0}$, $\boldsymbol{\zeta}(\boldsymbol{X})=[\boldsymbol{\zeta}_{11}^T S_1^2,\cdots,\boldsymbol{\zeta}_{1m}^T S_1S_m,\cdots,\boldsymbol{\zeta}_{m1}^T S_mS_1,\cdots,\boldsymbol{\zeta}_{mm}^T S_m^2]^T/\boldsymbol{S}_1^2\in\Re^{l_0}$, $l_0=\sum_{i=1}^m\sum_{j=1}^m l_{ij}$, $d_\mu(\boldsymbol{X})=\sum_{i=1}^m\sum_{j=1}^m\{d_{ij}(\boldsymbol{x})S_iS_j/\boldsymbol{S}_1^2\}$, $\boldsymbol{S}_1$ is defined in (27).

Remark 6: The class of systems described by (3) can be derived from a more general class of nonlinear systems described by:

$$\dot{\boldsymbol{x}} = \boldsymbol{p}(\boldsymbol{x}) + \sum_{k=1}^{m} \boldsymbol{g}_k(\boldsymbol{x}) u_k$$
$$y_i = h_i(\boldsymbol{x}) \tag{83}$$

where $\boldsymbol{x} = [x_1, \cdots, x_n]^T \in \Re^n$ and $\boldsymbol{p} \in \Re^n$, $\boldsymbol{g}_i \in \Re^n$, $h_i \in \Re$, i=1,2,…,m, are smooth functions. If the above system is feedback linearizable, it can be reduced to the system (3) as described in [3]. Let:

$$\boldsymbol{f}(\boldsymbol{x}) = \left[L_{\boldsymbol{p}}^{r_1}(h_1) \quad \cdots \quad L_{\boldsymbol{p}}^{r_m}(h_m) \right]^T \tag{84a}$$

$$\boldsymbol{B}(\boldsymbol{x}) = \begin{bmatrix} L_{\boldsymbol{g}_1}\left(L_{\boldsymbol{p}}^{r_1-1}(h_1)\right) & \cdots & L_{\boldsymbol{g}_m}\left(L_{\boldsymbol{p}}^{r_1-1}(h_1)\right) \\ \vdots & \ddots & \vdots \\ L_{\boldsymbol{g}_1}\left(L_{\boldsymbol{p}}^{r_m-1}(h_m)\right) & \cdots & L_{\boldsymbol{g}_m}\left(L_{\boldsymbol{p}}^{r_m-1}(h_m)\right) \end{bmatrix} \tag{84b}$$

where, the Lie Derivative expressions $L_{\boldsymbol{p}}$ and $L_{\boldsymbol{p}}^k$ are defined as:

$$L_{\boldsymbol{p}}(h) := < dh, \boldsymbol{p} > = \frac{\partial h}{\partial x_1} p_1(x) + \cdots + \frac{\partial h}{\partial x_n} p_n(x) \tag{85a}$$

$$L_{\boldsymbol{p}}^k(h) := L_{\boldsymbol{p}}\left(L_{\boldsymbol{p}}^{k-1}(h)\right) = < dL_{\boldsymbol{p}}^{k-1}(h), \boldsymbol{p} > \tag{85b}$$

Here, r_i is the equivalent linearizability index for output y_i, i.e., one needs to differentiate the output, y_i, r_i times until one of the control inputs is different from zero. $r = \sum_{i=1}^{m} r_i$ indicates the relative degree of the nonlinear system. Here we assume r=n such that there are no internal dynamics in the linearized plant.

4. SIMULATION RESULTS

We demonstrate the performance of the proposed adaptive control system using a dynamical model of a planar, two-link, articulated robotic manipulator [4],[42]. The dynamics of this robotic system are nonlinear with strong coupling between the two degrees of freedom. The equations of motion in terms of the generalized coordinates q_1 and q_2, representing the angular positions of joints 1 and 2 and applied torques τ_1 and τ_2 at these joints is given by:

$$\begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \begin{bmatrix} \ddot{q}_1 \\ \ddot{q}_2 \end{bmatrix} + \begin{bmatrix} -h\dot{q}_2 & -h(\dot{q}_1 + \dot{q}_2) \\ h\dot{q}_1 & 0 \end{bmatrix} \begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \end{bmatrix} = \begin{bmatrix} \tau_1 \\ \tau_2 \end{bmatrix} \tag{86}$$

where,

$$H_{11} = a_1 + 2a_3\cos(q_2) + 2a_4\sin(q_2)\,;\; H_{12} = H_{21} = a_2 + a_3\cos(q_2) + a_4\sin(q_2)\,;\; H_{22} = a_2\,;\; h = a_3\sin(q_2) - a_4\cos(q_2)$$

with

$$a_1 = I_1 + m_1 l_{c1}^2 + I_e + m_e l_{ce}^2 + m_e l_1^2\,;\; a_2 = I_e + m_e l_{ce}^2\,;\; a_3 = m_e l_1 l_{ce}\cos(\delta_e)\,;\; a_4 = m_e l_1 l_{ce}\sin(\delta_e)$$

The numerical values used for simulation purposes are:

$$m_1 = 1 \,,\ l_1 = 1 \,,\ m_e = 2 \,,\ \delta_e = 30^\circ \,,\ I_1 = 0.12 \,,\ l_{c1} = 0.5 \,,\ l_e = 0.25 \,,\ l_{ce} = 0.6$$

The robot initially is at rest at the position $q_1 = 0, q_2 = 0$. It is desired to determine control inputs $\tau_1(t)$ and $\tau_2(t)$ such that q_1 and q_2 follow a desired trajectory defined by:

$$q_{d_1}(t) = 30^\circ \cos(2\pi t)\,;\ q_{d_2}(t) = 45^\circ \cos(2\pi t) \tag{87}$$

with the tracking performance defined by:

$$|e_1| = \left|q_1 - q_{d_1}\right| \le 1.0^\circ\,;\ |e_2| = \left|q_2 - q_{d_2}\right| \le 1.0^\circ \tag{88}$$

