

ADAPTIVE CONTROL OF NONLINEAR SYSTEMS USING MULTIPLE
IDENTIFICATION MODELS

by

Ahmet Cezayirli

B.S., Electronics Engineering, Istanbul University, 1995

M.S., Systems and Control Engineering, Boğaziçi University, 1999

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

Graduate Program in Electrical and Electronics Engineering
Boğaziçi University

2007

To my brother, Osman Cezayirli

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my advisor, Prof. Kemal Cılız, for his guidance and invaluable help throughout the thesis. Working with him was really a privilege and good luck for me. He provided me with perception of many concepts as well as many academic assets. His careful and incisive yet constructive personality enforced me to produce always better, and took the thesis step-by-step further till this end. His contributions will be always effective in the rest of my academic life.

Special thanks to Prof. Yorgo İstefanopulos for his presence in every stage of my graduate study. He has played a major role in the completion of this study, with both his deep theoretical knowledge and advise at the very critical times. It was an honor for me to be a student of such a great person.

Prof. Eşref Eşkinat was very kind for participating all my progress report presentations and sharing his valuable comments. I was very lucky to have his contributions.

I am indebted much to Prof. Feza Kerestecioğlu. He has always supported me from the very beginning of my doctoral study, with his sincere comments, friendly attitude and excellent L^AT_EX knowledge.

Finally, it is a pleasure for me to thank Prof. Kadri Özçaldıran, dean of the Engineering Faculty, for his kindness to be in my PhD committee.

ABSTRACT

ADAPTIVE CONTROL OF NONLINEAR SYSTEMS USING MULTIPLE IDENTIFICATION MODELS

Adaptive control of nonlinear systems is considered in this study. Available methods in this field are reviewed first. Focusing on the minimum-phase, input-output linearizable and linearly parameterized nonlinear systems, direct and indirect adaptive controllers are developed for the cases of matched and unmatched uncertainties. A new methodology is proposed, which makes use of multiple identification models in order to improve the transient performance under large parametric uncertainties. Adaptation and switching mechanisms are developed based on the use of multiple Lyapunov functions and cost functions. The resulting closed loop systems are shown to be stable with the switching mechanisms. Combination of direct and indirect adaptive control schemes is also presented for a class of nonlinear systems which do not give rise to over-parametrization. The theoretical results obtained in this study are verified by computer simulations.

ÖZET

DOĞRUSAL OLMAYAN SİSTEMLERİN ÇOKLU TANIMLAMA MODELLERİ İLE UYARLAMALI DENETİMİ

Bu tez çalışmasında doğrusal olmayan sistemlerin uyarlamalı denetimi konu edilmiştir. Bu alandaki mevcut çalışmalar özetlendikten sonra, minimum-faz, giriş-çıkış doğrusallaştırılabilen ve doğrusal olarak parametrelendirilmiş doğrusal olmayan sistemlerde uygulanmak üzere, belirsizliklerin uyumlu ve uyumsuz olduğu durumlar için doğrudan ve dolaylı uyarlamalı denetleyiciler geliştirilmiştir. Parametrelerde büyük belirsizlikler olması durumunda sistemin geçici yanıtını iyileştirmek için, çoklu tanımlama modellerinin kullanıldığı yeni bir yöntem önerilmiştir. Çoklu Lyapunov fonksiyonlarını ve maliyet fonksiyonlarını temel alan uyarlama ve anahtarlama mekanizmaları geliştirilmiş ve oluşan kapalı çevrim sistemlerin bu anahtarlama mekanizması altında kararlı olduğu gösterilmiştir. Uyarlamalı denetimde doğrudan ve dolaylı yöntemler birleştirilerek, bunun aşırı parametrelendirmeye neden olmayan doğrusal olmayan sistemlere uygulaması yapılmıştır. Bu çalışmadaki kuramsal bulgular bilgisayar benzetimleri ile doğrulanmıştır.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	iv
ABSTRACT	v
ÖZET	vi
LIST OF FIGURES	x
LIST OF TABLES	xii
LIST OF SYMBOLS/ABBREVIATIONS	xiii
1. INTRODUCTION	1
1.1. Background and Motivation	1
1.2. Contributions	3
1.3. Outline of the Thesis	4
2. MATHEMATICAL PRELIMINARIES	6
2.1. Vector Fields	6
2.2. Lie Derivatives and Lie Bracket	7
2.3. Lipschitz Continuity	8
2.4. Vector Norms	8
2.5. Signal Norms and Signal Spaces	9
2.6. Signal Convergence Lemmas	11
2.7. Frobenius Theorem	13
2.8. Swapping Lemma	14
3. NONLINEAR DYNAMIC SYSTEMS	16
3.1. Nonlinear System Modeling	16
3.2. Linearization	18
3.2.1. Feedback Linearization	18
3.2.2. Input-State Linearization	19
3.2.3. Input-Output Linearization	21
3.2.4. Canonical Forms Used in Nonlinear Adaptive Control	25
3.2.4.1. Brunovsky Canonical Form	25
3.2.4.2. Strict-Feedback Form	26
3.2.4.3. Pure-Feedback Form	27

3.2.4.4. The Normal Form	27
3.3. Uncertainty	28
3.3.1. Exogenous Uncertainty and Disturbance Decoupling Problem	28
3.3.2. Structural Uncertainty	31
3.3.3. Parametric Uncertainty	33
3.4. Summary	34
4. PROBLEM STATEMENT	36
4.1. System Description	36
4.2. Performance in an Adaptive System	39
5. ADAPTIVE CONTROL OF NONLINEAR SYSTEMS	41
5.1. Identification of Unknown Parameters	42
5.1.1. Observer-Based Identification	43
5.1.2. Regressor-Filtering Based Identification	45
5.2. Adaptive Control of Nonlinear Systems	47
5.2.1. Adaptive Control under Matching Conditions	48
5.2.1.1. Matching Conditions Examples	48
5.2.1.2. Definition of the Matching Conditions	53
5.2.2. Adaptive Control under Growth Conditions	57
5.3. Indirect Adaptive Control	57
5.4. Model Reference Adaptive Control under Matching Conditions	67
5.5. Direct Adaptive Control	72
5.6. Summary	77
6. ADAPTIVE CONTROL USING MULTIPLE MODELS	79
6.1. Introduction	79
6.2. Switching between Multiple Models	84
6.3. Indirect Adaptive Control Using Multiple Models	
without Matching Conditions	87
6.3.1. Simulation Study	93
6.4. Summary	98
7. DIRECT ADAPTIVE METHODS USING MULTIPLE MODELS	99
7.1. Direct Adaptive Control Using Multiple Models	
under Matching Conditions	99

7.2. Direct Adaptive Control Using Multiple Models	
without Matching Conditions	105
7.2.1. Simulation Study	125
7.3. Summary	129
8. COMBINED DIRECT AND INDIRECT ADAPTIVE CONTROL USING MUL-	
TIPLE MODELS	130
8.1. The Over-parametrization Problem	131
8.2. Direct and Indirect Adaptive Control of Nonlinear Systems	137
8.2.1. Direct Adaptive Scheme	138
8.2.2. Indirect Adaptive Control and Nonlinear System Identification .	139
8.3. Combined Direct and Indirect Adaptive Scheme	140
8.4. Combined Direct and Indirect Adaptive Control	
Using Multiple Models	143
8.5. Summary	149
9. CONCLUSIONS	150
APPENDIX A: EXAMPLES	153
A.1. Simple Pendulum ($n = 2, \gamma = 2$)	153
A.2. Simple Pendulum ($n = 2, \gamma = 1$)	158
A.3. Fictitious System ($n = 2, \gamma = 2$)	159
A.4. Motor Controlled Nonlinear Valve ($n = 3, \gamma = 3$)	164
A.4.1. Case 1	165
A.4.2. Case 2	168
A.4.2.1. The Extended Matching Conditions	169
A.4.2.2. The Direct Adaptive Control Method	170
A.5. Field-Controlled DC Motor ($n = 2, \gamma = 1$)	172
A.6. Field-Controlled DC Motor ($n = 2, \gamma = 1$)	174
REFERENCES	177

LIST OF FIGURES

Figure 4.1.	Transient error characteristics	40
Figure 6.1.	Representation of parameter convergence in $\mathcal{S} \subset \mathbb{R}^p$ using only a single model.	80
Figure 6.2.	Representation of parameter convergence in $\mathcal{S} \subset \mathbb{R}^p$ using multiple adaptive models.	81
Figure 6.3.	Representation of parameter convergence with switching and tuning, using multiple fixed models and one adaptive model.	82
Figure 6.4.	Change in the plant parameter, followed by a switching and tuning, using multiple fixed models and one adaptive model.	83
Figure 6.5.	Block diagram of the indirect adaptive controller with N fixed models and one adaptive model.	94
Figure 6.6.	Simulation results for single model and multiple model cases of the indirect adaptive control scheme given in Section 6.3.1.	97
Figure 6.7.	Switching model index (0: adaptive model, 1-6: fixed models) of the indirect adaptive control scheme of Section 6.3.1.	98
Figure 7.1.	Block diagram of the direct adaptive controller with N fixed models and one adaptive model.	124
Figure 7.2.	Simulation results for single model and multiple model cases of the direct adaptive control scheme in Section 7.2.1.	128

Figure 7.3. Switching between the models (0: adaptive model, 1-6: fixed models) for the direct adaptive control scheme of Section 7.2.1.	129
Figure A.1. Block diagram of the nonlinear valve system	164

LIST OF TABLES

Table 4.1. Description of the symbols in transient error measurement 40

LIST OF SYMBOLS/ABBREVIATIONS

A^{-1}	Inverse of matrix A
A^T	Transpose of matrix A
$ad_f g$	Lie bracket vector fields f, g
E	Cost function for transient performance
e	Tracking error
$\mathcal{F}, \mathcal{G}$	Vectors of multi-linear Lie derivatives
$[f, g]$	Lie bracket
f, g, h	Vector fields of appropriate orders
I	Identity matrix of appropriate order
J	Cost function
J_j	Performance index of j^{th} -model
$L_f h$	Lie derivative of h w.r.t. f
$\mathcal{L}_2$	Space of functions having 2-norm
$\mathcal{L}_\infty$	Space of bounded functions
$\mathcal{L}_{\infty e}$	Space of bounded truncated functions
$\mathbb{R}$	Set of real numbers
$\mathbb{R}_+$	Set of strictly positive real numbers, $(0, \infty)$
$\mathbb{R}^n$	n -dimensional real Euclidean space
$\mathbb{R}^{n \times p}$	Space of $n \times p$ real matrices
T_{min}	Dwell time in switching
u	Control input of a system
$V_c(\cdot)$	Closed loop Lyapunov function
$V_j(\cdot)$	Identification Lyapunov function of j^{th} -model
ΔV	Change in Lyapunov function
v	Tracking control input
y	Output of a system
y_r	Reference output for tracking
x	State vector of a system
$\hat{x}$	Estimate of the state vector

$\tilde{x}$	State estimation error
$\dot{x}$	Time derivative of the state x
x_f	Low-pass filtered x
η	Vector of un-linearizable states (states of internal dynamics)
γ	Relative degree
ϕ	Linearizing diffeomorphism
ϕ^{-1}	Inverse of linearizing diffeomorphism
θ	Parameter vector
θ_j	Parameter vector of j^{th} -model
$\hat{\theta}$	Estimate of the parameter vector
$\tilde{\theta}$	Parameter estimation error
Θ	Over-parameterized parameter vector
ξ	Vector of linearized states
$ \cdot $	Absolute value of a scalar; 1-norm of a vector
$\ \cdot\ $	The 2-norm, a.k.a. Euclidean norm
$\ \cdot\ _\infty$	The infinity norm, a.k.a. essential bound
$\ \cdot\ _t$	The truncated infinity norm, a.k.a. $\mathcal{L}_{\infty e}$ norm
BIBS	Bounded-Input Bounded-State
BIBO	Bounded-Input Bounded-Output
MMAC	Multiple Model Adaptive Control
RLS	Recursive Least Squares
SISO	Single-Input Single-Output
SPR	Strictly Positive Real

1. INTRODUCTION

1.1. Background and Motivation

Control of nonlinear systems based on feedback linearization is considered in this thesis. If the nonlinear system is satisfactorily linearized, then several design and analysis tools available in linear control theory can be applied. The feedback linearization is based on the principle of cancelling the nonlinearities in the system via feedback. The feedback may be partial or full state feedback, as well as output feedback. Cancellation of the nonlinearities resembles the pole-zero cancellation in linear systems. If the control scheme relies on the exact cancellation of the nonlinearities, and if the applied linearization becomes inexact, it is highly probable that, this technique may give rise to stability problems. Even if the stability is not affected, the inexact linearization would probably cause performance deterioration in an unacceptable manner.

Reasons for the inexact linearization are usually imperfect mathematical modeling of the physical system and unpredictable changes in the physical environment. Control techniques, such as *robust control*, can yield good performance for these cases. It is possible to determine the bounds on the system uncertainties and design a robust control scheme, which guarantees stable control of the system as long as the uncertainties stay within these bounds. However, robust control does not enhance its performance with time. When a certain amount of error is present in a robust control system, the error lasts even in a repetitive task. Solution to this problem is the use of *adaptive control* methods.

If the uncertainties in the system are in form of some unknown parameters, then an adaptive scheme can be applied such that these unknown parameters are identified. The physical system whose perfect model is not available due to the parametric uncertainties can thus be perfectly linearized. Then a linear control technique can be applied for this linear model. If there are any changes in the system parameters afterwards, adaptive control accommodates for these new conditions.

The global stability proof of adaptive control of linear systems was developed in 1980 [1, 2, 3]. In the following years, the adaptive control theory was developed and became more structured. Direct and indirect techniques, model-reference and self-tuning techniques, continuous-time and discrete-time schemes were all proposed for linear systems [4, 5, 6]. On the other hand, adaptive control of nonlinear systems gained interest in the late 1980s, and it has been an active research topic since the introduction of geometric methods for the analysis of such systems.

A model-reference adaptive controller for pure-feedback nonlinear systems was proposed in [7]. A similar scheme with certain restrictions on the type of nonlinearities was studied in [8]. This approach required Lipschitz continuity on the nonlinearities and exponentially stable zero-dynamics. This direct adaptive control scheme also led to over-parametrization of the parameter space. Indirect adaptive control approaches which relax this over-parametrization were proposed in [9] and [10]. On the other hand, adaptive schemes with no restriction on the type of nonlinearities but with certain assumptions on the location of uncertainties were proposed as well. These schemes satisfying the so-called *matching conditions* were presented in [11] and [12]. These restrictive matching conditions were relaxed in [13] and named as the *extended matching conditions*. A methodology called the *backstepping method*, which removed the necessity for the matching conditions for parametric pure-feedback nonlinear systems was proposed in [14].

The adaptive schemes discussed above can be also classified as *direct* and *indirect* methods. In the direct scheme, the controller parameters which are based on system parameters are updated with the tracking error. However, in indirect schemes, identification and control are separated. First, unknown system parameters are estimated by an adaptive mechanism which is driven by the identification error, then these obtained parameters are used to build the *certainty equivalence* based control signal.

All of the above results on adaptive control of nonlinear systems provide asymptotic convergence and there is no focus on the transient behavior of the tracking error. However, when the parameter errors are large, the transient response of the system may

include unacceptably large peaks. Although the system is asymptotically stable, the adaptive control approach may be unapplicable for some systems due to these transient peaks. The objective of this study is to achieve improvement in the transient response of adaptive control of nonlinear systems using multiple models and switching.

The switching idea in adaptive control was first suggested in [15] in an effort to minimize the *a priori* assumptions for adaptively stabilizing a linear plant. This approach, where switching between control inputs was based on direct observation of the output, has been extended in [16]. A similar approach was proposed in [17] to adaptively control a plant with unknown relative degree. This methodology was further studied in [18, 19] and the references therein. Switching ideas for transient performance improvement were first proposed and studied in [20] and [21], where multiple adaptive models were used to improve the transient response of a linear dynamic system. These theoretical results are reviewed in [22] and applied for the adaptive control of linear plants with proposed modifications of the switching algorithm to enhance performance. The theory was later extended to switching and tuning in wider contexts as presented in [23, 24, 25, 26], which also included an application to robotic manipulators. Performance improvement methods based on parameter switching were also applied for adaptive backstepping control and sliding mode type controllers in [27] and [28], respectively. In this thesis, the above results are extended for a class of nonlinear systems.

1.2. Contributions

This study contributes to the theory of adaptive control of nonlinear systems in the following fields:

1. The multiple model based adaptive control (MMAC) concept has been extended to nonlinear systems. First, a MMAC scheme for state feedback linearizable systems has been developed. Then, the class of the nonlinear systems to which the multiple model adaptive control can be applied is extended. A more general

class of nonlinear SISO systems of the form

$$\dot{x} = f(x) + g(x)u$$

is considered. This system is assumed to have parametric uncertainties and Lipschitz nonlinearities [29].

2. MMAC is inherently an indirect adaptive control methodology. In this study, MMAC systems based on direct control scheme are proposed with stability proofs. The over-parameterized parameter estimate vector of the direct control is calculated using parameters of fixed models, and re-initialized with these fixed models based on a switching logic. The direct adaptive controllers are derived both for matched and mismatched uncertainties [30, 31].
3. Combination of the direct and indirect approaches has been proposed for the adaptive control of nonlinear systems using multiple models. This combined approach is presented for a class of nonlinear systems in this work [32].

1.3. Outline of the Thesis

This thesis is organized as follows: A general introduction and a summary of previous contributions in this field are given in Chapter 1. In Chapter 2, mathematical preliminaries, which include basic notions of differential geometric control theory and fundamental lemmas, that are used throughout the thesis are outlined. Characteristics of nonlinear dynamic systems, linearization techniques, canonical forms and uncertainty types are discussed in Chapter 3. The main problem of this study is defined in Chapter 4. Chapter 5 is devoted to the adaptive control of nonlinear systems where various adaptive control methods are reviewed. An indirect adaptive controller for nonlinear systems with mismatched uncertainties, and two direct adaptive controllers, one for nonlinear systems satisfying matching conditions and the other one for nonlinear systems with mismatched uncertainties are presented. The idea of using multiple models in adaptive control, switching and transient performance improvement is introduced in Chapter 6. An indirect adaptive controller using multiple models is also given in this chapter. Solution to the main problem and transient performance

enhancement are presented both theoretically and by computer simulations using the indirect scheme. The multiple models and switching idea is applied to the direct adaptive schemes in Chapter 7. Performance improvement on the transient response is also shown for the direct adaptive schemes. In Chapter 8, the combination of direct and indirect adaptive control schemes using multiple models is proposed and a combined methodology is derived. Conclusions of the thesis and suggestions for future work are presented in Chapter 9.

2. MATHEMATICAL PRELIMINARIES

Fundamental mathematical concepts are introduced in this chapter that are to be used throughout the thesis. A detailed discussion about these concepts and their relevance in control theory can be found in [33, 34, 35, 36].

2.1. Vector Fields

Definition 2.1 (Smoothness) *Let $\mathcal{U} \subset \mathbb{R}^n$ and $\mathcal{V} \subset \mathbb{R}^m$ be open sets. A function $f : \mathcal{U} \mapsto \mathcal{V}$ is called smooth, if its every component has continuous partial derivatives of all orders.*

Note that the above definition requires availability of infinite number of continuous partial derivatives. This is not necessary in some applications. Using a looser term, we say f is *sufficiently smooth*, if it has sufficient number of continuous partial derivatives.

Definition 2.2 (Vector field) *A vector field $f : \mathcal{D} \subset \mathbb{R}^n \mapsto \mathbb{R}^n$ is a function, which assigns an n -dimensional vector to every point p in the n -dimensional space $\mathcal{D}$.*

Definition 2.3 (Diffeomorphism) *A function $f : \mathcal{D} \subset \mathbb{R}^n \mapsto \mathbb{R}^n$ is called a local diffeomorphism, if*

1. *it is smooth (i.e. continuously differentiable on $\mathcal{D}$)*
2. *its inverse f^{-1} , defined by,*

$$f^{-1}(f(x)) \stackrel{\text{def}}{=} x \quad , \quad \forall x \in \mathcal{D} \tag{2.1}$$

exists and is smooth.

2.2. Lie Derivatives and Lie Bracket

Definition 2.4 (Lie derivative) *Having a smooth vector field $f : \mathcal{D} \subset \mathbb{R}^n \mapsto \mathbb{R}^n$ and a smooth scalar function $h : \mathcal{D} \subset \mathbb{R}^n \mapsto \mathbb{R}$, the Lie derivative of h with respect to the vector field f is a new function $L_f h : \mathcal{D} \subset \mathbb{R}^n \mapsto \mathbb{R}$ given by,*

$$L_f h(p) \stackrel{\text{def}}{=} f(h)(p) \quad (2.2)$$

In local coordinates, $L_f h(p)$ is represented by,

$$L_f h(x) \stackrel{\text{def}}{=} \frac{\partial h}{\partial x} f(x) \quad (2.3)$$

The Lie derivative of a function h with respect to a vector field f is the rate of change of h in the direction of f . We have

$$L_f^k h \stackrel{\text{def}}{=} L_f(L_f^{k-1} h) \quad , \quad k = 1, 2, \dots \quad (2.4)$$

with $L_f^0 h \stackrel{\text{def}}{=} h$, and

$$L_f L_g h \stackrel{\text{def}}{=} L_f(L_g h) \quad (2.5)$$

where f and g are vector fields.

Definition 2.5 (Lie bracket) *Having two smooth vector fields $f, g : \mathcal{D} \subset \mathbb{R}^n \mapsto \mathbb{R}^n$, the Lie bracket is a new vector field, denoted by $[f, g]$, which is defined by,*

$$[f, g](x) \stackrel{\text{def}}{=} \frac{\partial g}{\partial x} f(x) - \frac{\partial f}{\partial x} g(x) \quad (2.6)$$

An important notation about the Lie bracket is the following:

$$ad_f g \stackrel{def}{=} [f, g] \quad (2.7)$$

$$ad_f^k g \stackrel{def}{=} [f, ad_f^{k-1} g] \quad , \quad k = 1, 2, \dots \quad (2.8)$$

with $ad_f^0 g = g$.

2.3. Lipschitz Continuity

Definition 2.6 (Lipschitz Continuity) *A function $f(x) : \mathbb{R}^n \mapsto \mathbb{R}^m$ is said to be locally Lipschitz on an open set $\mathcal{D} \subset \mathbb{R}^n$, if every point of $\mathcal{D}$ has a neighborhood $\mathcal{D}_0 \subset \mathcal{D}$, over which there exists a constant $L \geq 0$, such that,*

$$\|f(x_1) - f(x_2)\| \leq L\|x_1 - x_2\| \quad , \quad \forall x_1, x_2 \in \mathcal{D} \quad (2.9)$$

The smallest constant L satisfying this inequality is called the Lipschitz constant. If, the function f satisfies (2.9) for $\mathcal{D} = \mathbb{R}^n$, then it is said to be globally Lipschitz.