Since the inertia matrix $\boldsymbol{H}(\boldsymbol{q})$ is positive definite, (86) can be written as:

$$\begin{bmatrix} \ddot{q}_1 \\ \ddot{q}_2 \end{bmatrix} = -\begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix}^{-1} \begin{bmatrix} -h\dot{q}_2 & -h(\dot{q}_1 + \dot{q}_2) \\ h\dot{q}_1 & 0 \end{bmatrix} \begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \end{bmatrix} + \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix}^{-1} \begin{bmatrix} \tau_1 \\ \tau_2 \end{bmatrix} \tag{89}$$

The above equation is in the general nonlinear form of (3) with $\boldsymbol{B}(\boldsymbol{x}) = \boldsymbol{H}^{-1}$. Because the workspace is a closed set, it is easy to show that both $\boldsymbol{H}$ and $\boldsymbol{H}^{-1}$ are uniformly positive definite for all positions $\boldsymbol{q}$ in the robot's workspace. In this simulation, the lower bound of the minimum singular value of $\boldsymbol{H}^{-1}$ is $\underline{b} = 0.152$. Assuming a completely unknown nonlinear plant, one layer radial basis neural network with a basis function $\zeta_i(\boldsymbol{x}) = \exp[-\pi\sigma^2 (\boldsymbol{x} - \xi_i)^T (\boldsymbol{x} - \xi_i)]$, where σ, ξ_i are design parameters, was used to approximate the unknown functions $\boldsymbol{f}$ and μ. The upper bounds for the approximation errors are estimated using offline training as $\left|f_i - f_i^a\right| \le 1.0$, i=1,2 and $\left|\mu - \mu^a\right| \le 0.01$, i.e., $\psi_f = 1.0$ and $\psi_\mu = 0.01$. Following the design procedure, the condition $(\underline{b} = 0.152) > (3\psi_\mu = 0.03)$ is satisfied. Then the values of δ_μ, Δ are chosen to be 0.002, 0.002 respectively such that $(\sqrt{\delta_\mu} + \Delta + 3\psi_\mu = 0.077) < (\underline{b} = 0.152)$. Taking $\Phi = 0.3$, $\lambda_1 = 20$, $\lambda_2 = 20$, the desired bounds for the steady state tracking errors are satisfied, i.e., $|e_1| \le \Phi / \lambda_1 = 0.9^\circ < 1^\circ$ and $|e_2| \le \Phi / \lambda_2 = 0.9^\circ < 1^\circ$. Using (65a-c), we calculate $\delta_1 = 0.12$, $\delta_2 = 0.141$, $\delta_3 = 0.141$. Based on conditions (78a-d), we choose $k_S = 20$, $\sigma_f = 0.15$, $\sigma_v = 0.15$ and $\sigma_\mu = 0.15$. In this simulation, the system is assumed initially at rest and the initial conditions for the parameter estimates are taken to be zero, reflecting the fact that the system is completely unknown. Figures 2,3 show the simulation results for the tracking errors. Figure 4 shows the action of the switching function $\rho(t)$.

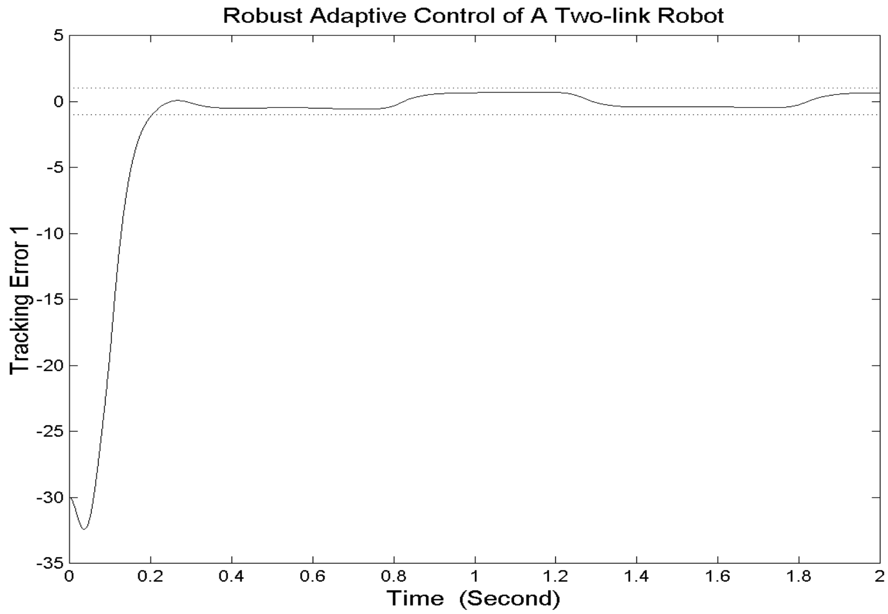

Fig. 2. Tracking error response of link 1 during the first 2 seconds
The dotted lines indicate the required error bound of ±1 degree

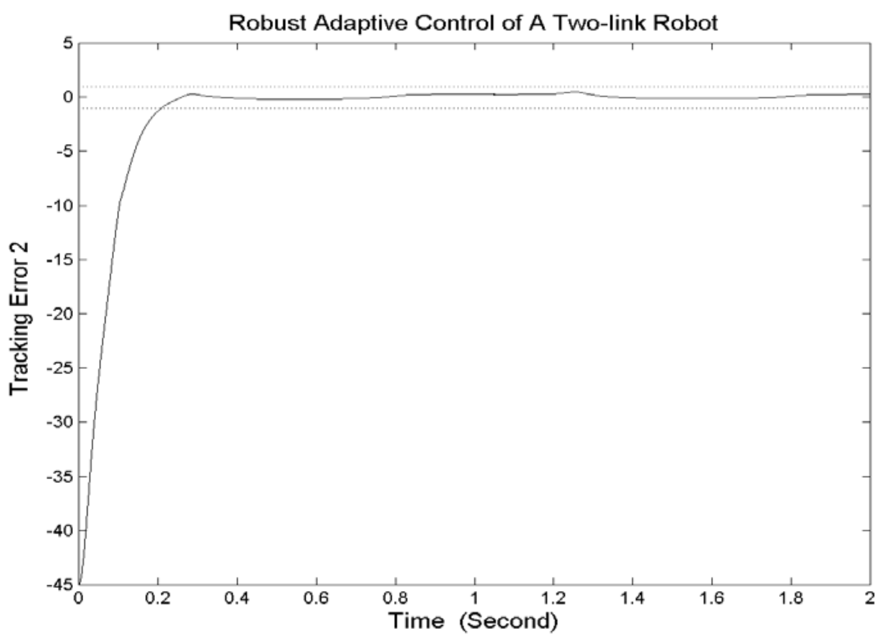

Fig. 3. Tracking error response of link 2 during the first 2 seconds
The dotted lines indicate the required error bound of ±1 degree

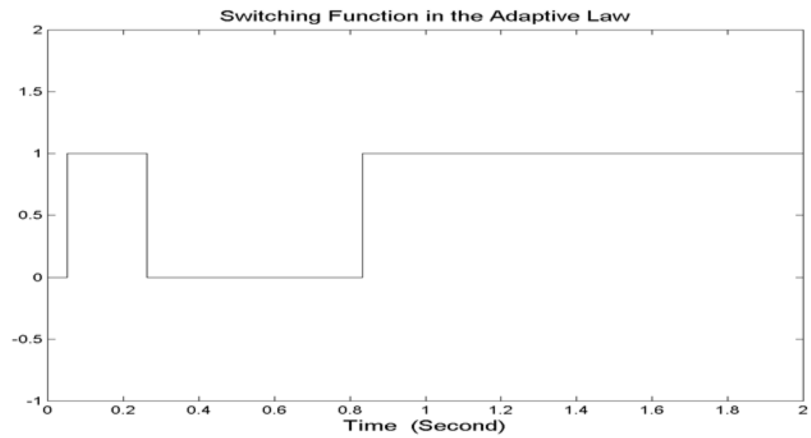

Fig. 4. The action of the switch function $\rho(t)$ during the first 2 seconds

The simulation results in Fig. 2,3 demonstrate the theory and calculated steady state error bounds. Let us now change the tracking error performance requirements to the following desired tracking error bounds

$$|e_1| = |q_1 - q_{d_1}| \le 0.5^\circ \;;\; |e_2| = |q_2 - q_{d_2}| \le 0.5^\circ \tag{90}$$

In this case we use the design inequalities (78a-d) to choose $\lambda_1 = 30$, $\lambda_2 = 30$, $\Phi = 0.26$ and $k_S = 20$, leading to $|e_1| \le \Phi / \lambda_1 = 0.5^\circ$ and $|e_2| \le \Phi / \lambda_2 = 0.5^\circ$. As demonstrated in Fig. 5,6 the algorithm meets the new error bound requirements at steady state. We should also note that as shown in Fig. 4,7 the switching function $\rho(t)$ reaches a steady state in a short period of time after which no more switching takes place.

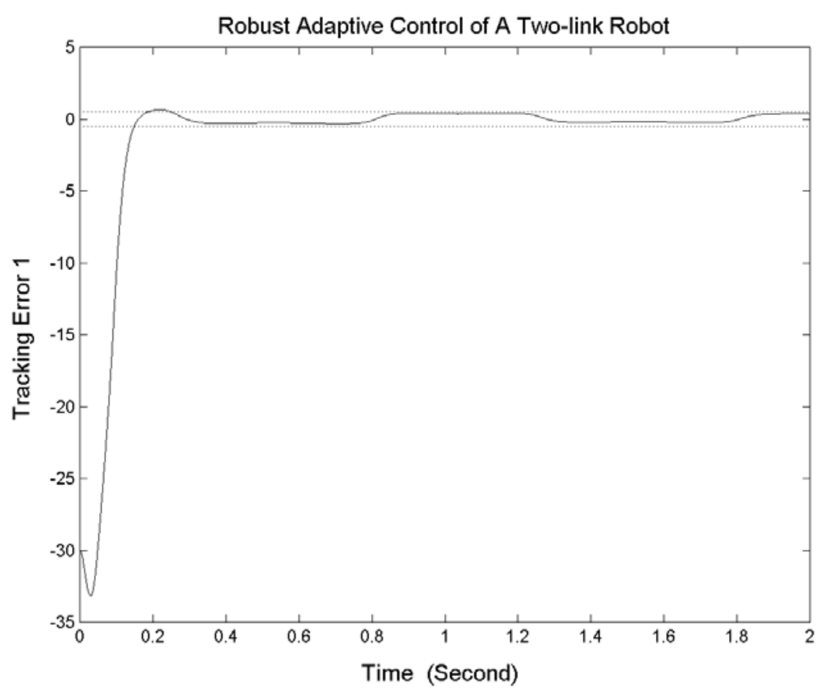

Fig. 5. Tracking error response of link 1 during the first 2 seconds
The dotted lines indicate the required error bounds ±0.5 degree