2.4. Vector Norms

Norms are generalization of the *size* or *distance* concept in vector spaces. Having a defined norm, a vector space becomes the normed vector space.

Definition 2.7 (Norm of a vector) *Consider the vector space $\mathbb{R}^n$, and a constant vector $x = [x_1, \dots, x_n]^T \in \mathbb{R}^n$. The p -norm in $\mathbb{R}^n$, denoted by $\|\cdot\|_p$, is a function defined by,*

$$\|x\|_p \stackrel{def}{=} (|x_1|^p + \dots + |x_n|^p)^{1/p} \quad , \quad p \in [1, \infty) \quad (2.10)$$

In particular, the 2–norm, defined as,

$$\|x\|_2 \stackrel{def}{=} \sqrt{|x_1|^2 + \cdots + |x_n|^2} \quad (2.11)$$

is the so-called *Euclidean norm*. Also, the ∞ –norm is defined as,

$$\|x\|_\infty \stackrel{def}{=} \max_i |x_i| \quad , \quad i = 1, \dots, n \quad (2.12)$$

Throughout this thesis, whenever we use the norm without subscript as $\|\cdot\|$, it will mean the 2–norm.

2.5. Signal Norms and Signal Spaces

Vector signals are time varying vector functions, defined for nonnegative time as $x(t) : \{0\} \cup \mathbb{R}_+ \mapsto \mathbb{R}^n$, where t denotes the time. Consider a vector signal $x(t) = [x_1(t), \dots, x_2(t)]^T \in \mathbb{R}^n$. At any time t , $x(t) = [x_1(t), \dots, x_2(t)]^T$ is a vector. As t changes, $x(t)$ represents a vector field. We define the vector norm of a vector signal similar to (2.10) as,

$$\|x(t)\|_p \stackrel{def}{=} (|x_1(t)|^p + \cdots + |x_n(t)|^p)^{1/p} \quad , \quad p \in [1, \infty) \quad (2.13)$$

We have defined the norms for a constant vector and a time varying vector (i.e. a vector signal). We now extend the norm definition to continuous-time vector signals.

Definition 2.8 (Norm of a signal) *For a continuous-time vector signal $x(t) \in \mathbb{R}^n$, its p –norm, denoted by $\|x(\cdot)\|_p$, in the signal space $\mathcal{L}_p$ is defined as,*

$$\|x(\cdot)\|_p \stackrel{def}{=} \left(\int_0^\infty \|x(t)\|^p dt \right)^{\frac{1}{p}} = \left(\int_0^\infty (|x_1(t)|^p + \cdots + |x_n(t)|^p) dt \right)^{\frac{1}{p}} \quad , \quad p \in [1, \infty) \quad (2.14)$$

In particular, the 2–norm and the ∞ –norm of signals are of special interest. These are defined as,

$$\|x(\cdot)\|_2 \stackrel{def}{=} \sqrt{\int_0^\infty (x_1^2(t) + \cdots + x_n^2(t)) dt} \quad (2.15)$$

$$\|x(\cdot)\|_\infty \stackrel{def}{=} \sup_{t \geq 0} \|x(t)\|_\infty = \sup_{t \geq 0} \max_{1 \leq i \leq n} |x_i(t)| \quad (2.16)$$

From now on, we will drop the term *vector* and simply say *signal* to mean vector signal.

Definition 2.9 (Signal space) *A functional space over which the signal norm exists is called signal space, and defined as,*

$$\mathcal{L}_p \stackrel{def}{=} \{x(t) \in \mathbb{R}^n : \|x(\cdot)\|_p < \infty\} \quad , \quad p \in [1, \infty) \quad (2.17)$$

We say that a signal $x(t) \in \mathcal{L}_p$ if $\|x(t)\|_p$ exists $\forall t \geq 0$. Especially $\mathcal{L}_2$ and $\mathcal{L}_\infty$ norms will be widely used in the study. A signal $x(t)$ is *bounded* if $x(t) \in \mathcal{L}_\infty$, that is, $\|x(t)\|_\infty \leq K$, $\forall t \geq 0$, for some $K > 0$.

Definition 2.10 (Extended $\mathcal{L}_p$ space) *The truncated function $f_s(t)$ is defined as,*

$$f_s = \begin{cases} f(t) & , \quad t \leq s \\ 0 & , \quad t > s \end{cases} \quad (2.18)$$

Then, using this truncated function notation, the extended $\mathcal{L}_p$ spaces are defined by,

$$\mathcal{L}_{pe} \stackrel{def}{=} \{f : \forall s < \infty, \quad f_s \in \mathcal{L}_p\} \quad (2.19)$$

Consequently, the extended $\mathcal{L}_\infty$ space is denoted by $\mathcal{L}_{\infty e}$.

Definition 2.11 (Truncated infinity norm) *The truncated ∞ -norm ($\mathcal{L}_{\infty e}$ norm) is defined as,*

$$\|x\|_t \stackrel{def}{=} \sup_{\tau \leq t} |x(\tau)| \quad (2.20)$$

Throughout the thesis, the norm $\|\cdot\|_t$ will refer to the $\mathcal{L}_{\infty e}$ norm.

2.6. Signal Convergence Lemmas

In adaptive control, a key task is to show that certain error signals, such as tracking error or state prediction error, converge to zero as time t goes to infinity. This is called *asymptotic convergence*. Such a convergence property is based on the properties of the signal and its derivative. The following lemmas state the requirements for the signal convergence. Since these are general lemmas, they will be given without proof. The proofs can be found in textbooks [6, 37, 38, 39].

Definition 2.12 (Uniformly continuous signal) *A signal $x(t)$ is uniformly continuous if for every $\epsilon > 0$ there exists a $\delta(\epsilon)$ such that,*

$$|t - \tau| < \delta \quad \Rightarrow \quad |x(t) - x(\tau)| < \epsilon \quad , \quad \forall t, \tau \geq 0 \quad (2.21)$$

Note that Lipschitz continuity given in Definition 2.6 is more restrictive than the uniform continuity of functions. Hence, if f is Lipschitz continuous in x , it is uniformly continuous in x , but the converse is not necessarily true. On the other hand, if f has bounded partial derivatives in x , then it is Lipschitz, but the converse is not necessarily true.

Lemma 2.1 (Barbalat) [39] *If a scalar function $x(t)$ is uniformly continuous such that $\lim_{t \rightarrow \infty} \int_0^t x(\tau) d\tau$ exists and is finite, then $\lim_{t \rightarrow \infty} x(t) = 0$.*

Lemma 2.2 [39] *If $x(t) \in \mathcal{L}_p$, $0 < p < \infty$, and $\dot{x}(t) \in \mathcal{L}_\infty$, then $\lim_{t \rightarrow \infty} x(t) = 0$.*

Either Lemma 2.1 or Lemma 2.2 can be used to draw the asymptotic convergence conclusion, that is, $\lim_{t \rightarrow \infty} x(t) = 0$. However, in an adaptive control system, the analyzed signal using these lemmas is generally a certain error signal $e(t)$. Besides the asymptotic convergence, it is always desired that the error signal must be bounded too, that is, $e(t) \in \mathcal{L}_\infty$. Additionally, as the quadratic form of the error signal is used in the stability analysis very frequently, it is more convenient to show that $e(t) \in \mathcal{L}_2$ rather than trying to prove that $e(t) \in \mathcal{L}_p$ for any other p value. Therefore, the above lemmas can be specialized as the following lemma which is more popular in the adaptive control literature.

Lemma 2.3 [39] *If $e(t) \in \mathcal{L}_2 \cap \mathcal{L}_\infty$ and $\dot{e}(t) \in \mathcal{L}_\infty$, then $\lim_{t \rightarrow \infty} e(t) = 0$.*

It is often seen in adaptive control systems that signals are low-pass filtered through a transfer function in the form of $\frac{b}{s+a}$. With slight abuse of notation, the following lemma gives a convergence result for filtered signals.

Lemma 2.4 [39] *If $x(t) \in \mathcal{L}_p$, $1 \leq p < \infty$, then $x_f(t) = \frac{b}{s+a}x(t) \in \mathcal{L}_\infty$ and $\lim_{t \rightarrow \infty} x_f(t) = 0$ for any constants $a > 0$ and $b > 0$.*

The low-pass filter $\frac{b}{s+a}$ above is in fact a special case of a general stable, strictly proper transfer function $H(s)$. A more general result involving $H(s)$ is available below.

Lemma 2.5 [39] *If $x(t) \in \mathcal{L}_p$, $1 \leq p < \infty$, then $x_f(t) = H(s)x(t) \in \mathcal{L}_\infty$ and $\lim_{t \rightarrow \infty} x_f(t) = 0$ for any stable and strictly proper transfer function $H(s)$.*

Definition 2.13 (Regular function) *Let $x : \mathbb{R}_+ \mapsto \mathbb{R}^n$, such that $x, \dot{x} \in \mathcal{L}_{\infty e}$. Then, x*

is called *regular* (or *holomorphic*) if,

$$|\dot{x}(t)| \leq k_1 \|x\|_t + k_2 \quad , \quad \forall t \geq 0 \quad (2.22)$$

holds for some $k_1, k_2 \geq 0$.

2.7. Frobenius Theorem

Definition 2.14 (Distribution) [33] *Given an open set $\mathcal{D} \subset \mathbb{R}^n$ and smooth vector fields $f_1, f_2, \dots, f_m : \mathcal{D} \mapsto \mathbb{R}^n$, a smooth distribution, denoted by Δ , is the process of assigning the subspace*

$$\Delta \stackrel{\text{def}}{=} \text{span}\{f_1, f_2, \dots, f_m\} \quad (2.23)$$

to each point $x \in \mathcal{D}$.

The dimension of the distribution $\Delta(x)$ at a point $x \in \mathcal{D}$ is the dimension of the subspace $\Delta(x)$. That is,

$$\dim(\Delta(x)) \stackrel{\text{def}}{=} \text{rank}([f_1, f_2, \dots, f_m]) \quad (2.24)$$

Definition 2.15 (Involutive distribution) [33] *A distribution Δ is said to be involutive, if,*

$$g_1 \in \Delta \quad \text{and} \quad g_2 \in \Delta \quad \Rightarrow \quad [g_1, g_2] \in \Delta \quad (2.25)$$

Definition 2.16 (Complete integrability) [33] *A linearly independent set of vector fields $f_1, f_2, \dots, f_m$ on $\mathcal{D} \subset \mathbb{R}^n$ is said to be completely integrable, if for each $x_0 \in \mathcal{D}$ there exist a neighborhood $\mathcal{N}$ of x_0 and $(n - m)$ real-valued smooth functions*

$h_1(x), \dots, h_{n-m}(x)$ satisfying the partial differential equation

$$\frac{\partial h_j}{\partial x} f_i(x) = 0 \quad i = 1, \dots, m; \quad j = 1, \dots, n - m \quad (2.26)$$

and the gradients ∇h_i are linearly independent.

Theorem 2.1 (Frobenius) *Let $f_1, f_2, \dots, f_m$ be a set of linearly independent vector fields. The set is completely integrable if and only if it is involutive.*

Proof: The proof of this theorem can be found in several textbooks such as [33]. $\square$

2.8. Swapping Lemma

The following lemma, known as the *swapping lemma* in adaptive control literature is common in derivation of many adaptive schemes [6].

Lemma 2.6 (Swapping lemma) *Let $\phi, w : \mathbb{R}_+ \mapsto \mathbb{R}^n$ and ϕ be differentiable. Let H be a proper, rational transfer function. If H is stable, with minimal realization*

$$H = c^T(sI - A)^{-1}b + d \quad (2.27)$$

then

$$H(w^T \phi) - H(w^T) \phi = H_c(H_b(w^T) \dot{\phi}) \quad (2.28)$$

where $H_b = (sI - A)^{-1}b$ and $H_c = -c^T(sI - A)^{-1}$.

Proof: Let $[A, b, c^T, d]$ be a minimal realization of H , with $A \in \mathbb{R}^{m \times m}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^m$,

and $d \in \mathbb{R}$. Let $x : \mathbb{R}_+ \mapsto \mathbb{R}^m$, and $y_1 : \mathbb{R}_+ \mapsto \mathbb{R}$ such that,

$$\left. \begin{aligned} \dot{x} &= Ax + b(w^T \phi) \\ y_1 &= c^T x \end{aligned} \right\} \quad (2.29)$$

and $W : \mathbb{R}_+ \mapsto \mathbb{R}^{m \times n}$, $y_2 : \mathbb{R}_+ \mapsto \mathbb{R}$ such that,

$$\left. \begin{aligned} \dot{W} &= AW + bw^T \\ y_2 &= c^T W \phi \end{aligned} \right\} \quad (2.30)$$

Thus,

$$H(w^T \phi) = y_1 + d(w^T \phi) \quad (2.31)$$

and

$$H(w^T) \phi = y_2 + (dw^T) \phi \quad (2.32)$$

Since

$$\frac{d}{dt}(W \phi) = \dot{W} \phi + W \dot{\phi} = AW \phi + bw^T \phi + W \dot{\phi} \quad (2.33)$$

it follows that,

$$\left. \begin{aligned} \frac{d}{dt}(x - W \phi) &= A(x - W \phi) - W \dot{\phi} \\ y_1 - y_2 &= c^T(x - W \phi) \end{aligned} \right\} \quad (2.34)$$

The result then follows, since

$$H(w^T \phi) - H(w^T) \phi = y_1 - y_2 = H_c(W \dot{\phi}) = H_c(H_b(w^T) \dot{\phi}) \quad (2.35)$$

□

3. NONLINEAR DYNAMIC SYSTEMS

3.1. Nonlinear System Modeling

A general representation of a physical system is given as,

$$\left. \begin{aligned} \dot{x} &= f(x, u, w, t) \\ y &= h(x, w) \\ e &= k(x, w) \end{aligned} \right\} \quad (3.1)$$

where x is a vector of states, u is a vector of inputs which are used for control purposes, w is a vector of uncontrollable inputs such as external disturbances, y is a vector of outputs available for measurement, e is a vector of outputs which are desired to be controlled; f , h and k are appropriate vector fields [34].

Control systems engineering deals usually with two tasks when such a model of a physical system is available; namely, analysis and synthesis. The analysis task is essentially based on solution of the system of differential equations given by (3.1) in order to see the behavior of the dynamic model under the applied input. Stability and performance are main concerns in the analysis, where the term *performance* is used loosely as it is highly dependent on the purposes of the physical system. However, there are some control techniques that provide valuable results about the system without actually solving the differential equations, although the applicability of these techniques varies according to the system description. Generally speaking, if f , h and k describe linear relationships between their parameters and equivalents, there is a nice collection of analysis tools [40]; otherwise the tools at hand might be very limited. The lack of mathematical competence in analysis of nonlinear systems has been decreased to some extent using numerical analysis techniques and computer simulation methods. The synthesis task, on the other hand, aims to find a suitable set of control inputs $\mathcal{U}$ that drives the controlled outputs vector e to desired location. As in the analysis task, it might be extremely difficult to find such an input if f is nonlinear.

Although the nonlinear dynamic model in (3.1) can be used to describe almost all known physical systems, it is too general to develop a mathematical treatment for provable results. The efforts in nonlinear systems field have been continuing with the ultimate objective that analysis and synthesis of any given nonlinear system become possible. A slightly less general representation of the nonlinear physical systems could be written as,

$$\left. \begin{aligned} \dot{x} &= f(x, u) \\ y &= h(x) \end{aligned} \right\} \quad (3.2)$$

where the system is *autonomous* with no disturbances. The output vector e is not expressed explicitly in (3.2) but it is usually considered as error dynamics, the error being the difference between the system output and a set of desired locations $\mathcal{Y}_d$. If the desired location for the output is fixed, then the control problem is referred to as “a *regulation*” about a set point; and if the desired location for the output varies with time, it is referred to as *tracking control* or *servo control*.

A widely accepted nonlinear model is also obtained in which the control vector u enters the dynamic equation linearly as,

$$\left. \begin{aligned} \dot{x} &= f(x) + g(x)u \\ y &= h(x) \end{aligned} \right\} \quad (3.3)$$

Dynamic systems in the form of (3.3) are known as *control-affine* (or “linear in control”) nonlinear dynamic systems. Fortunately, many physical systems can be modeled as in (3.3). Besides this control-affine property, we also assume that the nonlinear system has a single input and a single output. Hence, in (3.3), x is $n \times 1$ vector of state variables, u is the control input variable, y is the controlled output variable, f is $n \times 1$ vector of nonlinear functions, g is $n \times 1$ vector of nonlinear functions and h is a scalar nonlinear function. Furthermore, $f : \mathcal{D} \mapsto \mathbb{R}^n$, $g : \mathcal{D} \mapsto \mathbb{R}^n$ and $h : \mathcal{D} \mapsto \mathbb{R}$ are sufficiently smooth in a domain $\mathcal{D} \subset \mathbb{R}^n$ [33, 34, 35].

3.2. Linearization

The most important method in control of nonlinear systems is the method of linearization. The motivation to linearize a nonlinear system is obvious, since there are established control techniques developed for the linear systems. Hence, if the nonlinear system might be somehow transformed into a linear one having same dynamical characteristics, the linear control methods such as pole placement or optimal control could be easily applied.

Linearization might be performed by a first-order approximation around a local operating point (Jacobian approach) or feedback. Feedback linearization is a more general approach and will be studied in detail in the following section. The Jacobian linearization of the nonlinear system in (3.3) is done about an equilibrium point (x_0, y_0, u_0) as,

$$\left. \begin{aligned} \dot{x} &= \left[\frac{\partial f(x_0)}{\partial x} + \frac{\partial g(x_0)}{\partial x} u_0 \right] (x - x_0) + g(x_0)(u - u_0) \\ y - y_0 &= \frac{\partial h(x_0)}{\partial x} (x - x_0) \end{aligned} \right\} \quad (3.4)$$

Then the Jacobian model in (3.4) can be put into linear state-space form as,

$$\left. \begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx \end{aligned} \right\} \quad (3.5)$$

where A is an $n \times n$ matrix, B is an $n \times 1$ vector, and C is a $1 \times n$ row vector.

3.2.1. Feedback Linearization

The first-order linearized model is an exact representation of the nonlinear model only at the linearization point (x_0, y_0, u_0) . A control strategy based on the Jacobian linearized model may yield unsatisfactory performance at other operating points. A more general method is required which would give a linear model representing the nonlinear system over a large set of operating conditions. This method is known as

feedback linearization and is based on two operations: *Nonlinear state feedback* and *nonlinear change of coordinates*. Mathematically speaking, for the system given by (3.3), the feedback linearization method studies the question of whether there exist a state feedback control $u = \alpha(x) + \beta(x)v$ and a diffeomorphism $\phi(x)$ that transform the nonlinear system into an equivalent linear one.

There are two main approaches in feedback linearization, namely, *input-state linearization* and *input-output linearization*. In the input-state linearization approach, the goal is to completely linearize the state equation. This is achieved by deriving artificial output w that yields a feedback linearized model of degree n . A linear controller is then synthesized. In the input-output linearization approach, the objective is to linearize the map between the transformed input v and the actual output y . A linear controller is then designed for the linearized input-output model.

3.2.2. Input-State Linearization

We have a nonlinear system described by,

$$\dot{x} = f(x) + g(x)u \quad (3.6)$$

where $x \in \mathbb{R}^n$ is the vector of state variables, $u \in \mathbb{R}$ is the input variable, $f, g : \mathcal{D} \subset \mathbb{R}^n \mapsto \mathbb{R}^n$ are vector of sufficiently smooth nonlinear vector functions. We first formally define the input-state linearization.

Definition 3.1 (Input-state linearization) *A nonlinear system of the form (3.6) is said to be input-state linearizable, if there exist a diffeomorphism $\phi : \mathcal{D} \subset \mathbb{R}^n \mapsto \mathbb{R}^n$ which defines the coordinate transformation*

$$\xi = \phi(x) \quad (3.7)$$

and a control law of the form

$$u = \alpha(x) + \beta(x)v \quad (3.8)$$

that transform (3.6) into a linear system of the form

$$\dot{\xi} = A\xi + Bv \quad (3.9)$$

$$\text{where } A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix} \text{ and } B = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

Linearizing the nonlinear system in (3.6) as described in Definition 3.1 may not always be possible. When such an input-state linearization is possible, the linearizing diffeomorphism ϕ in (3.7) and the linearizing control u in (3.8) may not be unique. The following theorem states the input-state linearizability conditions for nonlinear systems in the form of (3.6).

Theorem 3.1 *The system in (3.6) is input-state linearizable on $\mathcal{D}_0 \subset \mathcal{D}$, if and only if the following conditions are satisfied:*

1. *The vector fields $\{g, \text{ad}_f g, \dots, \text{ad}_f^{n-1} g\}$ are linearly independent in $\mathcal{D}_0$. This condition is equivalent to saying that the $n \times n$ matrix*

$$C = \begin{bmatrix} g(x) & \text{ad}_f g(x) & \cdots & \text{ad}_f^{n-1} g(x) \end{bmatrix}$$

has rank n for all $x \in \mathcal{D}_0$.

2. *The distribution $\Delta = \text{span}\{g, \text{ad}_f g, \dots, \text{ad}_f^{n-2} g\}$ is involutive in $\mathcal{D}_0$.*

Proof: The proof of this theorem can be found in [33]. $\square$

If both conditions of the Theorem 3.1 hold, then there exists a smooth scalar function $\psi(x) : \mathcal{D} \mapsto \mathbb{R}$ such that

1. $L_g L_f^i \psi(x) = 0 \quad , \quad i = 0, \dots, n-2$
2. $L_g L_f^{n-1} \psi(x) \neq 0$

are satisfied. The function ψ is sometimes called as the *fictitious output* function. The linearized states are found by,

$$\xi_i = L_f^{i-1} \psi(x) \quad , \quad i = 1, \dots, n \quad (3.10)$$

and $\alpha(x)$ and $\beta(x)$ in the state feedback control law given in (3.8) are obtained as,

$$\alpha(x) = -\frac{L_f^n \psi(x)}{L_g L_f^{n-1} \psi(x)} \quad (3.11)$$

and

$$\beta(x) = \frac{1}{L_g L_f^{n-1} \psi(x)} \quad (3.12)$$

3.2.3. Input-Output Linearization

Consider the nonlinear system given in (3.3). The objective of the input-output linearization is to find a linear relationship between the system output y and the transformed input variable v . This type of linearization is especially important when an output tracking problem is of interest. Note that in the input-state linearization, we have not defined an output for the nonlinear system in (3.6). One might think that an output could have been defined as a function of the states and might expect a linear relationship between the input and the output. However, linearizing the state

equation via input-state linearization does not necessarily imply that the resulting map from the input to the output is linear. The reason for this is that, in the derivation of the coordinate transformation which is used to linearize the state equation, the output equation, hence the possible nonlinearities in it, are not taken into account. In the input-output linearization, the output equation is known from the beginning, and the linearizing diffeomorphism is determined using the output equation. Next, we define the *relative degree* of a nonlinear system.

Definition 3.2 (Relative degree) *The n^{th} -order single-input single-output nonlinear system given in (3.3) is said to have relative degree γ in a region $\mathcal{D}_0 \subset \mathcal{D}$ if*

1. $L_g L_f^i h(x) = 0$, $\quad \forall x \in \mathcal{D}_0, \quad i = 0, \dots, \gamma - 2$
2. $L_g L_f^{\gamma-1} h(x) \neq 0$, $\quad \forall x \in \mathcal{D}_0$

Basically, the relative degree γ is the number of times the output $y = h(x)$ has to be differentiated before the input u appears in its expression. Relative degree may not always be defined, especially when the output is not affected by the input. If the relative degree γ exists, then $\gamma \leq n$, where n is the order of the nonlinear system. When $\gamma = n$, then the input-output linearization leads to the input-state linearization. In fact, this is a preferred case, whenever possible.

If $\gamma < n$, then there exist some parts of the nonlinear dynamics which is unobservable and cannot be linearized. This unobservable part of the dynamics is called the *internal dynamics*. The following theorem states the conditions for input-output linearizability in the case of $\gamma < n$.

Theorem 3.2 *If the system in (3.3) has relative degree $\gamma \leq n$, $\forall x \in \mathcal{D}_0 \subset \mathcal{D}$, then for every $x_0 \in \mathcal{D}_0$, there exists a neighborhood Ω of x_0 and smooth vector function $\eta(x) = \begin{bmatrix} \eta_1 & \dots & \eta_{n-\gamma} \end{bmatrix}^T$ such that*

1. $L_g \eta_i(x) = 0$, $\quad i = 1, \dots, n - \gamma, \quad \forall x \in \Omega$

$$2. \phi(x) = \begin{bmatrix} \xi(x) \\ \eta(x) \end{bmatrix} \text{ is a diffeomorphism on } \Omega, \text{ where } \xi(x) = \begin{bmatrix} h(x) \\ L_f h(x) \\ \vdots \\ L_f^{\gamma-1} h(x) \end{bmatrix}.$$

Proof: The proof of this theorem can be found in [33]. □

As it is seen in the above theorem, the coordinate transformation function $\phi(x)$ consists of two parts, one for the linearized part and one for the nonlinear part of the system. Applying this coordinate transformation to the nonlinear system in (3.3) transforms it into the so-called *normal form* as,

$$\left. \begin{aligned} \dot{\xi}_1 &= \xi_2 \\ \dot{\xi}_2 &= \xi_3 \\ &\vdots \\ \dot{\xi}_\gamma &= b(\xi, \eta) + a(\xi, \eta)u \\ \dot{\eta} &= q(\xi, \eta) \\ y &= \xi_1 \end{aligned} \right\} \quad (3.13)$$

where $b(\xi, \eta) \stackrel{def}{=} L_f^\gamma h(x)$ and $a(\xi, \eta) \stackrel{def}{=} L_g L_f^{\gamma-1} h(x)$. The $\dot{\eta} = q(\xi, \eta)$ part of the dynamics in the normal form given in (3.13) represents the internal dynamics. Stability of the internal dynamics is of great importance, which will be explained in the sequel.

The feedback control

$$u = \frac{v - b(\xi, \eta)}{a(\xi, \eta)} \quad (3.14)$$

transforms the $\dot{\xi}_\gamma$ equation in (3.13) to $\dot{\xi}_\gamma = v$. Hence, the mapping between the transformed input v and the output y is exactly linear, and a linear state feedback controller, such as pole placement, can be designed to stabilize the ξ subsystem. The

feedback control law in (3.14) can be written in the original coordinates as,

$$u = \frac{v - L_f^\gamma h(x)}{L_g L_f^{\gamma-1} h(x)} \quad (3.15)$$

If the pole placement design is applied, the transformed input is expressed as,

$$v = \lambda_1 L_f^{\gamma-1} h(x) + \lambda_2 L_f^{\gamma-2} h(x) + \cdots + \lambda_\gamma h(x) \quad (3.16)$$

where $\lambda_1, \dots, \lambda_\gamma \in \mathbb{R}_+$. Substituting (3.16) in (3.15),

$$u = \frac{-L_f^\gamma h(x) + \lambda_1 L_f^{\gamma-1} h(x) + \cdots + \lambda_\gamma h(x)}{L_g L_f^{\gamma-1} h(x)} \quad (3.17)$$

defines the complete state feedback control law. If (3.17) is applied to the system in the normal form, (3.13) can be written as,

$$\left. \begin{aligned} \dot{\xi} &= A\xi \\ \dot{\eta} &= q(\xi, \eta) \\ y &= \xi_1 \end{aligned} \right\} \quad (3.18)$$

The state matrix A of the linearized system can be designed through the transformed control input v in such a way that state variables of the ξ subsystem converge to zero. In this case, the second equation of (3.18) becomes

$$\dot{\eta} = q(0, \eta) \quad (3.19)$$

Equation (3.19) is called the *zero-dynamics* of the nonlinear system given in (3.3). Since the zero-dynamics is unobservable part of the nonlinear system from the output, its asymptotic stability is of great importance. This is expressed in the following definition [41].

Definition 3.3 (Minimum-phase) *Assuming that $x = 0$ is an equilibrium point of the*

system in (3.3) and $h(0) = 0$, the dynamics $\dot{\eta} = q(0, \eta)$ is referred to as the zero-dynamics of the system. A system is said to be asymptotically minimum-phase if it has asymptotically stable zero-dynamics.

Definition 3.4 (Exponentially minimum-phase) *Referring to Definition 3.3, a system is said to be exponentially minimum-phase if it has exponentially stable zero-dynamics.*

Lemma 3.1 *Asymptotic stability of the zero dynamics is a necessary and sufficient condition for asymptotic stability of the feedback linearized system in (3.18)*

Proof: See [34].

3.2.4. Canonical Forms Used in Nonlinear Adaptive Control

The nonlinear system given in (3.3) is too general, and not appropriate for direct mathematical treatment. For that reason, it is first usually put into a more appropriate form, if this is possible, such as the normal form as given in (3.13). If such a transformation is not possible, then the nonlinear system subject to study is usually restricted to be already in one of the canonical forms. In this section, we outline the canonical forms which are frequently used in application of adaptive control to nonlinear systems [37, 42].

3.2.4.1. Brunovsky Canonical Form. Brunovsky canonical form has a quite simple structure. The nonlinearity and the control input appear only in the n^{th} -equation. However, very few systems can be put into this form. The general appearance of a

system in this form is given as,

$$\left. \begin{aligned} \dot{\xi}_1 &= \xi_2 \\ \dot{\xi}_2 &= \xi_3 \\ &\vdots \\ \dot{\xi}_{n-1} &= \xi_n \\ \dot{\xi}_n &= u + f(\xi_1, \xi_2, \xi_3, \dots, \xi_n) \end{aligned} \right\} \quad (3.20)$$

where $\xi_i \in \mathbb{R}$, $i = 1, \dots, n$; $u \in \mathbb{R}$ and $f : \mathbb{R}^n \mapsto \mathbb{R}$.

3.2.4.2. Strict-Feedback Form. This form is one of the forms which are said to be in *triangular* structure. The strict-feedback form is defined as,

$$\left. \begin{aligned} \dot{x} &= f(x) + g(x)\xi_1 \\ \dot{\xi}_1 &= f_1(x, \xi_1) + g_1(x, \xi_1)\xi_2 \\ \dot{\xi}_2 &= f_2(x, \xi_1, \xi_2) + g_2(x, \xi_1, \xi_2)\xi_3 \\ &\vdots \\ \dot{\xi}_{k-1} &= f_{k-1}(x, \xi_1, \dots, \xi_{k-1}) + g_{k-1}(x, \xi_1, \dots, \xi_{k-1})\xi_k \\ \dot{\xi}_k &= f_k(x, \xi_1, \dots, \xi_k) + g_k(x, \xi_1, \dots, \xi_k)u \end{aligned} \right\} \quad (3.21)$$

where $x \in \mathbb{R}^n$; $f, g : \mathbb{R}^n \mapsto \mathbb{R}^n$; $\xi_i \in \mathbb{R}$; $f_i, g_i : \mathbb{R}^{n+i} \mapsto \mathbb{R}$; $i = 1, \dots, k$ and $u \in \mathbb{R}$. In this form, there are k affine subsystems in terms of ξ , separating the input u from the dynamics of x . The reason for referring to the ξ -subsystem as 'strict-feedback' is that the nonlinearities f_i, g_i in the $\dot{\xi}_i$ equation depend only on $x, \xi_1, \dots, \xi_i$, that is, on state variables that are 'fed back'.

3.2.4.3. Pure-Feedback Form. A more general class of *triangular* systems than the strict-feedback systems are the systems in pure-feedback form:

$$\left. \begin{aligned} \dot{x} &= f(x) + g(x)\xi_1 \\ \dot{\xi}_1 &= f_1(x, \xi_1, \xi_2) \\ \dot{\xi}_2 &= f_2(x, \xi_1, \xi_2, \xi_3) \\ &\vdots \\ \dot{\xi}_{k-1} &= f_{k-1}(x, \xi_1, \dots, \xi_k) \\ \dot{\xi}_k &= f_k(x, \xi_1, \dots, \xi_k, u) \end{aligned} \right\} \quad (3.22)$$

where $x \in \mathbb{R}^n$; $f, g : \mathbb{R}^n \mapsto \mathbb{R}^n$; $\xi_i \in \mathbb{R}$; $f_i : \mathbb{R}^{n+i+1} \mapsto \mathbb{R}$; $i = 1, \dots, k$ and $u \in \mathbb{R}$. The systems in pure-feedback form are not restricted to the affine structure of the strict-feedback class in the ξ -subsystem.

3.2.4.4. The Normal Form. The general structure of a system of n^{th} -order in the normal form, having a relative degree $\gamma < n$, is

$$\left. \begin{aligned} \dot{\xi}_1 &= \xi_2 \\ \dot{\xi}_2 &= \xi_3 \\ &\vdots \\ \dot{\xi}_\gamma &= b(\xi, \eta) + a(\xi, \eta)u \\ \dot{\eta} &= q(\xi, \eta) \\ y &= \xi_1 \end{aligned} \right\} \quad (3.23)$$

where $\xi_i \in \mathbb{R}$, $i = 1, \dots, \gamma$; $a, b : \mathbb{R}^n \mapsto \mathbb{R}$; $u \in \mathbb{R}$; $\eta \in \mathbb{R}^{n-\gamma}$, $q : \mathbb{R}^n \mapsto \mathbb{R}^{n-\gamma}$ and $y \in \mathbb{R}$, as given in Section 3.2.3. Transformation conditions for the system in (3.3) into the normal form were stated in Theorem 3.2. The normal form representation will be in the core of the derivation of the adaptive scheme in this study, so it will be referred to in the following chapters.

Representation of nonlinear systems in one of these canonical forms is very important. For example, the backstepping design deals with the systems in strict-feedback

form [14, 37].

3.3. Uncertainty

Mathematical modeling is the central point in almost all control system design methods. However, no physical system can be perfectly modeled and there always exists some mismatch between the model and the physical system. Even if an exact model of the physical system is available, one type of uncertainty may still arise from some exogenous disturbances which might affect the control system via imperfect measurement of state or output [43]. On the other hand, other types of uncertainties may be present due to the lack of knowledge about the physical system. Even if the model is qualitatively correct, i.e. the structure of the model is exact, the actual parameter values of the physical system could hardly be obtained exactly. Therefore, both *qualitative* (structural) and *quantitative* (parametric) uncertainties will always be present when trying to model any physical system [6].

Robustness of a control system determines the capability of the controller to be unaffected by the uncertainties. Since uncertainties will always be present, robustness is always a desired property when designing control systems. The robustness issues will be discussed in the following sections.

3.3.1. Exogenous Uncertainty and Disturbance Decoupling Problem

Uncontrollable inputs give rise to exogenous uncertainty in the model. These are considered as disturbances since they are uncontrollable. The model of the physical system in (3.3) can be written as,

$$\left. \begin{aligned} \dot{x} &= f(x) + g(x)u + d(x)w \\ y &= h(x) \end{aligned} \right\} \quad (3.24)$$

where $d : \mathcal{D} \subset \mathbb{R}^n \mapsto \mathbb{R}^n$ is a known smooth vector field and $w \in \mathbb{R}$ is stochastic disturbance. The robust control problem associated with this kind of uncertainty is

usually called as the *disturbance decoupling* or *disturbance rejection* problem. When considered with input-output linearization of nonlinear systems, the disturbance rejection problem is to find a diffeomorphism $\phi(x)$ and a nonlinear static state feedback control law $u = \alpha(x) + \beta(x)v$ such that:

1. The map between the transformed input v and the output y is linear,
2. The output y is completely unaffected by the disturbance w .

The system in (3.24) can be transformed into the normal form using the diffeomorphism

$$\phi(x) = \begin{bmatrix} \xi \\ \eta \end{bmatrix} \text{ as,} \quad \left. \begin{aligned} \dot{\xi}_1 &= \xi_2 \\ \dot{\xi}_2 &= \xi_3 \\ &\vdots \\ \dot{\xi}_\gamma &= b(\xi, \eta) + a(\xi, \eta)u + s(\xi, \eta)w \\ \dot{\eta} &= q(\xi, \eta) + r(\xi, \eta)w \\ y &= \xi_1 \end{aligned} \right\} \quad (3.25)$$

where $s(\xi, \eta) \stackrel{\text{def}}{=} L_d L_f^{\rho-1} h(x)$ and $r_i(\xi, \eta) \stackrel{\text{def}}{=} L_d \eta_i(x)$ for $i = 1, \dots, n - \gamma$, and ρ is the relative degree of the disturbance w . Definition of the relative degree for the disturbance is analogous to the relative degree γ associated with the control input u , which was given in Definition 3.2. The disturbance w is said to have relative degree ρ in a region $\mathcal{D}_0 \subset \mathcal{D}$ if,

1. $L_d L_f^i h(x) = 0$, $\forall x \in \mathcal{D}_0$, $i = 0, \dots, \rho - 1$
2. $L_d L_f^{\rho-1} h(x) \neq 0$, $\forall x \in \mathcal{D}_0$

The following theorem states the solvability of the disturbance rejection problem:

Theorem 3.3 [35] *The disturbance rejection problem for the system in (3.24) is solv-*

able if and only if

$$L_d L_f^i h(x) = 0 \quad , \quad i = 0, \dots, \gamma - 1 \quad (3.26)$$

Proof: The conclusion can be directly drawn from the normal form given in (3.25). The term $s(\xi, \eta)$ must be zero in order for the output y to be completely unaffected by the disturbance w . This is possible only if the condition in (3.26) is satisfied [35]. $\square$

Theorem 3.3 is an expression of the fact that the relative degree of the disturbance to the output must be strictly greater than the relative degree of the unperturbed nonlinear system, that is, $\rho > \gamma$. This condition is called the *disturbance matching condition*. There are actually three cases according to relative degrees of the disturbance and the unperturbed system:

1. If $\rho > \gamma$, the output can be completely decoupled from the disturbance, as stated above.
2. If $\rho = \gamma$, the control law given in (3.14) does not decouple the output from the disturbance. However, if it is possible to measure the disturbance, the disturbance rejection problem is solvable with the following control law:

$$u = \frac{v - b(\xi, \eta) - s(\xi, \eta)w}{a(\xi, \eta)} \quad (3.27)$$

Then the conditions given in Theorem 3.3 can be modified as,

$$L_d L_f^i h(x) = 0 \quad , \quad i = 0, \dots, \gamma - 2 \quad (3.28)$$

3. If $\rho < \gamma$, then the disturbance affects the output more directly than does the control input and disturbance decoupling cannot be achieved with a static state feedback control law.

In reality, the disturbances are seldom measurable. Hence the conditions in (3.26) seem to be more applicable. Another important point is that the decoupling is achieved only between ξ variables of the normal form in (3.25) and the disturbance. The η variables associated with zero dynamics may be perturbed and furthermore the zero dynamics may be unstable due to the disturbance.

3.3.2. Structural Uncertainty

Structural uncertainty stems from inadequate, hence inaccurate, qualitative information about the physical system which is to be modeled. Consider the physical system given by (3.3) and assume that the uncertainty exists on the vector fields f and g . This type of uncertainty may be modeled as perturbations and the system can be expressed as,

$$\left. \begin{aligned} \dot{x} &= f(x) + \Delta f(x) + (g(x) + \Delta g(x))u \\ y &= h(x) \end{aligned} \right\} \quad (3.29)$$

where f and g are definite parts of the dynamics, and Δf and Δg represent uncertain parts associated with f and g , respectively, which can also be viewed as perturbations.

Proposition 3.1 [35] *Let the uncertain system in (3.29) be the perturbed form of the nominal system in (3.3) and assume the relative degree of the nominal system is γ . If the perturbations Δf and Δg satisfy*

$$\begin{aligned} L_{\Delta f} L_f^i h &= 0 \quad , \quad i = 0, \dots, \gamma - 1 \\ L_{\Delta g} L_f^i h &= 0 \quad , \quad i = 0, \dots, \gamma - 1 \end{aligned} \quad (3.30)$$

then the linearizing control law for the nominal system also linearizes the perturbed system.

Proof: In order for the linearizing control law for the system in (3.15) to linearize also the uncertain system in (3.29), the linearized parts (i.e. the ξ subsystems) of the

normal form representations of (3.6) and (3.29) must be same. We linearize the system in (3.29) using the transformations outlined in Section 3.2.3 as,

$$\left. \begin{aligned} \dot{\xi}_1 &= \xi_2 + L_{\Delta f}h + L_{\Delta g}hu \\ \dot{\xi}_2 &= \xi_3 + L_{\Delta f}L_f h + L_{\Delta g}L_f hu \\ &\vdots \\ \dot{\xi}_\gamma &= b(\xi, \eta) + a(\xi, \eta)u + L_{\Delta f}L_f^{\gamma-1}h + L_{\Delta g}L_f^{\gamma-1}hu \\ \dot{\eta} &= q(\xi, \eta) + L_{\Delta f}\eta + L_{\Delta g}\eta u \\ y &= \xi_1 \end{aligned} \right\} \quad (3.31)$$

It is seen that if $L_{\Delta f}L_f^i h = 0$ and $L_{\Delta g}L_f^i h = 0$ for all $i = 0, \dots, \gamma - 1$, then the uncertainty is confined in the zero dynamics only. Hence the control law in (3.15) provides a linear relationship between the transformed input v and the output y for the uncertain system in (3.29) as well. $\square$

Definition 3.5 (Matching conditions) *For the uncertain system given in (3.29), the conditions given in (3.30) are called the matching conditions.*

The matching conditions can also be expressed as,

$$\Delta f, \Delta g \in \ker \begin{bmatrix} dh \\ dL_f h \\ \vdots \\ dL_f^{\gamma-1} h \end{bmatrix} \quad (3.32)$$

Definition 3.6 (Generalized matching conditions) *For the uncertain system in (3.29), the conditions*

$$\begin{aligned} L_{\Delta f}L_f^i h &= 0 \quad , \quad i = 0, \dots, \gamma - 2 \\ L_{\Delta g}L_f^i h &= 0 \quad , \quad i = 0, \dots, \gamma - 1 \end{aligned} \quad (3.33)$$

are called the generalized matching conditions.

The generalized matching conditions are also known as the *extended matching conditions* and are weaker than those in (3.30), because the condition $L_{\Delta f} L_f^{\gamma-1} h = 0$ is no longer required [35]. The generalized matching conditions can also be expressed as,

$$\Delta f \in \ker \begin{bmatrix} dh \\ dL_f h \\ \vdots \\ dL_f^{\gamma-2} h \end{bmatrix}, \quad \Delta g \in \ker \begin{bmatrix} dh \\ dL_f h \\ \vdots \\ dL_f^{\gamma-1} h \end{bmatrix} \quad (3.34)$$

Under the generalized matching conditions, the γ^{th} -equation in the normal form given in (3.31) becomes

$$\dot{\xi}_\gamma = b(\xi, \eta) + a(\xi, \eta)u + L_{\Delta f} L_f^{\gamma-1} h \quad (3.35)$$

where the term $L_{\Delta f} L_f^{\gamma-1} h$ must be accounted for by the control u . This might be accomplished either by the sliding-mode control approach, or by incorporating the derivative of the uncertainty, if such a derivative is available.

3.3.3. Parametric Uncertainty

In case of parametric uncertainty, the nonlinear system in (3.3) is expressed as,

$$\left. \begin{aligned} \dot{x} &= f(x, \theta) + g(x, \theta)u \\ y &= h(x) \end{aligned} \right\} \quad (3.36)$$

where $\theta \in \mathbb{R}^p$ is the vector of unknown parameters. The vector fields f and g are now defined as $f, g : \mathcal{D}_x \subset \mathbb{R}^{n \times p} \mapsto \mathbb{R}^n$. The system dynamics might be a general nonlinear function of the parameter vector θ as above, or it might be *linear* in θ , which is an important property from adaptive control point of view. If the dynamics is linear in terms of the unknown parameter vector θ , then the vector fields f and g can be

expressed as,

$$f(x, \theta) = \sum_{k=1}^p \theta_k f_k(x) \quad (3.37)$$

$$g(x, \theta) = \sum_{k=1}^p \theta_k g_k(x) \quad (3.38)$$

where $f_k, g_k : \mathcal{D}_x \subset \mathbb{R}^n \mapsto \mathbb{R}^n$. The nonlinear system given in (3.36) with parametric linearities in (3.37) and (3.38) is the subject of nonlinear adaptive control research. The elements of the θ vector are unknown but constant parameters of the system. The term constant here is used somewhat loosely. It can be stated more precisely that the actual system parameters must be constant over a sufficiently large period of time within the operation of the system. When an adaptive controller is used, it is a necessary condition that θ does not change faster than the adaptation speed of the controller [5, 6, 41].