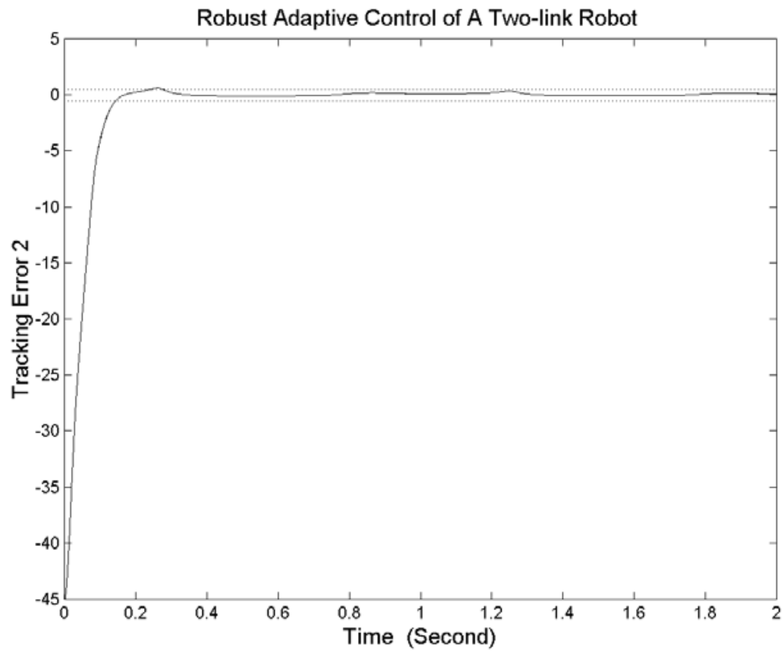

Fig. 6. Tracking error response of link 2 during the first 2 seconds
The dotted lines indicate the required error bounds ±0.5 degree

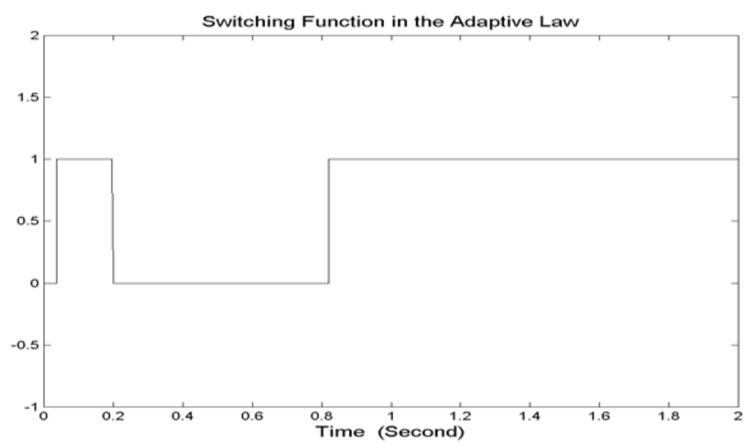

Fig. 7. The action of the switching function $\rho(t)$ during the first 2 seconds

5. CONCLUSIONS

In this paper, we consider the control problem of a class of nonlinear MIMO with unknown nonlinearities. The nonlinearities are assumed to be smooth functions and as such can be approximated and estimated on-line using a single layer neural network. A nonlinear robust adaptive control scheme is designed and analyzed. The controller guarantees closed loop semi-global stability and convergence of the tracking error to a small residual set even in the case where the estimated plant loses controllability. The semi-global 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. The size of the residual set for the tracking error depends solely on design parameters, which can be chosen to meet desired upper bounds for the tracking error. A design procedure is presented which shows how to choose the various parameters in the control law in order to guarantee that the steady state tracking error are within prespecified bounds.

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.

APPENDIX

A. Proof of Lemma 1

Let us write $\boldsymbol{B}(\boldsymbol{x})$ as the sum of a symmetric matrix and a skew-symmetric matrix as:

$$\boldsymbol{B}(\boldsymbol{x}) = \frac{\boldsymbol{B}(\boldsymbol{x}) + \boldsymbol{B}^T(\boldsymbol{x})}{2} + \frac{\boldsymbol{B}(\boldsymbol{x}) - \boldsymbol{B}^T(\boldsymbol{x})}{2} \tag{A1}$$

Then

$$\boldsymbol{S}^T(\boldsymbol{X})\boldsymbol{B}(\boldsymbol{x})\boldsymbol{S}(\boldsymbol{X}) = \boldsymbol{S}^T(\boldsymbol{X})\left(\frac{\boldsymbol{B}(\boldsymbol{x}) + \boldsymbol{B}^T(\boldsymbol{x})}{2}\right)\boldsymbol{S}(\boldsymbol{X}) \tag{A2}$$

because the quadratic form associated with a skew-symmetric matrix is always zero.

Let $\boldsymbol{S}_e(\boldsymbol{X}) = \boldsymbol{S}(\boldsymbol{X})/\|\boldsymbol{S}(\boldsymbol{X})\|$ be the unit vector. Then

$$\boldsymbol{S}^T(\boldsymbol{X})\left(\frac{\boldsymbol{B}(\boldsymbol{x}) + \boldsymbol{B}^T(\boldsymbol{x})}{2}\right)\boldsymbol{S}(\boldsymbol{X}) = \boldsymbol{S}_e^T(\boldsymbol{X})\left(\frac{\boldsymbol{B}(\boldsymbol{x}) + \boldsymbol{B}^T(\boldsymbol{x})}{2}\right)\boldsymbol{S}_e(\boldsymbol{X})\|\boldsymbol{S}(\boldsymbol{X})\|^2 \tag{A3}$$

Since $\frac{1}{2}(\boldsymbol{B}(\boldsymbol{x}) + \boldsymbol{B}^T(\boldsymbol{x}))$ is uniformly symmetric definite, all eigenvectors are orthogonal and all eigenvalues are real. Let $\boldsymbol{e}_1(\boldsymbol{x}), \cdots, \boldsymbol{e}_m(\boldsymbol{x})$ indicate a set of unit eigenvectors of the span of $\frac{1}{2}(\boldsymbol{B}(\boldsymbol{x}) + \boldsymbol{B}^T(\boldsymbol{x}))$. $\boldsymbol{e}_1(\boldsymbol{x}), \cdots, \boldsymbol{e}_m(\boldsymbol{x})$ are orthonormal and form a set of basis in space $\Re^m$. Thus the unit vector $\boldsymbol{S}_e(\boldsymbol{X})$ can be expressed as a linear combination of $\boldsymbol{e}_1(\boldsymbol{x}), \cdots, \boldsymbol{e}_m(\boldsymbol{x})$.