It is also possible that the uncertain system described in (3.36), (3.37) and (3.38) can be put into a form similar to (3.29) and disturbance rejection methods can be applied as long as matching conditions hold. For (3.36), however, it is more favorable to apply a methodology that estimates the parameter vector θ , denoted by $\hat{\theta}$, which leads to the convergence to the actual θ as time goes to infinity. In fact, matching conditions are necessary conditions for some schemes that are widely used in nonlinear adaptive control (as will be seen in the following chapter).

3.4. Summary

There are various representations to mathematically model nonlinear dynamic systems. An important class is the so called the “control-affine” nonlinear systems. Although there is no general framework to control a nonlinear system, the linearization via state feedback is a very important initial step. Under certain conditions, the nonlinear system can be put into a canonical form using the state feedback lineariza-

tion. The canonical forms are useful to be able to apply well known techniques of linear control. The most important canonical form which is used extensively in this thesis is the normal form.

In most cases, the exact model of a nonlinear physical system is not available. The uncertainty in the model can be expressed in several different ways. From the adaptive control point of view, the parametric uncertainties will be considered. When there are unknown parameters in the system dynamics, it is possible to build an adaptive system including an update dynamics for the unknown parameters.

4. PROBLEM STATEMENT

4.1. System Description

We consider a class of nonlinear systems which can be described by,

$$\left. \begin{aligned} \dot{x} &= f(x, \theta) + g(x, \theta)u \\ y &= h(x) \end{aligned} \right\} \quad (4.1)$$

where $x \in \mathbb{R}^n$ is the state vector, $f, g : \mathcal{D}_x \subset \mathbb{R}^{n \times p} \mapsto \mathbb{R}^n$ are sufficiently smooth vector fields, $h : \mathcal{D} \subset \mathbb{R}^n \mapsto \mathbb{R}$ is a sufficiently smooth scalar function, $\theta \in \mathcal{S}$ is the unknown parameter vector with $\mathcal{S} \subset \mathbb{R}^p$ being a compact set, $u \in \mathbb{R}$ is the input and $y \in \mathbb{R}$ is the output.

Assumption 1 *The system in (4.1) has strong relative degree γ for all $x \in \mathbb{R}^n$ and $\theta \in \mathcal{S}$, according to the Definition 3.2.*

Assumption 2 *The state vector x is available for feedback, and the system is input-output linearizable as stated in Theorem 3.2.*

As described in Chapter 3, the linearizing diffeomorphism is found in a neighborhood Ω of x , $\forall x \in \mathcal{D}$ and $\forall \theta \in \mathcal{S}$, as $\phi(x, \theta) = \begin{bmatrix} \xi(x, \theta) & \eta(x, \theta) \end{bmatrix}^T$ where

$$\xi(x, \theta) = \begin{bmatrix} h(x) \\ L_{f(x, \theta)} h(x) \\ \vdots \\ L_{f(x, \theta)}^{\gamma-1} h(x) \end{bmatrix} \quad (4.2)$$

and

$$\eta(x, \theta) = \begin{bmatrix} \eta_1(x, \theta) \\ \vdots \\ \eta_{n-\gamma}(x, \theta) \end{bmatrix}, \quad L_g \eta_i(x, \theta) = 0, \quad i = 1, \dots, n - \gamma \quad (4.3)$$

The diffeomorphism $\phi(x, \theta)$ transforms the nonlinear system in (4.1) into the normal form as,

$$\left. \begin{aligned} \dot{\xi}_1 &= \xi_2 \\ \dot{\xi}_2 &= \xi_3 \\ &\vdots \\ \dot{\xi}_\gamma &= L_f^\gamma h(x)(x, \theta) + L_g L_f^{\gamma-1} h(x)(x, \theta) u \\ \dot{\eta}_1 &= q_1(\xi, \eta) \\ &\vdots \\ \dot{\eta}_{n-\gamma} &= q_{n-\gamma}(\xi, \eta) \\ y &= \xi_1 \end{aligned} \right\} \quad (4.4)$$

The second part of (4.4), known as the *internal dynamics*, can be written in vector form as,

$$\dot{\eta} = q(\xi, \eta) \quad (4.5)$$

With $\xi = 0$, the internal dynamics becomes

$$\dot{\eta} = q(0, \eta) \quad (4.6)$$

which is referred to as the *zero-dynamics* of the system in (4.1).

Assumption 3 *The system defined in (4.1) is exponentially minimum-phase, and all the nonlinearities in the internal-dynamics are globally Lipschitz.*

As given in Definition 3.4, the system in (4.1) is exponentially minimum-phase, if its zero-dynamics given in (4.6) is exponentially stable. The global Lipschitz condition for the nonlinearities is defined in Definition 2.6.

Assumption 4 *The system dynamics is linear in the unknown parameters. The vector fields f and g in (4.1) can be expressed as,*

$$f(x, \theta) = \sum_{k=1}^p \theta_k f_k(x) \quad (4.7)$$

$$g(x, \theta) = \sum_{k=1}^p \theta_k g_k(x) \quad (4.8)$$

where $f_k, g_k : \mathcal{D}_x \subset \mathbb{R}^n \mapsto \mathbb{R}^n$. Without loss of generality, the elements θ_k of the parameter vector θ are the same for both vector fields f and g , as defined in (4.7) and (4.8).

As a technical remark, the above assumption modifies the definition of the vector fields f and g as $f, g : \mathcal{D} \subset \mathbb{R}^n \mapsto \mathbb{R}^n$ (cf. Equation (4.1)).

The main objective of this study is to adaptively control the nonlinear system defined in (4.1) under the assumptions Assumption 1-4, in such a way that the following conditions are achieved by the overall control system:

1. the overall nonlinear system must be BIBS and BIBO stable.
2. the initial transient error and the transients under large parameter changes in the system dynamics will be improved.

The term *improvement* of the transient response is used in different studies to express that the peak values of the transient error are reduced and/or the convergence time of the transient error is shortened. In the following section, we present a method

to evaluate the transient error quantitatively.

4.2. Performance in an Adaptive System

It is seen in almost all related studies on adaptive control that the transient response enhancement is asserted qualitatively. Although the performance is usually observed in terms of the number of peaks in the output signals, no formal quantitative description of superiority or inferiority of transient performance of an adaptive system has been given yet. The reason for this might be the fact that several such descriptions can be proposed depending on the application. For example, the convergence time of the transient error might be very important in an application, whereas the magnitude of the peaks might be critical in another application. Moreover, some applications may pay attention to the sign of the error signal e , as well as its magnitude $\|e\|$. Ignoring the sign of the error, but trying to be as general as possible, we state the desired properties of an error response as follows:

- Magnitude of the peaks should be small
- Duration of the peaks should be short
- The error should converge to zero as fast as possible
- No steady-state error in the output should be present

In Figure 4.1 an illustration of a typical transient error signal is seen. The terms used here are described in Table 4.1. Additionally, the integral of $\|e\|$ might be denoted by e_{total} . Using these terms, transient response of a system would be optimal if it provides

- minimum e_{peak}
- minimum $\sum_i \Delta t_i$
- minimum t_{low}
- minimum e_{total}

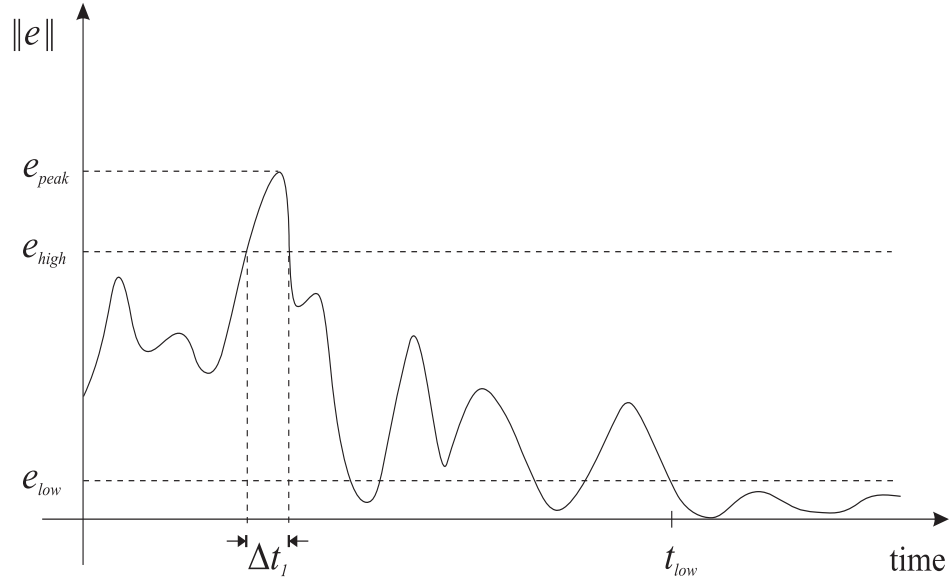


Figure 4.1. Transient error characteristics

Table 4.1. Description of the symbols in transient error measurement

symbol	description
e_{peak}	peak value of $\ e\ $
e_{high}	maximum acceptable value of $\ e\ $
e_{low}	upper limit of desired value of $\ e\ $
t_{low}	the time after which $\ e\ \leq e_{low}$, $\forall t \geq t_{low}$
Δt_i	i^{th} -period in which $\ e\ > e_{high}$

5. ADAPTIVE CONTROL OF NONLINEAR SYSTEMS

Mathematical modeling is a vital part of any given control system design methodology. If we have an exact model of the physical system, then it is relatively easy to design an automatic control system. However, the exact model is not always available in reality. Moreover, if the dynamics of a physical system contains nonlinearities and if the control approach is based on feedback linearization, the problem of inaccurate modeling is often encountered. This is an important problem, since even very small differences between the physical system and the mathematical model may lead to large output errors.

Another problem is the change of system dynamics due to changes in the environment in an unexpected manner. For example, although the exact mathematical model of a robot arm is usually available, there may be abrupt changes in the loads that the manipulator handles. This kind of unknown changes in the dynamics lead to an inexact model and create difficulties in its control.

Adaptive control tries to remove the effects of the uncertainties in the system model. Let $\theta \in \mathbb{R}^p$ be a vector of unknown parameters. Since θ is unknown, an estimate of it is used in the mathematical model of the physical system. If the vector of the estimated parameters are denoted by $\hat{\theta} \in \mathbb{R}^p$, the ideal condition is that $\hat{\theta} = \theta$. However, the parameters used in the model are not always identical, nor sufficiently close, to the true parameters of the physical plant. Objective of the adaptive control is to provide a mechanism such that the model parameters converge to the true values, even if the actual system parameters change unexpectedly. The term *adaptive* is used to indicate this property of this type of controllers [4, 5, 6].

All known adaptive control schemes are based on parametric uncertainties. There are usually fixed, or *slowly* varying, parameters in a physical system. The control scheme may depend on explicit identification of the system parameters, or may depend on calculation of controller parameters without requiring the estimates of the system

parameters. The former is known in the literature as *indirect* adaptive control scheme, and the latter is referred to as *direct* adaptive control scheme. These direct and indirect schemes will be explained in detail in the following sections.

5.1. Identification of Unknown Parameters

The unknown parameters in a parametrically modeled dynamic system are identified using an adaptive methodology. In the direct schemes, the identification is usually based on Lyapunov analysis. An update rule for the parameter vector is obtained which leads to the negative-definiteness of the derivative of the closed loop Lyapunov function. However, the arising parameter vector in this Lyapunov analysis is not necessarily the same parameter vector as in the system model. The identified parameters, also called the controller parameters, are implicitly based on the system parameters. Hence the identification on the system parameters are achieved as well. On the contrary, in the indirect schemes, the system parameters are identified separately, and the controller parameters are calculated using the estimates of the system parameters. Since the identification is performed separately in the indirect schemes, there are various identification methods which are applicable. Here we will analyze two of the most important ones, namely, the observer-based identification and regressor filtering-based identification. In both schemes, it is vital that the nonlinear parametric system can be expressed in linear regressor form.

We consider the nonlinear dynamics given by,

$$\dot{x} = f(x, \theta) + g(x, \theta)u \quad (5.1)$$

where $x \in \mathbb{R}^n$ is the state vector, $f, g : \mathcal{D} \subset \mathbb{R}^n \mapsto \mathbb{R}^n$ are sufficiently smooth vector fields, $\theta \in \mathcal{S} \subset \mathbb{R}^p$ is the unknown parameter vector, and $u \in \mathbb{R}$ is the input. The vector fields f and g have the form,

$$f(x, \theta) = \sum_{k=1}^p \theta_k f_k(x) \quad (5.2)$$

$$g(x, \theta) = \sum_{k=1}^p \theta_k g_k(x) \quad (5.3)$$

The regressor $w(x, u) \in \mathbb{R}^{p \times n}$ is formulated as,

$$w(x, u) = \begin{bmatrix} f_1(x) + g_1(x)u \\ \vdots \\ f_p(x) + g_p(x)u \end{bmatrix} \quad (5.4)$$

which includes all the nonlinearities in the system dynamics. Hence we can write (5.1) in the *regressor form* as,

$$\dot{x} = w^T(x, u)\theta \quad (5.5)$$

The system dynamics in regressor form will constitute a basis for the identifier structures in the following sections.

5.1.1. Observer-Based Identification

In this study, we assume availability of the full state vector, as stated in Assumption 2 in Chapter 4. The identifier system for the observer-based method is given with the following identification dynamics equations,

$$\left. \begin{aligned} \dot{\hat{x}} &= A(\hat{x} - x) + w^T(x, u)\hat{\theta} \\ \dot{\hat{\theta}} &= -w(x, u)P(\hat{x} - x) \end{aligned} \right\} \quad (5.6)$$

where $\hat{x} \in \mathbb{R}^n$ and $\hat{\theta} \in \mathbb{R}^p$ are estimates of x and θ respectively, $A \in \mathbb{R}^{n \times n}$ is a Hurwitz matrix and $P \in \mathbb{R}^{n \times n}$ is the positive definite solution of the Lyapunov equation

$$A^T P + P A = -Q \quad (5.7)$$

with $Q \in \mathbb{R}^{n \times n}$ being any symmetric, positive definite matrix. Defining $\tilde{x} = \hat{x} - x$ as the observer state error and $\tilde{\theta} = \hat{\theta} - \theta$ as the parameter error, we obtain the following

error system:

$$\left. \begin{aligned} \dot{\tilde{x}} &= A\tilde{x} + w^T(x, u)\tilde{\theta} \\ \dot{\tilde{\theta}} &= -w(x, u)P\tilde{x} \end{aligned} \right\} \quad (5.8)$$

Theorem 5.1 [9] *Let the identifier error dynamics of the system in (5.1) be given by (5.8). Then, $\tilde{\theta} \in \mathcal{L}_\infty$ and $\tilde{x} \in \mathcal{L}_2 \cap \mathcal{L}_\infty$. If the system given in (5.1) is bounded-input bounded-state (BIBS) stable as well, then $\dot{\tilde{x}} \in \mathcal{L}_\infty$ and $\tilde{x} \rightarrow 0$ as $t \rightarrow \infty$.*

Proof: We choose a Lyapunov function of the form

$$V(\tilde{x}, \tilde{\theta}) = \tilde{x}^T P \tilde{x} + \tilde{\theta}^T \tilde{\theta} \quad (5.9)$$

Differentiating (5.9), we get,

$$\dot{V}(\tilde{x}, \tilde{\theta}) = \dot{\tilde{x}}^T P \tilde{x} + \tilde{x}^T P \dot{\tilde{x}} + 2\tilde{\theta}^T \dot{\tilde{\theta}} \quad (5.10)$$

Substituting (5.7) and (5.8) in (5.10) and arranging the terms we obtain,

$$\dot{V} = -\tilde{x}^T Q \tilde{x} \leq 0 \quad (5.11)$$

This leads to $\tilde{x} \in \mathcal{L}_\infty$ and $\tilde{\theta} \in \mathcal{L}_\infty$. Integrating (5.11),

$$\int_0^t \dot{V}(\tilde{x}(\tau), \tilde{\theta}(\tau)) d\tau = \int_0^t -\tilde{x}^T(\tau) Q \tilde{x}(\tau) d\tau \quad (5.12)$$

$$V(\tilde{x}(t), \tilde{\theta}(t)) - V(\tilde{x}(0), \tilde{\theta}(0)) = - \int_0^t \tilde{x}^T(\tau) Q \tilde{x}(\tau) d\tau \quad (5.13)$$

We already know from (5.9) and boundedness of $\tilde{x}$ and $\tilde{\theta}$ that the left-hand side of (5.13) is bounded. Hence $\tilde{x} \in \mathcal{L}_2$. If the system is BIBS stable with bounded input, then the regressor w is bounded, and from (5.8), $\dot{\tilde{x}} \in \mathcal{L}_\infty$. Having $\tilde{x} \in \mathcal{L}_2 \cap \mathcal{L}_\infty$ and

$\dot{\tilde{x}} \in \mathcal{L}_\infty$, using Lemma 2.3, we get $\lim_{t \rightarrow \infty} \tilde{x} = 0$ □

5.1.2. Regressor-Filtering Based Identification

The system in regressor form given in (5.5) represents a dynamic system. The filtering-based approach first converts this dynamic equation into an algebraic one using appropriate filters. Then several identification schemes can be applied to this algebraic system.

We consider the system dynamics in regressor form given in (5.5). Let x_f and w_f be the filtered forms of x and w respectively, defined as,

$$\left. \begin{aligned} \sigma \dot{x}_f &= -x_f + x \\ \sigma \dot{w}_f^T &= -w_f^T + \sigma w^T(x, u) \end{aligned} \right\} \quad (5.14)$$

where $\sigma \in \mathbb{R}_+$ is the time constant of a 1^{st} -order filter which is used for eliminating the derivative term in (5.5) and obtaining a linear error equation. The state x can be reconstructed as,

$$x = w_f^T \theta + x_f + [x(0) - w_f^T(0)\theta - x_f(0)]e^{-t/\sigma} \quad (5.15)$$

The equivalence of this representation can be easily verified by taking the time derivative of (5.15) to obtain (5.5). Assuming that σ is chosen small enough, the steady state estimate is then given by,

$$\hat{x} = w_f^T \hat{\theta} + x_f \quad (5.16)$$

We define the state prediction error as, $\tilde{x} = \hat{x} - x$ and parameter error as, $\tilde{\theta} = \hat{\theta} - \theta$. Then, the state prediction error is obtained as,

$$\tilde{x} = w_f^T \tilde{\theta} \quad (5.17)$$

The linear error equation in (5.17) can now be used to identify the unknown parameter vector $\hat{\theta}$, using different identification schemes. Here we choose a simple gradient scheme for parameter identification.

Theorem 5.2 [9] *For the error equation given in (5.17), the gradient estimation scheme*

$$\dot{\tilde{\theta}} = -\alpha w_f \tilde{x} \quad , \quad \alpha \in \mathbb{R}_+ \quad (5.18)$$

leads to $\tilde{\theta} \in \mathcal{L}_\infty$ and $\tilde{x} \in \mathcal{L}_2$. If $w_f(x)$ is bounded as well, then $\tilde{x} \in \mathcal{L}_\infty$, $\dot{\tilde{x}} \in \mathcal{L}_\infty$ and $\tilde{x} \rightarrow 0$ as $t \rightarrow \infty$.

Proof: We choose a Lyapunov function of the form,

$$V(\tilde{\theta}) = \frac{1}{2} \tilde{\theta}^T \tilde{\theta} \quad (5.19)$$

Differentiating (5.19), we get,

$$\dot{V}(\tilde{\theta}) = \tilde{\theta}^T \dot{\tilde{\theta}} \quad (5.20)$$

Substituting (5.18) and (5.17) in (5.20) and arranging the terms we obtain,

$$\dot{V}(\tilde{\theta}) = -\alpha \tilde{x}^T \tilde{x} \quad (5.21)$$

which leads to $\dot{V}(\tilde{\theta}) \leq 0$. Hence $\tilde{\theta} \in \mathcal{L}_\infty$. Now integrating (5.21),

$$\int_0^t \dot{V}(\tilde{\theta}(\tau)) d\tau = \int_0^t -\alpha \tilde{x}^T(\tau) \tilde{x}(\tau) d\tau \quad (5.22)$$

$$V(\tilde{\theta}(t)) - V(\tilde{\theta}(0)) = -\alpha \int_0^t \tilde{x}^T(\tau) \tilde{x}(\tau) d\tau \quad (5.23)$$

We already know from (5.19) and boundedness of $\tilde{\theta}$ that the left-hand side of (5.23) is bounded. Hence $\tilde{x} \in \mathcal{L}_2$. If the system is BIBS stable, then $x \in \mathcal{L}_\infty$. This leads to $w(x) \in \mathcal{L}_\infty$. The filter for w in (5.14) is stable for $\sigma > 0$. The use of Lemma 2.4 leads to $w_f \in \mathcal{L}_\infty$. We use this in (5.17) and obtain $\tilde{x} \in \mathcal{L}_\infty$. Then using (5.18), we have $\dot{\tilde{x}} \in \mathcal{L}_\infty$. Having $\tilde{x} \in \mathcal{L}_2 \cap \mathcal{L}_\infty$ and $\dot{\tilde{x}} \in \mathcal{L}_\infty$, using Lemma 2.3 we get $\lim_{t \rightarrow \infty} \tilde{x} = 0$. $\square$

Lemma 5.1 [6] *If the regressor w in (5.5) is sufficiently rich, leading to persistent excitation such that,*

$$\beta_1 I > \int_t^{t+\delta} w_f(\tau) w_f^T(\tau) d\tau > \beta_2 I \quad , \quad \beta_1, \beta_2, \delta > 0 \quad (5.24)$$

holds, then both $\tilde{x}$ and $\tilde{\theta}$ converge to zero asymptotically.

Proof: The proof of this lemma is given in [6].

5.2. Adaptive Control of Nonlinear Systems

Adaptive control of nonlinear systems has been widely studied during the last two decades, and it is still an active research area. Most of the work published so far rely on availability of the full state vector for feedback. On the other hand, some output feedback approaches are present as well, as in [44, 45, 46]. One of the main assumptions in this thesis is that full state vector of the nonlinear plant is available for feedback. For that reason, the methods using state feedback will be emphasized in the rest of the thesis.

The major difficulty in the analysis of global stability of nonlinear adaptive systems arises from the fact that, in general, the closed loop system depends on the unknown parameters. There are different approaches which use state feedback, to overcome this difficulty [37, 47]:

1. The first approach is based on *uncertainty constraints*, such that the unknown parameters are restricted to appear in the same equation as the control (matching conditions), or at most one integrator away (extended matching conditions). The uncertainties due to the unknown parameters can then be rejected as disturbances. Locations of the parameters are restricted in this approach, but the nonlinearities can be of any smooth type [10, 11, 48].
2. The second approach is based on *nonlinearity constraints*, and there is no restriction imposed on the location of the uncertainties. The nonlinearities are restricted by the growth conditions. A typical scheme in this approach requires global linear growth condition (or Lipschitz condition) in the system nonlinearities [8, 9].
3. The third approach requires *geometric constraints*, such that the nonlinear system is required to be of a certain form, such as parametric strict-feedback form. But there is no restrictions on the location of the uncertainties and the type of the nonlinearities in the system dynamics. This approach is named as the *backstepping approach* [14].

5.2.1. Adaptive Control under Matching Conditions

This approach is introduced by Taylor *et al.* [11], extended by Kanellakopoulos, Kokotovic and Marino [13], and generalized by Pomet and Praly [49]. In simple terms, the matching condition can be described as the case when the unknown parameters are in the same equation with the control input. The formal definition of the matching conditions is given in Section 5.2.1.2. First, we will analyze the *matching* and *non-matching* cases by examples.

5.2.1.1. Matching Conditions Examples. Here we present three examples. The first example demonstrates the case where matching conditions are satisfied. The second example provides a system satisfying only the extended-matching conditions. The third example gives a system where the matching conditions are not satisfied at all.

Example 5.1 Consider the following scalar system:

$$\dot{x} = \theta x + u \quad (5.25)$$

In this system, $x \in \mathbb{R}$ is the state variable, $u \in \mathbb{R}$ is the control input, and $\theta \in \mathbb{R}$ is an unknown parameter. We see that the uncertainty is matched in here, since the unknown parameter θ is in the same equation as the control u . If a priori information is available about the bounds on θ such as $|\theta| < \bar{\theta}$, then a linear feedback control,

$$u = -kx - \bar{\theta}x \quad (5.26)$$

would stabilize the system, where $k > 0$ is a design parameter. If no information about the bounds on θ is available, then a nonlinear control,

$$u = -k_1x - k_2x^3 \quad (5.27)$$

with $k_1 > 0, k_2 > 0$ would stabilize the system.

Another method to achieve the same goal is *Lyapunov method*. Denote the estimate of unknown parameter θ at time t as $\hat{\theta}(t)$. A Lyapunov function $V : \mathbb{R}^2 \mapsto \mathbb{R}_+$ is then formed as,

$$V(x(t), \hat{\theta}(t)) = \frac{1}{2}x^2 + \frac{1}{2}(\hat{\theta} - \theta)^2 \quad (5.28)$$

with a parameter update law defined as,

$$\dot{\hat{\theta}} = \vartheta(x, \hat{\theta}) \quad (5.29)$$

and a control law given by,

$$u = \kappa(x, \hat{\theta}) \quad (5.30)$$

where $\vartheta, \kappa : \mathbb{R}^2 \mapsto \mathbb{R}$ are some functions of the state and the parameter estimate, which make (5.28) a non-increasing function of time, such that, $\dot{V} \leq -\lambda x^2$ must be satisfied for $\lambda > 0$. Differentiating (5.28) along the system trajectories, we get,

$$\dot{V} = x(u + \theta x^2) + (\hat{\theta} - \theta)\dot{\hat{\theta}} \leq -\lambda x^2 \quad (5.31)$$

and arranging the terms we obtain,

$$xu + \hat{\theta}\dot{\hat{\theta}} + \theta(x^2 - \dot{\hat{\theta}}) \leq -\lambda x^2 \quad (5.32)$$

If we set

$$\dot{\hat{\theta}} = x^2 \quad (5.33)$$

the term with the unknown factor θ is removed. Using the remaining part, $\kappa(x, \hat{\theta})$ can be determined and u is obtained as,

$$u = -(\lambda + \hat{\theta})x. \quad (5.34)$$

In the above Lyapunov method, we explicitly identified the plant parameter θ as well. As seen in both methods, when the matching condition is met, decoupling the parametric uncertainty is straightforward [37, 42].

Example 5.2 In the following system, the unknown parameter θ does not appear in the same equation as the control input u :

$$\left. \begin{aligned} \dot{x}_1 &= x_2 + \theta x_1^2 \\ \dot{x}_2 &= u \end{aligned} \right\} \quad (5.35)$$

where $x = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^T \in \mathbb{R}^2$ is the state vector, $\theta \in \mathbb{R}$ is an unknown parameter

and $u \in \mathbb{R}$ is the input. Hence the uncertainty and the control are *not matched* in this system. On the other hand, the uncertainty is only one integrator away from the control input, and this corresponds to the so called *extended matching* case. We seek for a diffeomorphism $\phi : x \mapsto \xi$, where $\xi = \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} \in \mathbb{R}^2$ is the linearized state vector. Choosing an artificial output as $y = x_1$, yields a relative degree two system, because

$$\dot{y} = \dot{x}_1 = x_2 + \theta x_1^2 \quad (5.36)$$

and

$$\ddot{y} = \dot{x}_2 + 2\theta x_1 \dot{x}_1 = u + 2\theta x_1(x_2 + \theta x_1) \quad (5.37)$$

Applying the certainty equivalence principle, we know from (3.7) that a linearizing diffeomorphism is given by,

$$\phi(x) = \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 + \hat{\theta} x_1^2 \end{bmatrix} \quad (5.38)$$

where $\hat{\theta} \in \mathbb{R}$ is the estimate of θ . Differentiating (5.38), we get,

$$\left. \begin{aligned} \dot{\xi}_1 &= \dot{x}_1 = x_2 + \theta x_1^2 = \xi_2 + (\theta - \hat{\theta})\xi_1^2 \\ \dot{\xi}_2 &= \dot{\hat{\theta}} x_1^2 + 2\hat{\theta} x_1(x_2 + \hat{\theta} x_1^2) + u \end{aligned} \right\} \quad (5.39)$$

Introducing the control law

$$u = -\alpha_1 \xi_1 - \alpha_2 \xi_2 - 2\hat{\theta} x_1(x_2 + \hat{\theta} x_1^2) - \dot{\hat{\theta}} x_1^2 + v \quad (5.40)$$

with $\alpha_1, \alpha_2 \in \mathbb{R}_+$ and $v \in \mathbb{R}$, we obtain,

$$\dot{\xi} = \begin{bmatrix} 0 & 1 \\ -\alpha_1 & -\alpha_2 \end{bmatrix} \xi + \begin{bmatrix} \xi_1^2 \\ 2\hat{\theta}\xi_1^3 \end{bmatrix} \tilde{\theta} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} v \quad (5.41)$$

where $\tilde{\theta} = \theta - \hat{\theta}$ is the parameter error and v is the reference input. Let $\xi_d \in \mathbb{R}^2$ be the vector of desired states. We can then define the error vector as $e = \xi - \xi_d$ and write the error dynamics as,

$$\dot{e} = \begin{bmatrix} 0 & 1 \\ -\alpha_1 & -\alpha_2 \end{bmatrix} e + \begin{bmatrix} \xi_1^2 \\ 2\hat{\theta}\xi_1^3 \end{bmatrix} \tilde{\theta} = Ae + B\tilde{\theta} \quad (5.42)$$

For $\alpha_1 > 0$ and $\alpha_2 > 0$, the matrix A has all its eigenvalues in the left half plane. Then it is possible to find a matrix P such that,

$$A^T P + PA = -Q \quad (5.43)$$

where Q is a positive definite matrix. Choosing a Lyapunov function as,

$$V = e^T P e + \frac{1}{2} \tilde{\theta}^2 \quad (5.44)$$

and differentiating (5.44) we obtain,

$$\dot{V} = e^T (A^T P + PA) e + \tilde{\theta} B^T P e + \tilde{\theta} \dot{\tilde{\theta}} \quad (5.45)$$

If the parameter update law is chosen as,

$$\dot{\tilde{\theta}} = B^T P e \quad (5.46)$$

this leads to

$$\dot{\tilde{\theta}} = \frac{d}{dt}(\theta - \hat{\theta}) = -\dot{\hat{\theta}} = -B^T P e \quad (5.47)$$

Substituting this in (5.45), we obtain,

$$\dot{V} = -e^T Q e \quad (5.48)$$

which is a negative definite function as long as any component of e is nonzero. Therefore, with the linearizing control law in (5.40), stable reference trajectory ξ_d , and the parameter update law in (5.46), the tracking error will always go to zero [5].

When the linearizing control law (5.40) is inspected, it is seen that it depends on the states x_1, x_2 , estimate of the parameter $\hat{\theta}$, and derivative of the parameter estimate $\dot{\hat{\theta}}$. These components are all available in this case, and stable adaptive control of the system (5.35) is thus possible.

Example 5.3 Consider the following system,

$$\left. \begin{aligned} \dot{x}_1 &= x_2 + \theta x_1^2 \\ \dot{x}_2 &= x_3 \\ \dot{x}_3 &= u \end{aligned} \right\} \quad (5.49)$$

Even the *extended* matching condition is not satisfied for this system. If a similar Lyapunov design is performed as in Example 5.2, it can be easily seen that the control law would have to include the second derivative of the parameter estimate, $\ddot{\hat{\theta}}$, which is not available. Therefore, one cannot obtain a stable adaptive controller for the system in (5.49) using standard Lyapunov design. When the matching conditions are not satisfied, an adaptive controller can still be obtained either using the nonlinearity-constraints approach, if the nonlinearities are Lipschitz; or using the backstepping approach, if the system is in parametric strict-feedback form. In this example, the Lipschitz condition is not satisfied either, because of the term x_1^2 , but the system is in the strict-feedback form.

5.2.1.2. Definition of the Matching Conditions. We will now analyze the matching conditions for the systems having parametric uncertainties in a formal manner. Consider the following nonlinear system:

$$\left. \begin{aligned} \dot{x} &= f(x, \theta) + g(x, \theta)u \\ y &= h(x) \end{aligned} \right\} \quad (5.50)$$

where $x \in \mathbb{R}^n$ is the state vector, $u \in \mathbb{R}$ is the input, $\theta \in \mathcal{S} \subset \mathbb{R}^p$ is the vector of unknown parameters, $f, g : \mathcal{D} \subset \mathbb{R}^{n \times p} \mapsto \mathbb{R}^n$ are sufficiently smooth vector fields, $h : \mathcal{D} \subset \mathbb{R}^n \mapsto \mathbb{R}$ is a sufficiently smooth function, and $y \in \mathbb{R}$ is the system output. Since the parameter vector θ is unknown, we use the estimate of the parameter vector and re-write the model as,

$$\left. \begin{aligned} \dot{x} &= f(x, \hat{\theta}) + \Delta f(x) + g(x, \hat{\theta})u + \Delta g(x)u \\ y &= h(x) \end{aligned} \right\} \quad (5.51)$$

where $\hat{\theta}$ is the estimate of the parameter vector θ , $\Delta f(x) = f(x, \theta) - f(x, \hat{\theta})$ and $\Delta g(x) = g(x, \theta) - g(x, \hat{\theta})$. In (5.51), $f(x, \hat{\theta})$ and $g(x, \hat{\theta})$ represent the *known* nominal parts because the state x is assumed to be measurable, and the vector of parameter estimates $\hat{\theta}$ is available. The uncertainty is considered in $\Delta f(x)$ and $\Delta g(x)$. Similar to the Proposition 3.1, we have the following proposition:

Proposition 5.1 [35] *Consider the system in (5.50) and the model in (5.51). If the relative degree of the system is γ , and if the uncertainties Δf and Δg satisfy*

$$\left. \begin{aligned} L_{\Delta f} L_{f(x, \hat{\theta})}^i h &= 0 \quad , \quad i = 0, \dots, \gamma - 1 \\ L_{\Delta g} L_{f(x, \hat{\theta})}^i h &= 0 \quad , \quad i = 0, \dots, \gamma - 1 \end{aligned} \right\} \quad (5.52)$$

then the linearizing control law for the system in (5.50) also linearizes the model in (5.51).

Proof: Since the uncertainty due to estimation errors is modeled as structural uncertainty, proof of this proposition is the same as the proof of Proposition 3.1. The linearizing control law for the nonlinear model in (5.51) is given as,

$$u = \frac{v - L_{f(x, \hat{\theta})}^\gamma h(x)}{L_{g(x, \hat{\theta})} L_{f(x, \hat{\theta})}^{\gamma-1} h(x)} \quad (5.53)$$

This control input also provides asymptotic stability for the system in (5.50) if v is

designed such that,

$$v = y_r^{(\gamma)} + \lambda_1(y_r^{(\gamma-1)} - L_{f(x,\hat{\theta})}^{\gamma-1}) + \cdots + \lambda_\gamma(y_r - h(x)) \quad (5.54)$$

where y_r is the reference output, $y_r^{(i)}$ is i^{th} -derivative of the reference output, and λ_i 's are all positive, real coefficients for $i = 1, \dots, \gamma$. $\square$

The definitions for the matching conditions and the extended matching conditions given in Definition 3.5 and Definition 3.6, respectively, are valid for the parametric model in (5.50). Another definition for the matching conditions is expressed through the so-called *triangularity* conditions [42].

Definition 5.1 (Matching conditions) *Consider the nonlinear system,*

$$\dot{x} = f(x, \theta) + (g_0(x) + g(x, \theta))u \quad (5.55)$$

with

$$f(x, \theta) = \sum_{i=1}^p \theta_i f_i(x) \quad (5.56)$$

and

$$g(x, \theta) = \sum_{i=1}^p \theta_i g_i(x) \quad (5.57)$$

where $g_0 : \mathbb{R}^n \mapsto \mathbb{R}^n$ is the nominal part of the system, which includes no uncertainty, and $\theta \in \mathcal{S} \subset \mathbb{R}^p$, with $\mathcal{S}$ being a compact set. The system in (5.55) is said to satisfy the matching conditions, if,

$$\forall f_i \in \text{span}\{g_0\}, \quad \forall g_i \in \text{span}\{g_0\} \quad , \quad i = 1, \dots, p \quad (5.58)$$

holds.

In [42], the extended matching condition is also defined similarly, for the systems which do not contain unknown parameters in the g vector. We next present an example to clarify the above matching condition definition.

Example 5.4 Consider a simple pendulum system:

$$\left. \begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -\frac{g}{l} \sin x_1 - \frac{k}{m} x_2 + \frac{1}{ml^2} T \end{aligned} \right\} \quad (5.59)$$

where T is the input torque of the system. Assume that the parameters g and k are not known. Then, the system can be written as,

$$\left. \begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -\theta_1 \sin x_1 - \theta_2 x_2 + cu \end{aligned} \right\} \quad (5.60)$$

where $c > 0$ is a constant, and θ_1, θ_2 are treated as unknown parameters. Writing the system in the matrix notation as,

$$\dot{x} = \begin{bmatrix} x_2 \\ 0 \end{bmatrix} 1 + \begin{bmatrix} 0 \\ -\sin x_1 \end{bmatrix} \theta_1 + \begin{bmatrix} 0 \\ -x_2 \end{bmatrix} \theta_2 + \begin{bmatrix} 0 \\ c \end{bmatrix} u \quad (5.61)$$

it is seen that,

$$\begin{bmatrix} 0 \\ -\sin x_1 \end{bmatrix} \in \text{span} \left\{ \begin{bmatrix} 0 \\ c \end{bmatrix} \right\} \quad (5.62)$$

and

$$\begin{bmatrix} 0 \\ -x_2 \end{bmatrix} \in \text{span} \left\{ \begin{bmatrix} 0 \\ c \end{bmatrix} \right\} \quad (5.63)$$

Hence, the matching conditions are satisfied. In words, this means that the uncertainties are in the *range space* of the control input, so it is possible to cancel these uncertainties by some appropriate control action.

5.2.2. Adaptive Control under Growth Conditions

A direct adaptive control methodology is proposed in [8] by Sastry and Isidori, where globally Lipschitz continuity of all nonlinear signals as well as exponentially asymptotically stability of the zero dynamics are required. In this scheme, parameterization of error dynamics also leads to *over-parameterization*, where the number of control parameters is larger than the parameters in the original system dynamics model. This over-parameterization problem is not present in indirect adaptive control schemes [9, 10].

The next three sections are devoted to the derivation of adaptive controllers under matching conditions and growth conditions.

5.3. Indirect Adaptive Control

In this section we consider a class of nonlinear systems which do not satisfy the matching conditions. The nonlinear system is as given in Chapter 4,

$$\left. \begin{aligned} \dot{x} &= f(x, \theta) + g(x, \theta)u \\ y &= h(x) \end{aligned} \right\} \quad (5.64)$$

where $x \in \mathbb{R}^n$ is the state vector, $f, g \in \mathbb{R}^{n \times p}$ and $h \in \mathbb{R}$ are sufficiently smooth vector fields, $\theta \in \mathcal{S}$ is the unknown parameter vector with $\mathcal{S} \subset \mathbb{R}^p$ being a compact set, $u \in \mathbb{R}$ is the input and $y \in \mathbb{R}$ is the output.

The indirect scheme estimates the plant parameters by incorporating a separate identifier model. Due to Assumption 4 given in Chapter 4, the system dynamics in (5.64) can be written in the regressor form as,

$$\dot{x} = w^T(x, u)\theta \quad (5.65)$$

with w being the regressor matrix and θ being the parameter vector as defined in

Chapter 4. That is,

$$f(x, \theta) = \sum_{k=1}^p \theta_k f_k(x) \quad (5.66)$$

$$g(x, \theta) = \sum_{k=1}^p \theta_k g_k(x) \quad (5.67)$$

The unknown parameter vector θ in (5.65) can be identified using the filtering-based identification method described in Section 5.1. As a result, x_f and w_f being the filtered forms of x and w respectively, a gradient update rule is derived as,

$$\dot{\hat{\theta}} = -\alpha w_f \tilde{x} \quad , \quad \alpha \in \mathbb{R}_+ \quad (5.68)$$

where $\tilde{x} = \hat{x} - x$ is the state prediction error. The asymptotic convergence of $\tilde{x}$ to zero is given by Theorem 5.2. Furthermore, the convergence of parameter estimate vector $\hat{\theta}$ to the actual system parameter vector θ is expressed in Lemma 5.1 under persistently exciting signal condition.

With the estimated system parameters at hand, in order to generate the control signal based on these estimates, we write the system dynamics in (5.64) in the normal form as,

$$\left. \begin{aligned} \dot{\xi}_1 &= \xi_2 \\ &\vdots \\ \dot{\xi}_\gamma &= L_f^\gamma h(x)(x, \theta) + L_g L_f^{\gamma-1} h(x)(x, \theta) \cdot u \\ \dot{\eta} &= q(\xi, \eta) \\ y &= \xi_1 \end{aligned} \right\} \quad (5.69)$$

where $\dot{\eta} = q(\xi, \eta)$ defines the internal dynamics with exponentially stable zero-dynamics.

Example 5.5 Consider the following nonlinear system:

$$\left. \begin{aligned} \dot{x}_1 &= \theta_1 \sin x_1 + x_2 + \theta_2 \\ \dot{x}_2 &= x_1 + \theta_3 x_2 u \\ y &= x_1 \end{aligned} \right\} \quad (5.70)$$

where $a, b, c \in \mathbb{R}$ are parameters of the system. Using (5.66),

$$f(x, \theta) = \underbrace{\begin{bmatrix} x_2 \\ x_1 \end{bmatrix}}_{f_1} \cdot 1 + \underbrace{\begin{bmatrix} \sin x_1 \\ 0 \end{bmatrix}}_{f_2} \cdot \theta_1 + \underbrace{\begin{bmatrix} 1 \\ 0 \end{bmatrix}}_{f_3} \cdot \theta_2 + \underbrace{\begin{bmatrix} 0 \\ 0 \end{bmatrix}}_{f_4} \cdot \theta_3 \quad (5.71)$$

and using (5.67),

$$g(x, \theta) = \underbrace{\begin{bmatrix} 0 \\ 0 \end{bmatrix}}_{g_1} \cdot 1 + \underbrace{\begin{bmatrix} 0 \\ 0 \end{bmatrix}}_{g_2} \cdot \theta_1 + \underbrace{\begin{bmatrix} 0 \\ 0 \end{bmatrix}}_{g_3} \cdot \theta_2 + \underbrace{\begin{bmatrix} 0 \\ x_2 \end{bmatrix}}_{g_4} \cdot \theta_3 \quad (5.72)$$

This system has three unknown parameters, namely θ_1 , θ_2 and θ_3 . But, according to (5.66) and (5.67), $p = 4$. Due to the terms in (5.70) which are not multipliers of the parameters, the constant 1 appears as the fourth parameter in (5.71) and (5.72).