$$\boldsymbol{S}_e(\boldsymbol{X}) = c_1(\boldsymbol{X})\boldsymbol{e}_1(\boldsymbol{x}) + c_2(\boldsymbol{X})\boldsymbol{e}_2(\boldsymbol{x}) + \cdots + c_m(\boldsymbol{X})\boldsymbol{e}_m(\boldsymbol{x}) \tag{A4}$$

where $c_1(\boldsymbol{X}), \cdots, c_m(\boldsymbol{X})$ are corresponding coefficients and $c_1^2(\boldsymbol{X}) + c_2^2(\boldsymbol{X}) + \cdots + c_m^2(\boldsymbol{X}) = 1$.

Let $\gamma_1(\boldsymbol{x}), \gamma_2(\boldsymbol{x}), \cdots, \gamma_m(\boldsymbol{x})$ be the eigenvalues of $\frac{1}{2}(\boldsymbol{B}(\boldsymbol{x}) + \boldsymbol{B}^T(\boldsymbol{x}))$. Note that $\boldsymbol{X} = [\boldsymbol{x}, \boldsymbol{y}_d]^T$, We have

$$\begin{aligned}
\boldsymbol{S}^T(\boldsymbol{X})\left(\frac{\boldsymbol{B}(\boldsymbol{x}) + \boldsymbol{B}^T(\boldsymbol{x})}{2}\right)\boldsymbol{S}(\boldsymbol{X}) &= (c_1\boldsymbol{e}_1^T + c_2\boldsymbol{e}_2^T + \cdots + c_m\boldsymbol{e}_m^T)(c_1\gamma_1\boldsymbol{e}_1 + c_2\gamma_2\boldsymbol{e}_2 + \cdots + c_m\gamma_m\boldsymbol{e}_m)\|\boldsymbol{S}(\boldsymbol{X})\|^2 \\
&= (c_1^2(\boldsymbol{X})\gamma_1(\boldsymbol{x}) + c_2^2(\boldsymbol{X})\gamma_2(\boldsymbol{x}) + \cdots + c_m^2(\boldsymbol{X})\gamma_m(\boldsymbol{x}))\|\boldsymbol{S}(\boldsymbol{X})\|^2 \\
&= \mu(\boldsymbol{X})\|\boldsymbol{S}(\boldsymbol{X})\|^2
\end{aligned} \tag{A5}$$

Where,

$$\mu(\boldsymbol{X}) = c_1^2(\boldsymbol{X})\gamma_1(\boldsymbol{x}) + c_2^2(\boldsymbol{X})\gamma_2(\boldsymbol{x}) + \cdots + c_m^2(\boldsymbol{X})\gamma_m(\boldsymbol{x}) \tag{A6}$$

and,

$$|\mu(\boldsymbol{X})| \geq \min(|\gamma_1|, |\gamma_2|, \cdots, |\gamma_m|) \geq \underline{b} \tag{A7}$$

$\mu(\boldsymbol{X})$ is a linear combination of all eigenvalues of the matrix $\frac{1}{2}(\boldsymbol{B}(\boldsymbol{x}) + \boldsymbol{B}^T(\boldsymbol{x}))$ and is real for all $\boldsymbol{x}$.

Let $\boldsymbol{X}_0$ denotes the value of $\boldsymbol{X}$ such that $\|\boldsymbol{S}(\boldsymbol{X}_0)\| = 0$. In this case (A5) always holds. However, the value of $\mu(\boldsymbol{X}_0)$ is arbitrary. Define:

$$\mu(\boldsymbol{X}_0) := \lim_{X \to (X_0 - 0)} \mu(\boldsymbol{X}) \tag{A8}$$

Note that both $\boldsymbol{S}(\boldsymbol{X})$, $\boldsymbol{B}(\boldsymbol{x})$ are smooth functions of $\boldsymbol{X}$. From (A5), It is easy to verify that $\mu(\boldsymbol{X})$ is a smooth function of $\boldsymbol{X}$.

B. Proof of the inequality (66).

In this appendix, we prove inequality (66) used in the proof of theorem 2. Let us start with the following equation:

$$d_\mu(\boldsymbol{X})u_0 - \delta_\mu u' = \rho\{d_\mu(\boldsymbol{X})u_0 - \delta_\mu u'\} + (1-\rho)\{d_\mu(\boldsymbol{X})u_0 - \delta_\mu u'\} \tag{B1}$$

From (60), we obtain

$$\hat{\mu}^a(\boldsymbol{X},t)u_0 = \frac{(\hat{\mu}^a(\boldsymbol{X},t))^2}{(\hat{\mu}^a(\boldsymbol{X},t))^2 + \delta_\mu}\overline{u} \tag{B2}$$

Substituting $\hat{\mu}^a(\boldsymbol{X},t) = \mu^a(\boldsymbol{X}) + \tilde{\mu}^a(\boldsymbol{X},t)$ into (B2), it follows that

$$\mu^a(\boldsymbol{X})u_0 = \frac{(\hat{\mu}^a(\boldsymbol{X},t))^2}{(\hat{\mu}^a(\boldsymbol{X},t))^2 + \delta_\mu}\overline{u} - \tilde{\mu}^a(\boldsymbol{X},t)u_0 \tag{B3}$$

Therefore, the scalar control law u_0 can be expressed as:

$$u_0 = \left\{\frac{(\hat{\mu}^a(\boldsymbol{X},t))^2}{(\hat{\mu}^a(\boldsymbol{X},t))^2 + \delta_\mu}\right\}\frac{1}{\mu^a(\boldsymbol{X})}\overline{u} - \frac{1}{\mu^a(\boldsymbol{X})}\tilde{\mu}^a(\boldsymbol{X},t)u_0 \tag{B4}$$

Using the fact $|\mu^a(\boldsymbol{X})| = |\mu(\boldsymbol{X}) - d_\mu(\boldsymbol{X})| \geq \underline{b} - \psi_\mu$, one has

$$\begin{aligned} |u_0| &\leq \frac{1}{|\mu^a(\boldsymbol{X})|}|\overline{u}| + \frac{1}{|\mu^a(\boldsymbol{X})|}|\tilde{\mu}^a(\boldsymbol{X},t)||u_0| \\ &\leq \frac{1}{\underline{b} - \psi_\mu}|\overline{u}| + \frac{1}{\underline{b} - \psi_\mu}|\tilde{\mu}^a(\boldsymbol{X},t)||u_0| \end{aligned} \tag{B5}$$

From the expression of (55),(61), the scalar function u' can be written as:

$$u' = \frac{1}{(\hat{\mu}^a(\boldsymbol{X},t))^2 + \delta_\mu}\overline{u} \tag{B6}$$

Then we have

$$\{[\mu^a(\boldsymbol{X}) + \tilde{\mu}^a(\boldsymbol{X},t)]^2 + \delta_\mu\}u' = \overline{u} \tag{B7}$$

and

$$\{(\mu^a(\boldsymbol{X}))^2 + (\tilde{\mu}^a(\boldsymbol{X},t))^2 + \delta_\mu\}u' = \bar{u} - 2\mu^a(\boldsymbol{X})\tilde{\mu}^a(\boldsymbol{X},t)u' \tag{B8}$$

From (B8), u' can be expressed as

$$u' = \frac{1}{(\mu^a(\boldsymbol{X}))^2 + (\tilde{\mu}^a(\boldsymbol{X},t))^2 + \delta_\mu}\bar{u} - \frac{2\mu^a(\boldsymbol{X})}{(\mu^a(\boldsymbol{X}))^2 + (\tilde{\mu}^a(\boldsymbol{X},t))^2 + \delta_\mu}\tilde{\mu}^a(\boldsymbol{X},t)u' \tag{B9}$$

Therefore, we have

$$\begin{aligned}
|u'| &\le \frac{1}{(\mu^a(\boldsymbol{X}))^2 + (\tilde{\mu}^a(\boldsymbol{X},t))^2 + \delta_\mu}|\bar{u}| + \frac{2|\mu^a(\boldsymbol{X})|}{(\mu^a(\boldsymbol{X}))^2 + (\tilde{\mu}^a(\boldsymbol{X},t))^2 + \delta_\mu}|\tilde{\mu}^a(\boldsymbol{X},t)||u'| \\
&\le \frac{1}{(\mu^a(\boldsymbol{X}))^2 + \delta_\mu}|\bar{u}| + \frac{2|\mu^a(\boldsymbol{X})|}{(\mu^a(\boldsymbol{X}))^2 + \delta_\mu}|\tilde{\mu}^a(\boldsymbol{X},t)||u'| \\
&\le \frac{1}{(\mu^a(\boldsymbol{X}))^2 + \delta_\mu}|\bar{u}| + \frac{2}{|\mu^a(\boldsymbol{X})|}|\tilde{\mu}^a(\boldsymbol{X},t)||u'| \\
&\le \frac{1}{(\underline{b}-\psi_\mu)^2 + \delta_\mu}|\bar{u}| + \frac{2}{\underline{b}-\psi_\mu}|\tilde{\mu}^a(\boldsymbol{X},t)||u'|
\end{aligned} \tag{B10}$$

The absolute value of the first term in (B1) can be written in the following form:

$$|\rho\{d_\mu(\boldsymbol{X})u_0 - \delta_\mu u'\}| \le \rho|d_\mu(\boldsymbol{X})||u_0| + \rho\delta_\mu|u'| \tag{B11}$$

Using (B5), (B10), and $|d_\mu(\boldsymbol{X})| \le \psi_\mu$, (B11) can be rewritten as

$$\begin{aligned}
|\rho\{d_\mu(\boldsymbol{X})u_0 - \delta_\mu u'\}| &\le \rho\left\{\frac{\psi_\mu}{\underline{b}-\psi_\mu} + \frac{\delta_\mu}{(\underline{b}-\psi_\mu)^2 + \delta_\mu}\right\}|\bar{u}| + \rho\frac{2\delta_\mu}{\underline{b}-\psi_\mu}|\tilde{\mu}^a(\boldsymbol{X},t)||u'| + \rho\frac{\psi_\mu}{\underline{b}-\psi_\mu}|\tilde{\mu}^a(\boldsymbol{X},t)||u_0| \\
&\le \rho\left\{\frac{\psi_\mu}{\underline{b}-\psi_\mu-\Delta} + \frac{\delta_\mu}{(\underline{b}-\psi_\mu-\Delta)^2 + \delta_\mu}\right\}|\bar{u}| + \rho\frac{2\delta_\mu}{\underline{b}-\psi_\mu}|\tilde{\mu}^a(\boldsymbol{X},t)||u'| + \rho\frac{\psi_\mu}{\underline{b}-\psi_\mu}|\tilde{\mu}^a(\boldsymbol{X},t)||u_0|
\end{aligned} \tag{B12}$$

Also, for the second term in (B1), we obtain:

$$(1-\rho)\{d_\mu(\boldsymbol{X})u_0 - \delta_\mu u'\} = (1-\rho)\frac{d_\mu(\boldsymbol{X})\hat{\mu}^a(\boldsymbol{X},t)}{(\hat{\mu}^a(\boldsymbol{X},t))^2 + \delta_\mu}\bar{u} - (1-\rho)\frac{\delta_\mu}{(\hat{\mu}^a(\boldsymbol{X},t))^2 + \delta_\mu}\bar{u} \tag{B13}$$