Relative degree of this system is two. Using (5.69), the linearizing diffeomorphism is obtained as,

$$\xi = \phi(x, \theta) = \begin{bmatrix} x_1 \\ \theta_1 \sin x_1 + x_2 + \theta_2 \end{bmatrix} \quad (5.73)$$

which leads to the linearized system in the normal form as,

$$\left. \begin{aligned} \dot{\xi}_1 &= \xi_2 \\ \dot{\xi}_2 &= \xi_1 + \theta_1 \xi_2 \cos \xi_1 + (\theta_3 \xi_2 - \theta_1 \theta_3 \sin \xi_1 - \theta_2 \theta_3) u \\ y &= \xi_1 \end{aligned} \right\} \quad (5.74)$$

Note that $L_f^\gamma h(x)(x, \theta)$ and $L_g L_f^{\gamma-1} h(x)(x, \theta)$ in (5.69) can now be expanded in terms of multi-linear parameter elements as,

$$L_f^\gamma h(x)(x, \theta) = \underbrace{\sum_{i_1}^p \sum_{i_2}^p \cdots \sum_{i_\gamma}^p \frac{\partial}{\partial x} \left(\frac{\partial}{\partial x} \left(\cdots \left(\frac{\partial h}{\partial x} f_{i_1}(x) \right) \cdots \right) f_{i_{\gamma-1}}(x) \right) f_{i_\gamma}(x)}_{\mathcal{F}^T(x)} \underbrace{\theta_{i_1} \theta_{i_2} \cdots \theta_{i_\gamma}}_{\Theta} \quad (5.75)$$

and

$$L_g L_f^{\gamma-1} h(x)(x, \theta) = \underbrace{\sum_j^p \sum_{i_1}^p \cdots \sum_{i_{\gamma-1}}^p \frac{\partial}{\partial x} \left(\frac{\partial}{\partial x} \left(\cdots \left(\frac{\partial h}{\partial x} f_{i_1}(x) \right) \cdots \right) f_{i_{\gamma-1}}(x) \right) g_j(x)}_{\mathcal{G}^T(x)} \underbrace{\theta_j \theta_{i_1} \cdots \theta_{i_{\gamma-1}}}_{\Theta} \quad (5.76)$$

The above representation is based on the definitions of f and g given in equations (5.66) and (5.67), respectively. For the brevity of the analysis to follow, equations (5.75) and (5.76) can be written in vector notation as,

$$L_f^\gamma h(x)(x, \theta) \stackrel{def}{=} \mathcal{F}^T(x) \Theta \quad (5.77)$$

and

$$L_g L_f^{\gamma-1} h(x)(x, \theta) \stackrel{def}{=} \mathcal{G}^T(x) \Theta \quad (5.78)$$

where Θ represents the multi-linear parameter vector as defined in (5.75) and (5.76). Using this notation, the normal form given in (5.69) can be written as,

$$\left. \begin{aligned} \dot{\xi}_1 &= \xi_2 \\ &\vdots \\ \dot{\xi}_\gamma &= \mathcal{F}^T(x) \Theta + \mathcal{G}^T(x) \Theta \cdot u \\ \dot{\eta} &= q(\xi, \eta) \\ y &= \xi_1 \end{aligned} \right\} \quad (5.79)$$

Example 5.6 Consider the system in Example 5.5. The $\dot{\xi}_2$ equation of the linearized system dynamics in (5.74) can be written as,

$$\dot{\xi}_2 = \underbrace{\begin{bmatrix} \xi_1 & \xi_2 \cos \xi_1 & 0 & 0 & 0 \end{bmatrix}}_{\mathcal{F}^T} \Theta + \underbrace{\begin{bmatrix} 0 & 0 & \xi_2 & -\sin \xi_1 & -1 \end{bmatrix}}_{\mathcal{G}^T} \Theta \cdot u \quad (5.80)$$

where the multi-linear parameter vector Θ is given by,

$$\Theta = \begin{bmatrix} 1 \\ \theta_1 \\ \theta_3 \\ \theta_1 \theta_3 \\ \theta_2 \theta_3 \end{bmatrix} \quad (5.81)$$

We continue the derivation of the indirect controller to obtain the control input “ u ”. Based on the parameter estimates obtained by the identification scheme, we can now generate the control signal u using the certainty equivalence principle as,

$$u = \frac{1}{\mathcal{G}^T(x)\hat{\Theta}}(-\mathcal{F}^T(x)\hat{\Theta} + v) \quad (5.82)$$

where $\hat{\Theta}$ is the estimate of the multi-linear parameter vector and v is the tracking control signal to be designed using the given desired trajectory information, as,

$$v = y_r^{(\gamma)} + \lambda_1(y_r^{(\gamma-1)} - y^{(\gamma-1)}) + \dots + \lambda_\gamma(y_r - y) \quad (5.83)$$

where $\lambda_1, \dots, \lambda_\gamma$ are chosen positive constants such that,

$$s^\gamma + \lambda_1 s^{\gamma-1} + \dots + \lambda_\gamma = 0 \quad (5.84)$$

represents a Hurwitz polynomial. Note that in (5.83), $y_r, \dot{y}_r, \dots, y_r^{(\gamma)}$ are already available. On the other hand, $\dot{y}, \dots, y^{(\gamma)}$ need to be calculated from $L_f h, L_f^2 h, \dots, L_f^{\gamma-1} h$,

respectively. However, since these Lie derivatives are not free from unknown parameters, we apply the certainty equivalence principle once more, to obtain,

$$\hat{y}^{(i)} = \dot{\xi}_i = \sum_{j_1}^p \cdots \sum_{j_k}^p \frac{\partial}{\partial x} (\cdots (\frac{\partial h}{\partial x} f_{j_1}) \cdots) f_{j_k} \hat{\theta}_{j_1} \cdots \hat{\theta}_{j_k} \quad , \quad i = 1, \dots, \gamma \quad (5.85)$$

Using (5.85), we obtain the certainty equivalence based tracking control estimate $\hat{v}$ as,

$$\hat{v} = y_r^{(\gamma)} + \lambda_1(y_r^{(\gamma-1)} - \hat{y}^{(\gamma-1)}) + \cdots + \lambda_\gamma(y_r - \hat{y}) \quad (5.86)$$

Now, the control input is now fully constructable from the known signals as,

$$u = \frac{1}{\mathcal{G}^T(x)\hat{\Theta}}(-\mathcal{F}^T(x)\hat{\Theta} + \hat{v}) \quad (5.87)$$

Plugging the control signal u in the system dynamics given in (5.79), one can now obtain the closed loop error dynamics. In order to do this, first we rewrite the $\dot{\xi}_\gamma$ term in (5.79) as,

$$\dot{\xi}_\gamma = \mathcal{F}^T(x)\Theta + \mathcal{G}^T(x)\Theta \cdot u - [\mathcal{F}^T(x)\hat{\Theta} + \mathcal{G}^T(x)\hat{\Theta} \cdot u] + [\mathcal{F}^T(x)\hat{\Theta} + \mathcal{G}^T(x)\hat{\Theta} \cdot u] \quad (5.88)$$

Then, defining the multi-linear parameter error vector as,

$$\tilde{\Theta} \stackrel{def}{=} \hat{\Theta} - \Theta \quad (5.89)$$

we obtain,

$$\dot{\xi}_\gamma = \mathcal{F}^T(x)\hat{\Theta} + \mathcal{G}^T(x)\hat{\Theta} \cdot u + \varphi_1^T(x, u)\tilde{\Theta} \quad (5.90)$$

where $\varphi_1 = -[\mathcal{F}(x) + \mathcal{G}(x)u]$. If we now plug in the certainty equivalence control u into (5.90), we get,

$$\dot{\xi}_\gamma = \hat{v} + \varphi_1^T(x, u)\tilde{\Theta} \quad (5.91)$$

Note that $\hat{v}$ can be written as,

$$\hat{v} = y_r^{(\gamma)} + \lambda_1(y_r^{(\gamma-1)} - y^{(\gamma-1)}) + \lambda_1(y^{(\gamma-1)} - \hat{y}^{(\gamma-1)}) + \dots + \lambda_\gamma(y_r - y) + \lambda_\gamma(y - \hat{y}) \quad (5.92)$$

which is the exact tracking control v plus an offset term which is a function of parameter errors,

$$\hat{v} = v + \varphi_2^T(x, u)\tilde{\Theta} \quad (5.93)$$

where φ_2 is the regressor vector satisfying,

$$\varphi_2^T(x, u)\tilde{\Theta} = \lambda_1(y^{(\gamma-1)} - \hat{y}^{(\gamma-1)}) + \dots + \lambda_{\gamma-1}(\dot{y} - \dot{\hat{y}}) \quad (5.94)$$

Then, equation (5.91) can now be written as,

$$\dot{\xi}_\gamma = v + \varphi^T(x, u)\tilde{\Theta} \quad (5.95)$$

where $\varphi = \varphi_1 + \varphi_2$ includes all the offset terms due to parameter errors. Next, defining an error term as,

$$e_i = \xi_i - y_r^{(i-1)} \quad , \quad i = 1, \dots, \gamma \quad (5.96)$$

the closed loop error dynamics takes the following form,

$$\left. \begin{aligned} \xi_i &= e_i + y_r^{(i-1)} \\ \dot{e} &= Ae + \varphi^T(\xi, \eta, u)\tilde{\Theta} \\ \dot{\eta} &= q(\xi, \eta) \end{aligned} \right\} \quad (5.97)$$

where A is a Hurwitz matrix, with parameters defined in (5.84), and $\varphi(\xi, \eta, u)$ represents the regressor vector defined in (5.95).

Theorem 5.3 [9] *Consider the nonlinear system defined in (5.64) with a bounded*

tracking signal y_r , with bounded derivatives as $y_r, \dot{y}_r, \dots, y_r^{(\gamma)}$; and with the internal dynamics $q(\xi, \eta)$ which is globally Lipschitz in ξ, η . If the regressor vector φ is bounded for bounded ξ, η, u , such that $\|\varphi^T(\xi, \eta, u)\| \leq b_\varphi(\|\xi\| + \|\eta\|)$ holds, and if the input is persistently exciting such that $\tilde{\Theta} \rightarrow 0$ as $t \rightarrow \infty$, then the control laws given by (5.86) and (5.87) along with the parameter update rule in (5.68) result in a closed loop stable system with tracking convergence such that $y(t) \rightarrow y_r(t)$ as $t \rightarrow \infty$.

Proof: The closed loop error dynamics of the system is now given in (5.97). Since the zero dynamics is exponentially stable, the converse Lyapunov theorem [34] states that, there exists a Lyapunov function $V_q(\eta)$ such that,

$$\left. \begin{aligned} c_1 \|\eta\|^2 &\leq V_q(\eta) \leq c_2 \|\eta\|^2 \\ \frac{\partial V_q}{\partial \eta} q(0, \eta) &\leq -c_3 \|\eta\|^2 \\ \left\| \frac{\partial V_1(\eta)}{\partial \eta} \right\| &\leq c_4 \|\eta\| \end{aligned} \right\} \quad (5.98)$$

with c_1, c_2, c_3 and c_4 being all positive constants, where $\|\cdot\|$ represents the Euclidean norm. Consider the following Lyapunov function for the error dynamics given by,

$$V_c(e, \eta) = e^T P e + \mu V_q(\eta) \quad , \quad \mu \in \mathbb{R}_+ \quad (5.99)$$

where P is the solution of the Lyapunov equation,

$$A^T P + P A = -I \quad (5.100)$$

Differentiating (5.99), we get,

$$\begin{aligned} \dot{V}_c(e, \eta) &= e^T (A^T P + P A) e + 2e^T P \varphi^T(\xi, \eta, u) \tilde{\Theta} + \mu \frac{\partial V_q}{\partial \eta} q(\xi, \eta) \\ &= -e^T e + 2e^T P \varphi^T(\xi, \eta, u) \tilde{\Theta} \\ &\quad + \mu \frac{\partial V_q}{\partial \eta} q(0, \eta) + \left(\mu \frac{\partial V_q}{\partial \eta} q(\xi, \eta) - \mu \frac{\partial V_q}{\partial \eta} q(0, \eta) \right) \end{aligned} \quad (5.101)$$

The internal-dynamics of the system is assumed to be globally Lipschitz continuous in all of its arguments. Hence, we have,

$$\|q(\xi, \eta) - q(0, \eta)\| \leq b_q \|\xi\| \quad (5.102)$$

for all ξ , where $b_q \in \mathbb{R}_+$ is an upper bound. Since $y_r, \dot{y}_r, \dots, y_r^{(\gamma-1)}$ are also bounded, denoting b_r as the upper bound of y_r and its derivatives, we have,

$$\|\xi\| \leq \|e\| + b_r \quad (5.103)$$

Due to bounded $\|\xi\|$ and $\|\eta\|$, we also have as the condition of the theorem,

$$\|\varphi^T(\xi, \eta, u)\| \leq b_\varphi(\|\xi\| + \|\eta\|) \quad (5.104)$$

Using (5.104) and P from (5.100), we get,

$$\|2P\varphi^T(\xi, \eta, u)\| \leq b_1(\|\xi\| + \|\eta\|) \quad , \quad b_1 > 0 \quad (5.105)$$

Using (5.98), (5.102), (5.103) and (5.105) in (5.101), we obtain,

$$\dot{V}_c(e, \eta) \leq -\|e\|^2 + b_1\|e\|(\|e\| + \|\eta\| + b_r)\|\tilde{\Theta}\| - \mu c_3\|\eta\|^2 + \mu c_4 b_q \|\eta\|(\|e\| + b_r) \quad (5.106)$$

which can be further expressed as,

$$\begin{aligned} \dot{V}_c(e, \eta) &\leq -\left(\frac{1}{2}\|e\| - b_1 b_r \|\tilde{\Theta}\|\right)^2 - \left(\frac{1}{2}\sqrt{\mu c_3}\|\eta\| - \sqrt{\frac{\mu}{c_3}}c_4 b_q b_r\right)^2 \\ &\quad - \frac{3}{4}\|e\|^2 + (b_1 b_r \|\tilde{\Theta}\|)^2 - \frac{3}{4}\mu c_3\|\eta\|^2 + \frac{\mu}{c_3}(c_4 b_q b_r)^2 \\ &\quad + b_1 \|\tilde{\Theta}\| \|e\|^2 + (b_1 \|\tilde{\Theta}\| + \mu c_4 b_q)\|e\| \|\eta\| \end{aligned} \quad (5.107)$$

$$\begin{aligned} \dot{V}_c(e, \eta) &\leq \left(-\frac{3}{4} + b_1 \|\tilde{\Theta}\|\right)\|e\|^2 + (b_1 \|\tilde{\Theta}\| + \mu c_4 b_q)\|e\| \|\eta\| - \frac{3}{4}\mu c_3\|\eta\|^2 \\ &\quad + (b_1 b_r \|\tilde{\Theta}\|)^2 + \frac{\mu}{c_3}(c_4 b_q b_r)^2 \end{aligned} \quad (5.108)$$

Equation (5.108) can be also written in the following form,

$$\dot{V}_c(e, \eta) \leq - \begin{bmatrix} \|e\| \\ \|\eta\| \end{bmatrix}^T Q_1 \begin{bmatrix} \|e\| \\ \|\eta\| \end{bmatrix} + (b_1 b_r \|\tilde{\Theta}\|)^2 + \frac{\mu}{c_3} (c_4 b_q b_r)^2 \quad (5.109)$$

where

$$Q_1 = \begin{bmatrix} \frac{3}{4} - b_1 \|\tilde{\Theta}\| & -\frac{1}{2}(b_1 \|\tilde{\Theta}\| + \mu c_4 b_q) \\ -\frac{1}{2}(b_1 \|\tilde{\Theta}\| + \mu c_4 b_q) & \frac{3}{4} \mu c_3 \end{bmatrix} \quad (5.110)$$

The right-hand side of (5.109) can be negative only if Q_1 is a positive definite matrix, and the other two terms are small compared to the first term. The positive definiteness of Q_1 is provided if its first element and its determinant are both positive. These two conditions correspond, respectively, to

1. $\frac{3}{4} - b_1 \|\tilde{\Theta}\| > 0$, and
2. $\det(Q_1) = (\frac{3}{4} - b_1 \|\tilde{\Theta}\|) \frac{3}{4} \mu c_3 - \frac{1}{4} (b_1 \|\tilde{\Theta}\| + \mu c_4 b_q)^2 > 0$

It is easy to see that both of the above conditions can be satisfied, if,

$$\|\tilde{\Theta}\| \leq \frac{1}{2b_1} \quad (5.111)$$

and

$$\mu \leq \frac{3c_3}{4(c_4 b_q)^2} \quad (5.112)$$

Since the input is persistently exciting as an assumption of the theorem, we have $\|\tilde{\Theta}\| \rightarrow 0$ as $t \rightarrow \infty$. This implies that, after some time T_1 , we will have $\|\tilde{\Theta}\| \leq \frac{1}{2b_1}$. On the other hand, the only restriction on μ is that $\mu > 0$. We can easily set a positive value for μ which satisfies (5.112). Then, the matrix Q_1 in (5.109) becomes positive definite. The second term on the right-hand side of (5.109) (i.e. $(b_1 b_r \|\tilde{\Theta}\|)^2$) is approaching zero asymptotically. Finally, the last term in (5.109) is a small positive

constant, since μ can be selected as a small positive value. Therefore, for all $t \geq T_1$, whenever $\|e\|$ or $\|\eta\|$ is large, $\dot{V}_c < 0$. This implies that $\|e\| \in \mathcal{L}_\infty$ and $\|\eta\| \in \mathcal{L}_\infty$. From (5.103), $\|\xi\| \in \mathcal{L}_\infty$. Hence, $\varphi^T(\xi, \eta, u)$ is bounded as well, and,

$$\dot{e} = Ae + \varphi^T(\xi, \eta, u)\tilde{\Theta} \quad (5.113)$$

is an exponentially stable linear system having an input which approaches to zero asymptotically. Consequently, $e_i \rightarrow 0$ as $t \rightarrow \infty$. Since $e_1 = \xi_1 - y_r$, then $\xi_1 = y \rightarrow y_r$ as $t \rightarrow \infty$. $\square$

Remark 5.1 *Note that the indirect scheme discussed above basically results in asymptotic tracking error convergence, if the multi-linear parameter error vector $\tilde{\Theta}$ converges to zero. Hence, it is intuitively clear that faster convergence of parameters to their true values would directly increase the transient performance of the indirect scheme. This is actually the direct implication of the certainty equivalence based control.*

We have now completed derivation of indirect adaptive controller for a class of control-affine nonlinear systems. In the following section we consider a restricted set of the above class of nonlinear systems and we will build a model reference adaptive control (MRAC) system using the direct scheme.

5.4. Model Reference Adaptive Control under Matching Conditions

Here we introduce a MRAC scheme for a subset of the system defined in (4.1). Consider the following system,

$$\dot{x} = f(x) + g(x)u + \sum_{k=1}^p q_k(x)\theta_k \quad (5.114)$$

where $x \in \mathbb{R}^n$ is the state vector, $f, g \in \mathbb{R}^n$ are known smooth vector fields with $f(0) = 0$ and $g(0) \neq 0$, and $\theta = [\theta_1, \dots, \theta_p]^T$ is the constant but unknown parameter vector which takes values on a compact set $\mathcal{S} \subset \mathbb{R}^p$. $q_k \in \mathbb{R}^n$ are known vector fields

and $u \in \mathbb{R}$ is the input. We assume that the origin is the equilibrium point of (5.114), without loss of generality.

The nominal part (f, g) of the system in (5.114) is given by,

$$\dot{x} = f(x) + g(x)u \quad (5.115)$$

and is assumed to be input-state linearizable via state feedback, as described in Section 3.2.2. Since the input-state linearizability conditions stated by Theorem 3.1 hold, there exists a smooth scalar function $\psi(x) : \mathbb{R}^n \mapsto \mathbb{R}$ with $\psi(0) = 0$ such that,

$$L_g L_f^i \psi(x) = 0 \quad , \quad i = 0, \dots, n-2 \quad (5.116)$$

$$L_g L_f^{n-1} \psi(x) \neq 0 \quad (5.117)$$

Then we obtain the linearized state variables using the linearizing diffeomorphism $\xi = \phi(x)$ as,

$$\xi_i = L_f^{i-1} \psi(x) \quad , \quad i = 1, \dots, n \quad (5.118)$$

The state feedback control law given by,

$$u = \frac{1}{L_g L_f^{n-1} \psi(x)} (-L_f^n \psi(x) + v) \quad (5.119)$$

converts the nominal system in (5.115) into a linear system in Brunowsky canonical form with the new state vector ξ as,

$$\dot{\xi}_i = \xi_{i+1} \quad , \quad i = 1, \dots, n-1 \quad (5.120)$$

$$\dot{\xi}_n = v \quad (5.121)$$

which is equivalent to,

$$\dot{\xi} = A\xi + bv \quad (5.122)$$

where (A, b) is a controllable pair.

Referring to Definition 5.1, we see that the system given by (5.114) satisfies the matching conditions. Since the nominal system (f, g) in (5.115) is also input-state linearizable, we can now implement a model reference adaptive controller for the original system.

Theorem 5.4 *For the nonlinear system defined in (5.114) with input-state linearizable nominal part (f, g) and the linearizing diffeomorphism $\xi = \phi(x)$, if the matching conditions defined in Definition 5.1 are satisfied, then for a given reference model,*

$$\dot{z} = A_r z + b_r v_r \quad , \quad z \in \mathbb{R}^n \quad (5.123)$$

with A_r being a chosen Hurwitz matrix, (A_r, b_r) being a controllable pair, and v_r being the reference input, then there exists an adaptive controller with control input and parameter update dynamics defined as,

$$u = \chi_1(x, \xi_r, v_r, \hat{\theta}) \quad (5.124)$$

$$\dot{\hat{\theta}} = \chi_2(x, \xi_r, v_r, \hat{\theta}) \quad (5.125)$$

such that $x(t) \in \mathcal{L}_\infty$, $\hat{\theta} \in \mathcal{L}_\infty$ and $\lim_{t \rightarrow \infty} |x(t) - \phi^{-1}(\xi_r)| = 0$.

Proof: Based on the matching conditions given in Definition 5.1, there exists a θ independent state-space diffeomorphism $\xi = \phi(x)$ and the state feedback control,

$$u = \alpha(x) + \beta(x)v \quad (5.126)$$

such that the original system in (5.114) is transformed into a linear system,

$$\dot{\xi}_i = \xi_{i+1} \quad , \quad i = 1, \dots, n-1 \quad (5.127)$$

$$\dot{\xi}_n = \sum_{k=1}^p \theta_k w_k(\xi) + v = w^T(\xi)\theta + v \quad (5.128)$$

where $w_k(\xi) = q_k(\phi^{-1}(\xi))$. For the reference model, we have,

$$\dot{z}_i = z_{i+1} \quad , \quad i = 1, \dots, n-1 \quad (5.129)$$

$$\dot{z}_n = -\sum_{k=1}^n a_{k-1} z_k + v_r \quad (5.130)$$

where the coefficients $a_0, \dots, a_{n-1}$ constitute $(s^n + a_{n-1}s^{n-1} + \dots + a_0)$, which is a Hurwitz polynomial of choice. The reference model can also be written as,

$$\dot{z} = A_r z + b_r v_r \quad (5.131)$$

with

$$A_r = \begin{bmatrix} 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ -a_0 & -a_1 & \cdots & -a_{n-1} \end{bmatrix} \quad , \quad b_r = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \quad (5.132)$$

We define the control input v as,

$$v = -\sum_{i=1}^n a_{i-1} \xi_i - w^T(\xi)\hat{\theta} + v_r \quad (5.133)$$

where $\hat{\theta}$ is the estimate of θ . Plugging v in the equations (5.127) and (5.128) and using the reference dynamics given in (5.131), we obtain the error dynamics as,

$$\dot{e} = A_r e + b_r w^T(\xi)\tilde{\theta} \quad (5.134)$$

where $e = \xi - z$ is the state error and $\tilde{\theta} = \hat{\theta} - \theta$ is the parameter error vector. We can now use a Lyapunov function of the form,

$$V(e, \tilde{\theta}) = e^T P e + \tilde{\theta}^T \Gamma^{-1} \tilde{\theta} \quad (5.135)$$

with $\Gamma^{-1} \in \mathbb{R}^{p \times p}$ being a positive definite gain matrix, and $P \in \mathbb{R}^{n \times n}$ being the positive definite solution of the Lyapunov equation,

$$A_r^T P + P A_r = -Q \quad (5.136)$$

and $Q \in \mathbb{R}^{n \times n}$ is any symmetric positive definite matrix. Taking the time derivative of V and using the dynamics in (5.134), we obtain,

$$\dot{V}(e, \tilde{\theta}) = e^T \underbrace{(A_r^T P + P A_r)}_{-Q} e + 2e^T P b_r w^T \tilde{\theta} + 2\tilde{\theta}^T \Gamma^{-1} \dot{\tilde{\theta}} \quad (5.137)$$

If we choose,

$$\dot{\tilde{\theta}} = \hat{\dot{\theta}} = -\Gamma w b_r^T P e \quad (5.138)$$

as the update rule for the parameter vector, we obtain,

$$\dot{V}(e, \tilde{\theta}) = -e^T Q e \leq 0 \quad (5.139)$$

Equation (5.139) implies $e \in \mathcal{L}_\infty$ and $\tilde{\theta} \in \mathcal{L}_\infty$. Integrating both sides of (5.139) we get,

$$\int_0^t \dot{V}(e(\tau), \tilde{\theta}(\tau)) d\tau = \int_0^t -e^T(\tau) Q e(\tau) d\tau \quad (5.140)$$

$$V(e(t), \tilde{\theta}(t)) - V(e(0), \tilde{\theta}(0)) = \int_0^t -e^T(\tau) Q e(\tau) d\tau \quad (5.141)$$

Since e and $\tilde{\theta}$ are bounded, the left-hand side of (5.141) is bounded. This means the

integral on the right-hand side exists and $e \in \mathcal{L}_2$. From (5.134) we also have $\dot{e} \in \mathcal{L}_\infty$. Since $e \in \mathcal{L}_2 \cap \mathcal{L}_\infty$ and $\dot{e} \in \mathcal{L}_\infty$, applying Lemma 2.3 we have $e \rightarrow 0$ as $t \rightarrow \infty$. This is equivalent to $\lim_{t \rightarrow \infty} |x(t) - \phi^{-1}(z)| = 0$. $\square$

Next, we will derive a direct adaptive control scheme for the class of nonlinear systems without requiring the matching conditions.