Using the fact $(1-\rho) \ne 0$ only if $|\hat{\mu}^a(\boldsymbol{X},t)| > \underline{b} - \psi_\mu - \Delta$, we have

$$\begin{aligned}
|(1-\rho)\{d_\mu(\boldsymbol{X})u_0 - \delta_\mu u'\}| &\le (1-\rho)\frac{|d_\mu(\boldsymbol{X})|}{|\hat{\mu}^a(\boldsymbol{X},t)|}|\bar{u}| + (1-\rho)\frac{\delta_\mu}{(\hat{\mu}^a(\boldsymbol{X},t))^2 + \delta_\mu}|\bar{u}| \\
&\le (1-\rho)\left\{\frac{\psi_\mu}{\underline{b}-\psi_\mu-\Delta} + \frac{\delta_\mu}{(\underline{b}-\psi_\mu-\Delta)^2 + \delta_\mu}\right\}|\bar{u}|
\end{aligned} \tag{B14}$$

Combining the results of (B12) and (B14), it follows that:

$$
\begin{aligned}
\left|d_\mu(\boldsymbol{X})u_0-\delta_\mu u'\right| &\le \left\{\frac{\psi_\mu}{\underline{b}-\psi_\mu-\Delta}+\frac{\delta_\mu}{(\underline{b}-\psi_\mu-\Delta)^2+\delta_\mu}\right\}\left|\overline{u}\right|+\rho\frac{2\delta_\mu}{\underline{b}-\psi_\mu}\left|\widetilde{\mu}^a(\boldsymbol{X},t)\right|\left|u'\right|+\rho\frac{\psi_\mu}{\underline{b}-\psi_\mu}\left|\widetilde{\mu}^a(\boldsymbol{X},t)\right|\left|u_0\right| \\
&= \delta_1\left|\overline{u}\right|+\rho\delta_2\left|\widetilde{\mu}^a(\boldsymbol{X},t)\right|\left|u'\right|+\rho\delta_3\left|\widetilde{\mu}^a(\boldsymbol{X},t)\right|\left|u_0\right|
\end{aligned}
\tag{B15}
$$

where δ_1, δ_2, δ_3 are defined in (65a-c).

REFERENCES

[1] P. A. Ioannou and J. Sun, *Robust Adaptive Control*, Upper Saddle River, NJ: Prentice Hall, 1996.

[2] P. Kokotovic and M. Arcak, "Constructive nonlinear control: a historical perspective", *Automatica*, vol. 37, pp. 637-662, 2001.

[3] B. Fernandez R. and J. K. Hedrick, "Control of multivariable non-linear systems by the sliding model", *Int. J. Control*, vol. 46, no. 3, pp. 1019-1040, 1987.

[4] J.-J. E. Slotine and W. Li, *Applied Nonlinear Control*, Englewood Cliffs, NJ: Prentice Hall, 1991.

[5] D. N. Godbole and S. S. Sastry, "Approximate decoupling and asymptotic tracking for MIMO systems", *IEEE Trans. Automat. Contr.*, vol. 40, no. 3, March 1995.

[6] I. Kanellakopoulos, P. V. Kokotovic, and A. S. Morse, "Systematic design of adaptive controllers for feedback linearizable systems", *IEEE Trans. Automatic. Contr.*, vol. 36, no. 11, pp. 1241-1253, 1991.

[7] E. B. Kosmatopoulos and P. A. Ioannou, "A switch adaptive controller for feedback linearizable systems", *IEEE Trans. Automat. Contr.*, vol. 44, no.4, pp. 742-750, April 1999.

[8] M. Krstic, I. Kanellakopoulos, and P. Kokotovic, *Nonlinear and Adaptive Control Design*, New York: Wiley Interscience, 1995.

[9] S. S. Sastry and A. Isidori, "Adaptive control of linearizable systems", *IEEE Trans. Automat. Contr.*, vol. 34, no. 11, pp. 1123-1131, Nov. 1989.

[10] J.-J. E. Slotine, and J. A. Coetsee, "Adaptive sliding controller synthesis for non-linear systems", *Int. J. Control*, vol. 43, no. 6, 1631-1651, 1986.

[11] D. G. Taylor, P. V. Kokotovic, R. Marino and I. Kanellakopoulos, "Adaptive regulation of nonlinear systems with unmodeled dynamics", *IEEE Trans. Automat. Contr.*, vol. 34, no. 4, pp. 405-412, April 1989.

[12] D. Seto, A. M. Annaswamy, and J. Baillieul, "Adaptive control of nonlinear systems with a triangular structure", *IEEE Trans. Automat. Contr.*, vol. 39, pp. 1411-1428, 1994.

[13] H. K. Khalil, "Adaptive output feedback control of nonlinear systems represented by input-output models", *IEEE Trans. Automat. Contr.*, vol. 41, no. 2, 177-188, 1996.

[14] B. Aloliwi and H. K. Khalil, "Robust adaptive output feedback control of nonlinear systems without persistence of excitation", *Automatica*, vol. 33, no. 11, pp. 2025-2032, 1997.

[15] Y.-C. Chang, "An adaptive H^{∞} tracking control for a class of nonlinear multi-input-multi-output (MIMO) systems", *IEEE Trans. Automat. Contr.*, vol. 46, no. 9, pp. 1432-1437, September 2001.

[16] B. Yao and M. Tomizuka, "Adaptive robust control of MIMO nonlinear systems with guaranteed transient performance", in *Proc. 34th IEEE Conf. Decision & Contr.*, New Orleans, LA, pp. 2346-2351, December 1995.

[17] J. J. Craig, P. Hsu, and S. S. Sastry, "Adaptive control of mechanical manipulator", *International Journal of Robotics Research*, vol. 6, no. 2, pp. 17-28, 1987.

[18] J.-J. Slotine and W. Li, "Composite adaptive control of robot manipulators", *Automatica*, vol. 25, no. 4, pp. 509-519, 1989.