5.5. Direct Adaptive Control

We consider the nonlinear system given in Chapter 4, as,

$$\left. \begin{aligned} \dot{x} &= f(x, \theta) + g(x, \theta)u \\ y &= h(x) \end{aligned} \right\} \quad (5.142)$$

where $x \in \mathbb{R}^n$ is the state vector, $f, g \in \mathbb{R}^{n \times p}$ and $h \in \mathbb{R}$ are sufficiently smooth vector fields, $\theta \in \mathcal{S}$ is the unknown parameter vector with $\mathcal{S} \subset \mathbb{R}^p$ being a compact set, $u \in \mathbb{R}$ is the input and $y \in \mathbb{R}$ is the output. The system given in (5.142) is a quite general class of nonlinear systems and it does not necessarily satisfy the matching conditions.

The direct adaptive scheme estimates the controller parameters rather than the plant parameters. We first rewrite the system dynamics in the normal form as,

$$\left. \begin{aligned} \dot{\xi}_1 &= \xi_2 \\ \vdots & \\ \dot{\xi}_\gamma &= L_f^\gamma h(x)(x, \theta) + L_g L_f^{\gamma-1} h(x)(x, \theta) \cdot u \\ \dot{\eta} &= q(\xi, \eta) \\ y &= \xi_1 \end{aligned} \right\} \quad (5.143)$$

where $\dot{\eta} = q(\xi, \eta)$ defines the internal dynamics with exponentially stable zero-dynamics. Note that $L_f^\gamma h(x)(x, \theta)$ and $L_g L_f^{\gamma-1} h(x)(x, \theta)$ in (5.143) can now be expanded in terms

of multi-linear parameter elements as,

$$L_f^\gamma h(x)(x, \theta) = \sum_{i_1}^p \sum_{i_2}^p \cdots \sum_{i_\gamma}^p \underbrace{\theta_{i_1} \theta_{i_2} \cdots \theta_{i_\gamma}}_{\Theta^T} \underbrace{\frac{\partial}{\partial x} \left(\frac{\partial}{\partial x} \left(\cdots \left(\frac{\partial h}{\partial x} f_{i_1}(x) \right) \cdots \right) f_{i_{\gamma-1}}(x) \right) f_{i_\gamma}(x)}_{\mathcal{F}} \quad (5.144)$$

and

$$L_g L_f^{\gamma-1} h(x)(x, \theta) = \sum_j^p \sum_{i_1}^p \cdots \sum_{i_{\gamma-1}}^p \underbrace{\theta_j \theta_{i_1} \cdots \theta_{i_{\gamma-1}}}_{\Theta^T} \underbrace{\frac{\partial}{\partial x} \left(\frac{\partial}{\partial x} \left(\cdots \left(\frac{\partial h}{\partial x} f_{i_1}(x) \right) \cdots \right) f_{i_{\gamma-1}}(x) \right) g_j(x)}_{\mathcal{G}} \quad (5.145)$$

The above representation is based on the definitions of f and g given in equations (4.7) and (4.8), respectively. The equations (5.144) and (5.145) can be written in vector notation as,

$$L_f^\gamma h(x)(x, \theta) \stackrel{def}{=} \Theta^T \mathcal{F}(x) \quad (5.146)$$

and

$$L_g L_f^{\gamma-1} h(x)(x, \theta) \stackrel{def}{=} \Theta^T \mathcal{G}(x) \quad (5.147)$$

where Θ represents the multi-linear parameter vector as defined in (5.144) and (5.145).

Using this notation, the normal form given in (5.143) can be written as,

$$\left. \begin{aligned} \dot{\xi}_1 &= \xi_2 \\ \vdots & \\ \dot{\xi}_\gamma &= \Theta^T \mathcal{F}(x) + \Theta^T \mathcal{G}(x) \cdot u \\ \dot{\eta} &= q(\xi, \eta) \\ y &= \xi_1 \end{aligned} \right\} \quad (5.148)$$

Based on the parameter estimates obtained by the identification scheme which will be defined in the sequel, one can generate the control signal u using the certainty

equivalence principle as,

$$u = \frac{1}{\hat{\Theta}^T \mathcal{G}(x)} (-\hat{\Theta}^T \mathcal{F}(x) + v) \quad (5.149)$$

where $\hat{\Theta}$ is the estimate of the multi-linear parameter vector and v is the tracking control signal to be designed using the given desired trajectory information, as,

$$v = y_r^{(\gamma)} - \alpha_1(y^{(\gamma-1)} - y_r^{(\gamma-1)}) - \dots - \alpha_\gamma(y - y_r) \quad (5.150)$$

where $\alpha_1, \dots, \alpha_\gamma$ are chosen positive constants, such that,

$$s^\gamma + \alpha_1 s^{\gamma-1} + \dots + \alpha_\gamma = 0 \quad (5.151)$$

represents a Hurwitz polynomial. Note that in (5.150), $y_r, \dot{y}_r, \dots, y_r^{(\gamma)}$ are already available. On the other hand, $\dot{y}, \dots, y^{(\gamma)}$ need to be calculated from $L_f h, L_f^2 h, \dots, L_f^{\gamma-1} h$, respectively. However, since these Lie derivatives are not free from unknown parameters, we apply the certainty equivalence principle once more, to obtain,

$$\hat{y}^{(k)} = \sum_{j_1}^p \dots \sum_{j_k}^p \frac{\partial}{\partial x} (\dots (\frac{\partial h}{\partial x} f_{j_1}) \dots) f_{j_k} \hat{\theta}_{j_1} \dots \hat{\theta}_{j_k} \quad (5.152)$$

where $1 \leq k \leq \gamma$. Using (5.152), we obtain the certainty equivalence based tracking control estimate $\hat{v}$ as,

$$\hat{v} = y_r^{(\gamma)} - \alpha_1(\hat{y}^{(\gamma-1)} - y_r^{(\gamma-1)}) - \dots - \alpha_\gamma(\hat{y} - y_r) \quad (5.153)$$

Using (5.153), the control input is now fully constructable from the known signals given by,

$$u = \frac{1}{\hat{\Theta}^T \mathcal{G}(x)} (-\hat{\Theta}^T \mathcal{F}(x) + \hat{v}) \quad (5.154)$$

We next define the tracking error as,

$$e \stackrel{def}{=} y - y_r \quad (5.155)$$

and the multi-linear parameter error vector as,

$$\tilde{\Theta} \stackrel{def}{=} \hat{\Theta} - \Theta \quad (5.156)$$

Plugging the control signal u in the system dynamics given in (5.148), we can write the $\dot{\xi}_\gamma$ equation as,

$$\dot{\xi}_\gamma = \Theta^T \mathcal{F}(x) + \frac{\Theta^T \mathcal{G}(x)}{\hat{\Theta}^T \mathcal{G}(x)} (-\hat{\Theta}^T \mathcal{F}(x) + \hat{v}) \quad (5.157)$$

Adding $-v$ to both sides, we get,

$$\dot{\xi}_\gamma - v = \Theta^T \mathcal{F}(x) + \frac{\Theta^T \mathcal{G}(x)}{\hat{\Theta}^T \mathcal{G}(x)} (-\hat{\Theta}^T \mathcal{F}(x) + \hat{v}) - v \quad (5.158)$$

Since $\dot{\xi}_\gamma = y^{(\gamma)}$, using (5.150), we have the closed loop error dynamics as,

$$e^{(\gamma)} + \alpha_1 e^{(\gamma-1)} + \dots + \alpha_\gamma e = \Theta^T \mathcal{F}(x) + \frac{\Theta^T \mathcal{G}(x)}{\hat{\Theta}^T \mathcal{G}(x)} (-\hat{\Theta}^T \mathcal{F}(x) + \hat{v}) - v \quad (5.159)$$

$$= \underbrace{-\tilde{\Theta}^T \mathcal{F}(x) - \tilde{\Theta}^T \mathcal{G}(x) \frac{-\hat{\Theta}^T \mathcal{F}(x) + \hat{v}}{\hat{\Theta}^T \mathcal{G}(x)}}_{\tilde{\Theta}^T W_1} + \underbrace{\hat{v} - v}_{\tilde{\Theta}^T W_2} \quad (5.160)$$

$$= \tilde{\Theta}^T W_1 + \tilde{\Theta}^T W_2 \quad (5.161)$$

Summing W_1 and W_2 and denoting the sum as W , we can express the error dynamics in a more compact form as,

$$e^{(\gamma)} + \alpha_1 e^{(\gamma-1)} + \dots + \alpha_\gamma e = \tilde{\Theta}^T W \quad (5.162)$$

where $W = W_1 + W_2$ is the complete regressor vector and $\tilde{\Theta}$ is the multi-linear parameter error vector. Note that the regressor vector W is a function of x and $\hat{\theta}$. It is also implicitly a function of y_r and its derivatives. In (5.162), the polynomial on the left-hand side is Hurwitz. In order to be able to use arguments similar to those in derivation of the adaptive control of linear systems, we need a strictly positive real (SPR) transfer function. This requires an error signal ε of the form,

$$\varepsilon = \beta_1 e^{(\gamma-1)} + \dots + \beta_\gamma e \quad (5.163)$$

such that,

$$\frac{\beta_1 s^{\gamma-1} + \dots + \beta_\gamma}{s^\gamma + \alpha_1 s^{\gamma-1} + \dots + \alpha_\gamma} \quad (5.164)$$

is SPR. However, $\dot{e}, \dots, e^{(\gamma-1)}$ are not measurable since they depend on unknown parameters as,

$$\begin{aligned} \dot{e} &= L_f h(x)(x, \theta) - \dot{y}_r \\ \ddot{e} &= L_f^2 h(x)(x, \theta) - \ddot{y}_r \\ &\vdots \\ e^{(\gamma-1)} &= L_f^{\gamma-1} h(x)(x, \theta) - y_r^{(\gamma-1)} \end{aligned}$$

To overcome this difficulty, the error augmentation technique is used as proposed in [8]. We first define,

$$M(s) \stackrel{def}{=} \frac{1}{s^\gamma + \alpha_1 s^{\gamma-1} + \dots + \alpha_\gamma} \quad (5.165)$$

such that, (5.162) can be written as,

$$e = M(s) \cdot \tilde{\Theta}^T W \quad (5.166)$$

We now define the augmented error as,

$$\varepsilon \stackrel{def}{=} e + \hat{\Theta}^T M(s)\{W\} - M(s)\{\hat{\Theta}^T W\} \quad (5.167)$$

where in the last two terms, the regressor W is filtered by $M(s)$ before being multiplied by $\hat{\Theta}^T$ and after being multiplied by $\hat{\Theta}^T$, respectively. The augmented error defined in (5.167) is calculated using the measurable signals, unlike the error signal in (5.163).

We can also write,

$$\varepsilon = e + \tilde{\Theta}^T M(s)\{W\} - M(s)\{\tilde{\Theta}^T W\} \quad (5.168)$$

so that using (5.166), we obtain,

$$\varepsilon = \tilde{\Theta}^T M(s)W \quad (5.169)$$

Next, let us define,

$$\mu \stackrel{def}{=} M(s)W \quad (5.170)$$

which helps us to write $\varepsilon = \tilde{\Theta}^T \mu$ which is an algebraic error equation. This leads to a parameter update law given as,

$$\dot{\tilde{\Theta}} = \frac{-\varepsilon \mu}{1 + \mu^T \mu} \quad (5.171)$$

The boundedness properties of this normalized update law and asymptotic stability of the closed loop system using the control laws defined in (5.153)-(5.154) and the parameter update law defined in (5.171) can be found in [8].

5.6. Summary

Adaptive control is based on the identification of unknown parameters in the system model. In indirect adaptive schemes, the system parameters are identified explic-

itly and then the control parameters are calculated based on these system parameters. In direct adaptive control, the controller parameters are identified directly. Lyapunov analysis is used in the derivation of control parameters for direct adaptive schemes. In indirect schemes, various parameter identification methods can be utilized once an error equation is obtained using either observer based or filtering based methods on the system dynamics equation.

The location of the unknown parameters in the nonlinear dynamics is important for the adaptive control of such systems. The matching conditions assert that the uncertainties must appear in the same equation as the control input. Whenever the matching conditions hold, the control task is relatively easy. However, this is rather a restrictive approach. When the location of the uncertainties are free, it is still possible to build an adaptive controller if the nonlinearities satisfy certain growth conditions. This corresponds generally to Lipschitz continuity of the nonlinearities.

In this chapter, three adaptive control schemes were presented from the literature. The first one was the indirect adaptive controller for a class of nonlinear systems without requiring the matching conditions. The second one was the direct adaptive controller for a restricted class of nonlinear systems satisfying the matching conditions. The third one was a direct adaptive controller for a class of systems which do not satisfy the matching conditions.

6. ADAPTIVE CONTROL USING MULTIPLE MODELS

6.1. Introduction

The idea of multiple models in control theory is not new. A brief history of the development of this idea can be found in [23]. Use of multiple models in adaptive control for achieving better transient performance was proposed by Narendra and Balakrishnan for linear dynamical systems in [20, 21]. Later on, this method was extended to switching and tuning in different contexts in [23, 25, 50, 51].

Here our objective of including multiple models and switching in an adaptive control framework is to improve the transient response of the system. Main drawback of adaptive control, as stated by many authors, is its slow adaptation and poor transient response. This handicap is more significant if initial estimates of the parameters are far from their actual values, or if there exist large and abrupt changes in the parameters during the execution of a tracking task. It is commonly accepted that the adaptation might be made faster by choosing a larger adaptation gain. However, this is not a remedy to the problem, since higher adaptation gains may lead to performance deterioration as they may increase steady-state noise sensitivity. The proposed scheme allows us to improve the transient response with low adaptation gains.

General mechanism of the proposed system is simple. There are N identification models with $N > 1$. The adaptive control schemes introduced in Chapter 5 might be viewed as special cases with $N = 1$. These N identification models can be either fixed or adaptive. The adaptive and fixed models cases will be explained in the following sections. First, to better visualize the adaptation in the parameter space, we provide Figure 6.1. There is only one identification model in this figure. The vector of actual system parameters θ is represented by the ★-shaped point in the parameter space $\mathcal{S} \subset \mathbb{R}^p$. The initial value of the parameter estimate vector, $\hat{\theta}(0)$, is another point in the same space. The path connecting $\hat{\theta}(0)$ and θ represents the adaptation. When the adaptation begins, the parameter estimate vector $\hat{\theta}(t)$ moves from its initial value

$\hat{\theta}(0)$ toward the actual system parameter vector θ through the indicated path. This corresponds to the asymptotic convergence of the $\hat{\theta}(t)$ to θ . The shape of the path and the travelling time (i.e. the convergence time) is determined by the identification algorithm.

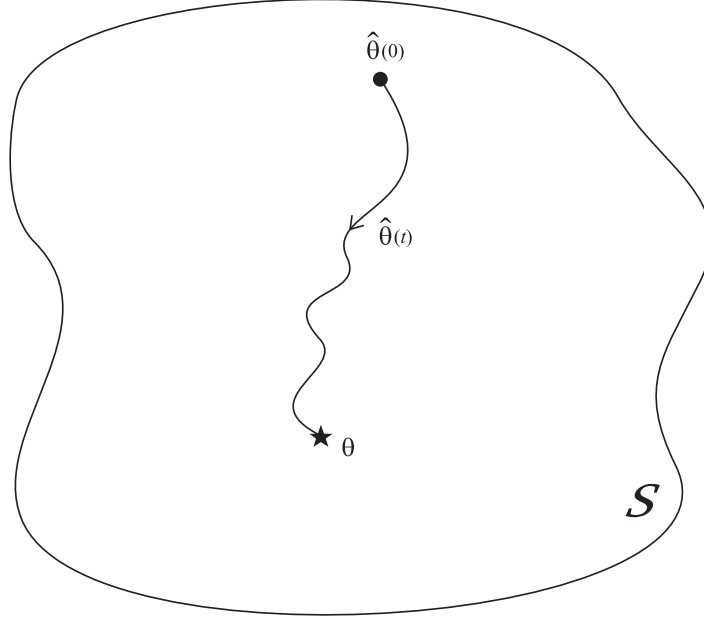


Figure 6.1. Representation of parameter convergence in $\mathcal{S} \subset \mathbb{R}^p$ using only a single model.

A multiple model adaptive control system may consist of all adaptive models. Moreover, each adaptive model can have a different identification algorithm. At each instant, a switching scheme decides on estimates of which model will be used by the controller. If all the adaptive models are asymptotically stable, then arbitrary switching does not affect the stability of the closed loop system [52]. Besides switching, a *composite-path* approach, which does not invoke switching but forms a path using the information provided by the all adaptive models, is proposed by the author and simulated for the linear systems in [53].

An illustration of the parameter convergence in a multiple model adaptive control system with 6 adaptive models is shown in Figure 6.2. In this figure, all six models are convergent. It is easy to visualize that arbitrary switching between these models takes

the parameter estimate to the actual parameter value, because all the paths converge to the actual value θ . On the other hand, starting from one initial point and not switching at all provides asymptotic convergency too, due to the same reason.

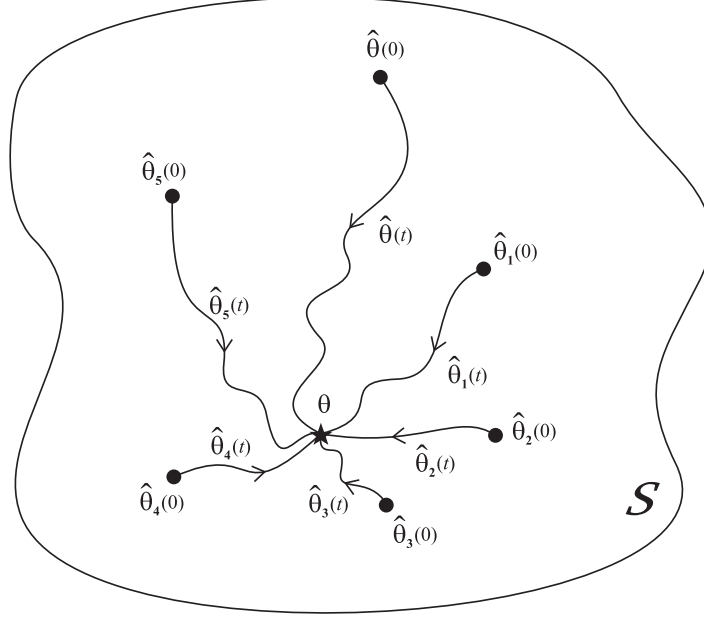


Figure 6.2. Representation of parameter convergence in $\mathcal{S} \subset \mathbb{R}^p$ using multiple adaptive models.

The advantages of the multiple model based adaptive control with all adaptive models can be summarized as follows:

1. Different identification algorithms can be used in each model, such as the (recursive) least-squares, gradient, neural-network based, etc.
2. Switching is not critical if all models are asymptotically stable. Even arbitrary switching leads to a stable system.

Besides these advantages, all adaptive models approach has disadvantages as well. These disadvantages are,

1. All models converge to the same value, i.e. to the actual system parameter vector. If the system parameter changes during the execution of the controller,

the information on the preset points in the parameter space are lost. This is equivalent to a single model case.

2. Cost of making all models adaptive is relatively high.

The first one especially is a serious drawback. If the initial values of the adaptive models are stored for future use and the adaptive models are re-initialized to these values during execution, then this drawback might be eliminated.

On the other hand, N fixed models can be used as well with an adaptive model to construct a multiple model adaptive system. The adaptive model $\hat{\theta}(t)$ is switched to one of the fixed models $\hat{\theta}_j$ which is closer to the actual parameter θ . This is demonstrated in Figure 6.3 for 5 fixed models and one adaptive model. In this figure, the proximity of each fixed model and the adaptive model is evaluated using a performance index, or a switching function. According to the scenario in Figure 6.3, after the adaptation begins at time $t = 0$, the adaptive model is reset to the fixed model $\hat{\theta}_3$ at a time $t_1 > 0$. This resetting is realized via switching. Then, the adaptive model $\hat{\theta}(t)$ continues adaptation from $\hat{\theta}_3$, which is called *tuning*. In fact, more than one switching can take place before the convergence.