[19] A. M. Annaswamy, F. P. Skantze, and A.-P. Loh, "Adaptive control of continuous time systems with convex/concave parameterization", *Automatica*, vol. 34, no. 1, pp. 33-49, 1998.

[20] A. Kojic, A. M. Annaswamy, A.-P. Loh, and R. Lozano, "Adaptive control of a class of nonlinear systems with convex/concave parameterization", *Systems & Control Letters*, vol. 37, pp. 267-274, 1999.

[21] A. Kojic and A. M. Annaswamy, "Adaptive control of nonlinearly parameterized systems with a triangular structure", *Automatica*, vol. 38, pp. 115-123, 2002.

[22] M. S. Netto, A. M. Annaswamy, R. Ortega, and P. Moya, "Adaptive control of a class of non-linearly parametrized systems using convexification", *Int. J. Control*, vol. 73, no. 14, pp. 1312-1321, 2000.

[23] S.-H. Yu and A. M. Annaswamy, "Adaptive control of nonlinear dynamic systems using θ-adaptive neural networks", *Automatica*, vol. 33, no. 11, pp. 1975-1995, 1997.

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

[25] M. M. Polycarpou, "Stable adaptive neural control scheme for nonlinear systems", *IEEE Trans. Automat. Contr.*, vol. 41, no. 3, pp. 447-451, Mar. 1996.

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

[27] M. M. Polycarpou and P. A. Ioannou, "Modeling, identification and stable adaptive control of continuous-time nonlinear dynamical systems using neural networks", in *Proc. 1992 Amer. Contr. Conf.*, June 1992, pp. 36-40.

[28] S. Seshagiri and H. Khalil, "Output feedback control of nonlinear systems using RBF neural networks", *IEEE Trans. Automat. Contr.*, vol. 11, no. 1, pp. 69-79, Jan. 2000.

[29] G. A. Rovithakis, "Stable adaptive neuro-control design via Lyapunov function derivative estimation", *Automatica*, vol. 37, pp. 1213-1221, 2001.

[30] F.-C. Chen and H. K. Khalil, "Adaptive control of nonlinear systems using neural networks", *Int. J. Control*, vol. 55, no. 6, pp. 1299-1317, 1992.

[31] F.-C. Chen and C.-C. Liu, "Adaptively controlling nonlinear continuous-time systems using multilayer neural networks", *IEEE Trans. Automat. Contr.*, vol. 39, no. 6, pp. 1306-1310, June 1994.

[32] J. T. Spooner and K. M. Passino, "Stable adaptive control using fuzzy systems and neural networks", *IEEE Trans. Neural Networks*, vol. 4, no. 3, pp. 339-359, August 1996.

[33] R. Sanner and J.-J. E. Slotine, "Gaussian networks for direct adaptive control", *IEEE Trans. Neural Networks*, vol. 3, no. 6, pp. 837-863, Nov. 1992.

[34] H. Xu and P. A. Ioannou, "Robust adaptive control of linearizable nonlinear single input systems", in *Proc 2002 XV IFAC world congress*, Barcelona, Spain, July 2002.

[35] A. Yesildirek and F. L. Lewis, "Feedback linearization using neural networks", *Automatica*, vol. 31, no. 11, pp. 1659-1664, 1995.

[36] S. Jagannathan and F. L. Lewis, "Multilayer discrete-time neual-net controller with guaranteed performance", *IEEE Trans. Automat. Contr.*, vol. 7, no. 1, pp. 107-130, January 1996.

[37] C. Kwan, and F. L. Lewis, "Robust backstepping control of nonlinear systems using neural networks", *IEEE Trans. Syst. Man, Cybern. A*, vol. 30, no. 6, pp. 753-766, Nov. 2000.

[38] Y. H. Kim and F. L. Lewis, "Neural network output feedback control of robot manipulators", *IEEE Trans. Robot. Automat.*, vol. 15, no. 2, pp. 301-309, April 1999.

[39] F. L. Lewis, A. Yesildirek, and K. Liu, "Multilayer neural-net robot controller with guaranteed tracking performance", *IEEE Trans. Neural Networks*, vol. 7, no. 2, pp. 388-399, Mar. 1996.

[40] F. L. Lewis, K. Liu, and A. Yesildirek, "Neural net robot controller with guaranteed tracking performance", *IEEE Trans. Neural Networks*, vol. 6, no. 3, pp. 703-715, May 1995.

[41] T. A. Johansen, "Fuzzy model based control: stability, robustness, and performance issues", *IEEE Trans. Fuzzy Systems*, vol. 2, no. 3, pp. 221-234, August 1994.

[42] C.-C. Liu and F.-C. Chen, "Adaptive control of nonlinear continuous-time systems using neural networks – general relative degree and MIMO cases", *Int. J. Control*, vol. 58, no. 2, pp. 317-335, 1993.

[43] F.-C. Chen and H. K. Khalil, "Adaptive control of linearizable discrete systems", in *Proc. 34th IEEE Conf. Decision & Contr.*, New Orleans, LA, pp. 2427-2432, December 1995.

[44] R. Ordonez and K. M. Passino, "Stable multi-input multi-output adaptive fuzzy/neural control", *IEEE Trans. Fuzzy Systems*, vol. 7, no. 3, pp. 345-353, June 1999.

[45] M. M. Polycarpou and P. A. Ioannou, "On the existence and uniqueness of solutions in adaptive control systems", *IEEE Trans. Automat. Contr.*, vol. 38, no. 3, pp. 474-479, Mar. 1993.

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

[47] K. Funahashi, "On the approximate realization of continuous mappings by neural networks", *Neural Networks*, vol. 2, pp. 183-192, 1989.

[48] J. Park and I. W. Sandberg, "Approximation and radial-basis-function networks", *Neural Computation*, vol. 5, pp. 305-316, 1993.