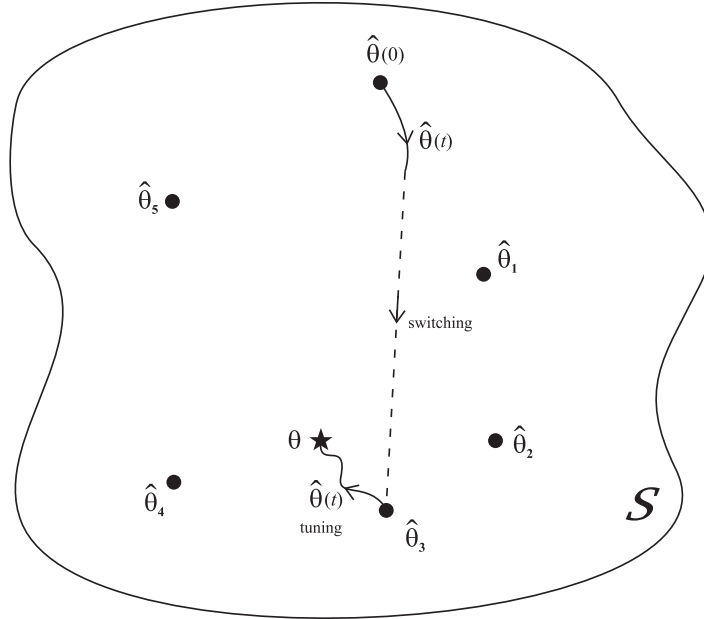


Figure 6.3. Representation of parameter convergence with switching and tuning, using multiple fixed models and one adaptive model.

Figure 6.4 displays the second part of the scenario. The actual system parameter suddenly changes at time t_2 . The new system parameter is another point in the parameter space $\mathcal{S}$. The adaptive system begins adaptation for the new value just after time t_2 . At time t_3 , the switching mechanism decides that the fixed model $\hat{\theta}_5$ is closer to the actual system parameter. Hence, a switching occurs, that is, the adaptive model $\hat{\theta}(t)$ is reset to the fixed model $\hat{\theta}_5$ at time t_3 . The adaptation continues from that point till the convergence is achieved.

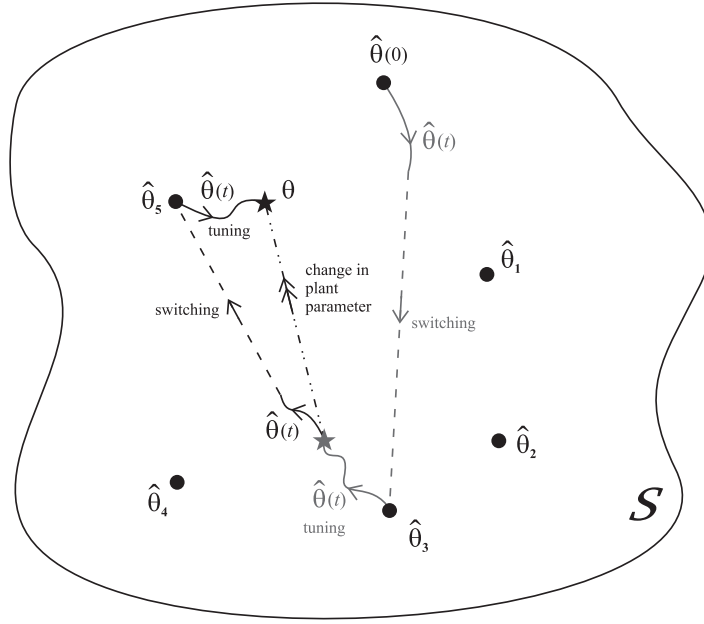


Figure 6.4. Change in the plant parameter, followed by a switching and tuning, using multiple fixed models and one adaptive model.

The above scenario reflects a simple situation. As expressed above, the time and number of switchings are determined by the switching scheme and an appropriately defined performance index. More than one fixed models can be used during adaptation.

The advantage of having multiple models is the capability of quickly switching to the most appropriate model. The overall system should ensure finite switching in order not to have chattering. Therefore we define a permissible switching mechanism to formally express this idea.

Definition 6.1 A finite sequence $T_i \in \mathbb{R}_+$ is defined as a switching sequence if $T_0 = 0$ and $T_i < T_{i+1}$ for all i . Additionally, if there is a number $T_{min} > 0$ such that $T_{i+1} - T_i \geq T_{min}$ for all i , then the sequence is called a permissible switching scheme.

This definition basically states that infinite switching is not allowed in finite time, in a permissible switching scheme. Based on the above definitions, N fixed identification models are represented by N fixed parameter estimates, $\hat{\theta}_j$, with $j = 1, \dots, N$. Assuming a switching period T_{min} , as defined in Definition 6.1, if a switching occurs at time t , the parameter estimate vector $\hat{\theta}(t)$ is instantaneously reset to $\hat{\theta}_j$, where j denotes the index of the selected model.

Next we have to define a switching logic which would lead to better tracking performance. Since the proposed adaptive control schemes are based on the certainty equivalence principle, better parameter identification will naturally lead to better tracking performance.

6.2. Switching between Multiple Models

The multiple model adaptive control requires system identification error to determine the closeness of the parameter estimates to the actual parameter vector in the parameter space. Hence, a performance index can be formed using the identification error. Let $\tilde{x} \in \mathbb{R}^n$ be the vector of identification errors. A general form of the performance index is given as,

$$J_j(t) \stackrel{def}{=} \tilde{x}_j(t)^T G_1 \tilde{x}_j(t) + \int_0^t e^{-\lambda(t-\tau)} \tilde{x}_j(\tau)^T G_2 \tilde{x}_j(\tau) d\tau \quad , \quad j = 0, \dots, N \quad (6.1)$$

where $G_1, G_2 \in \mathbb{R}^{n \times n}$ are positive (semi-)definite weight matrices and $\lambda \geq 0$ is a scalar forgetting factor. In (6.1), the index *zero* (i.e. $j = 0$) corresponds to the adaptive model, where the indices 1 through N correspond to the N fixed models. One of the weight matrices G_1, G_2 can be zero but not both. The first term in the right-hand side of (6.1) considers the instantaneous errors, and the second term accounts

for cumulative identification errors. If $\lambda > 0$, then the cumulative errors are included in the performance index with exponential forgetting. Generally speaking, when G_1 is made larger when G_2 is held constant, the number of switchings increases, as the switching scheme becomes more sensitive to the instantaneous errors. We define the performance index based switching logic as follows:

Definition 6.2 *Given the identification error $\tilde{x}_0$ of the adaptive model and N fixed identification models with the identification errors $\tilde{x}_j$ where $j = 1, \dots, N$, and a permissible switching time T_{min} given in Definition 6.1, the switching logic which is based on the performance index in (6.1) is given as,*

$$j^* = \{j : \min J_j\} \quad , \quad j = 0, \dots, N \quad (6.2)$$

where j^* gives the index for the chosen identification model.

According to the above definition, if the chosen index is zero, the adaptive model is already the best model and no switching occurs at that evaluation instant. If the chosen index is nonzero, this means the fixed model having the selected index is the best model at that instant. Then, the parameter estimate vector $\hat{\theta}(t)$ of the adaptive model is reset to the parameter vector $\hat{\theta}_{j^*}$ of the selected fixed model. The adaptive model continues adaptation re-initialized with this fixed model.

Besides using a cost function for switching, one can also utilize a result from [54] which states the stability condition for a dynamic system with multiple Lyapunov functions.

Theorem 6.1 [54] *Suppose that there are candidate Lyapunov functions V_j , $j = 1, \dots, N$, for the system defined in (4.1) based on multiple parameter vectors $\hat{\theta}_j$, $j = 1, \dots, N$, chosen from a compact parameter set $\mathcal{S} \subset \mathbb{R}^p$. Let T_i be a permissible switching se-*

quence defined in Definition 6.1 associated with the system. If,

$$V_j(x(T_{i+1})) < V_j(x(T_i))$$

holds for all j , then the switching system is stable in sense of Lyapunov.

Proof: The proof of this theorem can be found in [54]. □

The Lyapunov functions used here are the ones which are formed for the indirect identification dynamics defined in (5.9) for an observer-based identification, or as in (5.19) for a filtering-based identification. Denoting these identification Lyapunov functions for each model V_j , and assuming N fixed models in the identification process, the changes in the multiple Lyapunov functions associated with the identification dynamics are monitored, and the switching logic is based on the following condition,

$$\Delta V_j = V_j(t + T_{min}) - V(t) < 0 \quad , \quad j = 1, \dots, N \quad (6.3)$$

where $V(t)$ is the Lyapunov function of the identification dynamics of the adaptive model, and $V_j(t + T_{min})$ is the Lyapunov function of the j^{th} fixed model. If $\Delta V_j(t) \leq -\kappa$ for some $\kappa > 0$ is satisfied, then the parameter estimates of the adaptive model are reset to $\hat{\theta}_j$, and adaptation is restarted from that point. The positive threshold κ provides some sort of hysteresis to reduce the number of switchings. If more than one model makes $\Delta V_j(t) \leq -\kappa$, then the model providing the largest κ value is selected. If no models gives a negative jump in the Lyapunov function, then no switching takes place, and adaptation continues without any resets.

Definition 6.3 *With N fixed identification models and with a permissible switching time T_{min} given in Definition 6.1, the switching logic is given as,*

$$j^* = \{j : \{\Delta V_j(t) = V_j(t + T_{min}) - V(t) < 0\} \cap \{\min \Delta V_j\}\} \quad , \quad j = 1, \dots, N \quad (6.4)$$

where j^* defines the index for the chosen identification model. $V(t)$ is the Lyapunov function for the update dynamics of the adaptive model, $V_j(t + T_{min})$ is the same Lyapunov function evaluated with fixed parameter estimates at the switching instants.

We now apply the multiple models and switching logic to the indirect adaptive controller derived in Section 5.4.

6.3. Indirect Adaptive Control Using Multiple Models without Matching Conditions

In this section we extend the indirect adaptive system derived in Section 5.3 for the case of multiple fixed models [29]. First we modify the identification scheme defined in Section 5.3 such that the parameter update equation in (5.68) re-initializes itself with a fixed model which would generate a negative jump in the Lyapunov function defined in (5.19) for the update process. This would in turn speed up the convergence of the parameter estimates, while preserving the stability of the update mechanism. Here we utilize a result from [54] which states the stability condition for a dynamic system with multiple Lyapunov functions, as given in Theorem 6.1. Assuming N fixed models in the identification process, we monitor the changes in the multiple Lyapunov functions associated with the identification dynamics and the switching logic is based on the following condition,

$$\Delta V_j(t) = V_j(t + T_{min}) - V(t) < 0 \quad , \quad j = 1, \dots, N \quad (6.5)$$

where

$$V(t) = \frac{1}{2} \tilde{\theta}^T(t) \tilde{\theta}(t) \quad (6.6)$$

and

$$V_j(t + T_{min}) = \frac{1}{2} \tilde{\theta}_j^T \tilde{\theta}_j \quad (6.7)$$

If $\Delta V_j(t) \leq -\kappa$ for some $\kappa > 0$ is satisfied, then the parameter estimates of the adaptive model are reset to $\hat{\theta}_j$, and adaptation is restarted from that point. If more than one model makes $\Delta V_j(t) \leq -\kappa$, then the model providing the largest κ value is selected. If no models gives a negative jump in the Lyapunov function, then no switching takes place, and adaptation continues without any resets.

For simplicity, we will denote $(t + T_{min})$ by (t^+) . Then, using the Lyapunov function for the identification dynamics given in (5.19), we have,

$$\Delta V_j(t) = \frac{1}{2} \tilde{\theta}_j^T(t^+) \tilde{\theta}_j(t^+) - \frac{1}{2} \tilde{\theta}^T(t) \tilde{\theta}(t) \quad (6.8)$$

Using $\tilde{\theta} = \hat{\theta} - \theta$ and $\tilde{\theta}_j = \hat{\theta}_j - \theta$, after some manipulation we can write (6.8) as,

$$\Delta V_j(t) = -\frac{1}{2} \left(\hat{\theta}_j(t^+) - \hat{\theta}(t) \right)^T \left(\hat{\theta}_j(t^+) - \hat{\theta}(t) \right) + \left(\hat{\theta}_j(t^+) - \hat{\theta}(t) \right)^T \left(\hat{\theta}_j(t^+) - \theta \right) \quad (6.9)$$

where θ is the actual parameter value which is not available for computation. Here we use a similar approach given in [27] to evaluate (6.9). Since $\hat{\theta}_j$'s are fixed, the above equation is written as,

$$\Delta V_j(t) = -\frac{1}{2} \left(\hat{\theta}_j - \hat{\theta}(t) \right)^T \left(\hat{\theta}_j - \hat{\theta}(t) \right) + \left(\hat{\theta}_j - \hat{\theta}(t) \right)^T \left(\hat{\theta}_j - \theta \right) \quad (6.10)$$

Using the identification error given in (5.17), we obtain,

$$\underbrace{\int_{t-\delta}^t w_f(\tau) \tilde{x}_j(\tau) d\tau}_{\Upsilon_j(t)} = \underbrace{\int_{t-\delta}^t w_f(\tau) w_f^T(\tau) d\tau}_{\Psi(t)} \cdot (\hat{\theta}_j - \theta) \quad (6.11)$$

where $\delta > 0$ is a constant time window. Pulling $(\hat{\theta}_j - \theta)$ from (6.11), ΔV_j can now be computed as,

$$\Delta V_j(t) = -\frac{1}{2} \left(\hat{\theta}_j - \hat{\theta}(t) \right)^T \left(\hat{\theta}_j - \hat{\theta}(t) \right) + \left(\hat{\theta}_j - \hat{\theta}(t) \right)^T \Psi^{-1}(t) \Upsilon_j(t) \quad (6.12)$$

Remark 6.1 *Note that invertibility condition on Ψ corresponds to the persistent excitation condition given in (5.24). Since we assume persistently exciting regressor, Ψ^{-1} will be available for computation.*

Now, having ΔV_j available for all N fixed models, we state the switching logic in the following definition.

Definition 6.4 *With N fixed identification models defined with the error equations, (as in (5.17)),*

$$\tilde{x}_j = w_f^T \tilde{\theta}_j \quad (6.13)$$

where $\tilde{\theta}_j = \hat{\theta}_j - \theta$ with $\hat{\theta}_j$, $j = 1, \dots, N$, being the fixed parameter estimates, and with a permissible switching time T_{min} given in Definition 6.1, the switching logic which would increase parameter convergence while preserving stability of the identification dynamics defined in (5.68) is given as,

$$j^* = \{j : \{\Delta V_j = V_j(t^+) - V(t) < 0\} \bigcap \{\min \Delta V_j\}\} \quad , \quad j = 1, \dots, N$$

where j^* gives the index for the chosen identification model. $V(t)$ is the Lyapunov function for the update dynamics given in (5.19), $V_j(t^+)$ is the same Lyapunov function evaluated with fixed parameter estimates at the switching instants. In the above equation, $(t + T_{min})$ is denoted by (t^+) for simplicity.

Under this switching rule, the multi-linear parameter estimate vector $\hat{\Theta}$, hence the parameter error vector $\tilde{\Theta}$ defined in (5.75)-(5.76), will be in the form of,

$$\hat{\Theta}(t) = \hat{\Theta}_{z(t)} \quad (6.14)$$

$$\tilde{\Theta}(t) = \tilde{\Theta}_{z(t)} \quad (6.15)$$

where $z(t)$ is defined as $z(t) : \mathbb{R}_+ \mapsto 0, \dots, N$ such that if $t \in [T_i, T_{i+1})$ for some $i < \infty$, then $z(t) = z(T_i)$. Here, T_i is a permissible switching sequence defined in Definition 6.1.

With this switching scheme in place for identification, we now have to establish the closed loop stability of the error dynamics defined in (5.97).

Theorem 6.2 *Consider the indirect adaptive scheme with the control signal given in (5.87) and the adaptation dynamics given in (5.68). With N fixed identification models defined in (6.13) and the switching logic for parameter update dynamics defined in Definition 6.4, the closed loop error dynamics is asymptotically stable, such that $y \rightarrow y_r$ as $t \rightarrow \infty$, with enhanced transient performance.*

Proof: In the proposed system, we have N fixed models and one adaptive model which switches between the fixed parameter estimates. Hence the control signal in (5.87) and the dynamics $\dot{\xi}_\gamma$ in (5.95) are modified with the new parameter estimate vector $\hat{\Theta}_{z(t)}$, which is defined in (6.14). The certainty equivalence control “ u ” will have the form,

$$u = \frac{1}{\mathcal{G}^T(x)\hat{\Theta}_{z(t)}}(-\mathcal{F}^T(x)\hat{\Theta}_{z(t)} + \hat{v}) \quad (6.16)$$

Hence, $\dot{\xi}_\gamma$ can be written as,

$$\dot{\xi}_\gamma = v + \varphi^T(x, u)\tilde{\Theta}_{z(t)} \quad (6.17)$$

Then, the resulting closed loop error dynamics will be the same as before, with parameter error vector $\tilde{\Theta}$ being reset at predetermined switching instants,

$$\left. \begin{aligned} \xi &= e + r \\ \dot{e} &= Ae + \varphi^T(\xi, \eta, u)\tilde{\Theta}_{z(t)} \\ \dot{\eta} &= q(\xi, \eta) \end{aligned} \right\} \quad (6.18)$$

where $r = \begin{bmatrix} y_r & \dot{y}_r & \cdots & y_r^{(\gamma-1)} \end{bmatrix}^T$. We know from Theorem 5.3 that using the adaptive model alone, without switching, will result in stable asymptotic tracking. A Lyapunov function for the closed loop dynamics can be formed as,

$$V_c(e, \eta) = e^T P e + \mu V_q(\eta) \quad , \quad \mu \in \mathbb{R}_+ \quad (6.19)$$

Using previous derivations in (5.101) through (5.109), we obtain,

$$\dot{V}_c(e, \eta) \leq - \begin{bmatrix} \|e\| \\ \|\eta\| \end{bmatrix}^T Q_1 \begin{bmatrix} \|e\| \\ \|\eta\| \end{bmatrix} + (b_1 b_r \|\tilde{\Theta}_{z(t)}\|)^2 + \frac{\mu}{c_3} (c_4 b_q b_r)^2 \quad (6.20)$$

where

$$Q_1 = \begin{bmatrix} \frac{3}{4} - b_1 \|\tilde{\Theta}_{z(t)}\| & -\frac{1}{2}(b_1 \|\tilde{\Theta}_{z(t)}\| + \mu c_4 b_q) \\ -\frac{1}{2}(b_1 \|\tilde{\Theta}_{z(t)}\| + \mu c_4 b_q) & \frac{3}{4} \mu c_3 \end{bmatrix} \quad (6.21)$$

If,

1. $\mu \leq \frac{3c_3}{4(c_4 b_q)^2}$, and
2. $\|\tilde{\Theta}_{z(t)}\| \leq \frac{1}{2b_1}$

then the matrix Q_1 will be positive definite, as discussed in the proof of Theorem 5.3. The first condition depends only on the static design parameters and can be satisfied easily. One assumption of Theorem 5.3 is the persistently exciting reference signal, hence it is assumed that $\tilde{\Theta} \rightarrow 0$ as $t \rightarrow \infty$. Therefore, after some time T_1 , we will have,

$$\|\tilde{\Theta}_{z(t)}\| \leq \frac{1}{2b_1} \quad , \quad \forall t \geq T_1 \quad (6.22)$$

Since $\tilde{\Theta}$ is defined as the multi-linear parameter error vector, we will have,

$$\|\tilde{\Theta}_{z(t^+)}\| < \|\tilde{\Theta}_{z(t)}\| \quad (6.23)$$

whenever $\|\tilde{\theta}(t^+)\| < \|\tilde{\theta}(t)\|$. Let us suppose, a switching occurs at time t^+ , and the parameter estimate vector $\hat{\theta}$ is reset to $\hat{\theta}_j$ based on the switching logic given in Definition 6.4. This can happen only if $\Delta V_j \leq -\kappa$ for some $\kappa \in \mathbb{R}_+$, where ΔV_j is computed using (6.12). Then, from (6.8) we have,

$$\Delta V_j(t) = \frac{1}{2}\tilde{\theta}_j^T(t^+)\tilde{\theta}_j(t^+) - \frac{1}{2}\tilde{\theta}^T(t)\tilde{\theta}(t) \leq -\kappa \quad (6.24)$$

which is equivalent to,

$$\tilde{\theta}_j^T(t^+)\tilde{\theta}_j(t^+) + 2\kappa \leq \tilde{\theta}^T(t)\tilde{\theta}(t) \quad (6.25)$$

$$(\hat{\theta}_j - \theta)^T(\hat{\theta}_j - \theta) + 2\kappa \leq (\hat{\theta}(t) - \theta)^T(\hat{\theta}(t) - \theta) \quad (6.26)$$

$$\|\hat{\theta}_j - \theta\|^2 + 2\kappa \leq \|\hat{\theta}(t) - \theta\|^2 \quad (6.27)$$

Consequently, we get,

$$\|\hat{\theta}_j - \theta\| < \|\hat{\theta}(t) - \theta\| \quad (6.28)$$

and

$$\|\tilde{\theta}(t^+)\| < \|\tilde{\theta}(t)\| \quad (6.29)$$

Based on (6.29), we obtain,

$$\|\tilde{\Theta}_{z(t^+)}\| < \|\tilde{\Theta}_{z(t)}\| \quad (6.30)$$

which means that the multi-linear parameter error vector $\tilde{\Theta}_{z(t)}$ decreases with each

reset. Then, using (6.22) and (6.30) we can write,

$$\|\tilde{\Theta}_{z(t^+)}\| \leq \frac{1}{2b_1} \quad , \quad \forall t^+ \geq T_2 \quad (6.31)$$

for some T_2 , such that $T_2 < T_1$. On the other hand, for the term $(b_1 b_r \|\tilde{\Theta}_{z(t)}\|)^2$ in (6.20), we have,

$$(b_1 b_r \|\tilde{\Theta}_{z(t^+)}\|)^2 < (b_1 b_r \|\tilde{\Theta}_{z(t)}\|)^2 \quad (6.32)$$

Since the switching can only speed up the convergence of $\tilde{\Theta}$, it clearly decreases the time required for positive definiteness of Q_1 . Examining (6.20), (6.31) and (6.32), it is seen that the switching leads faster decay and convergence of the closed loop Lyapunov function to zero; hence, faster convergence of the tracking error to zero. Therefore the transient response is improved. $\square$

The block diagram of the proposed indirect adaptive control system with N fixed identification models and one adaptive model is displayed in Figure 6.5. Next, we test the performance of the proposed scheme using a simulation study. The nonlinear system chosen for this test is the same as in [9], and the multiple model based scheme is applied utilizing the Lyapunov-function based switching. Another simulation study with this scheme can be found in [55], where a 1-link robot arm driven by a brushed DC-motor system is simulated using the performance-index based switching.

6.3.1. Simulation Study

In this section we would like to verify the theoretical results with computer simulation. We use the same nonlinear system which was used in [9], for comparison

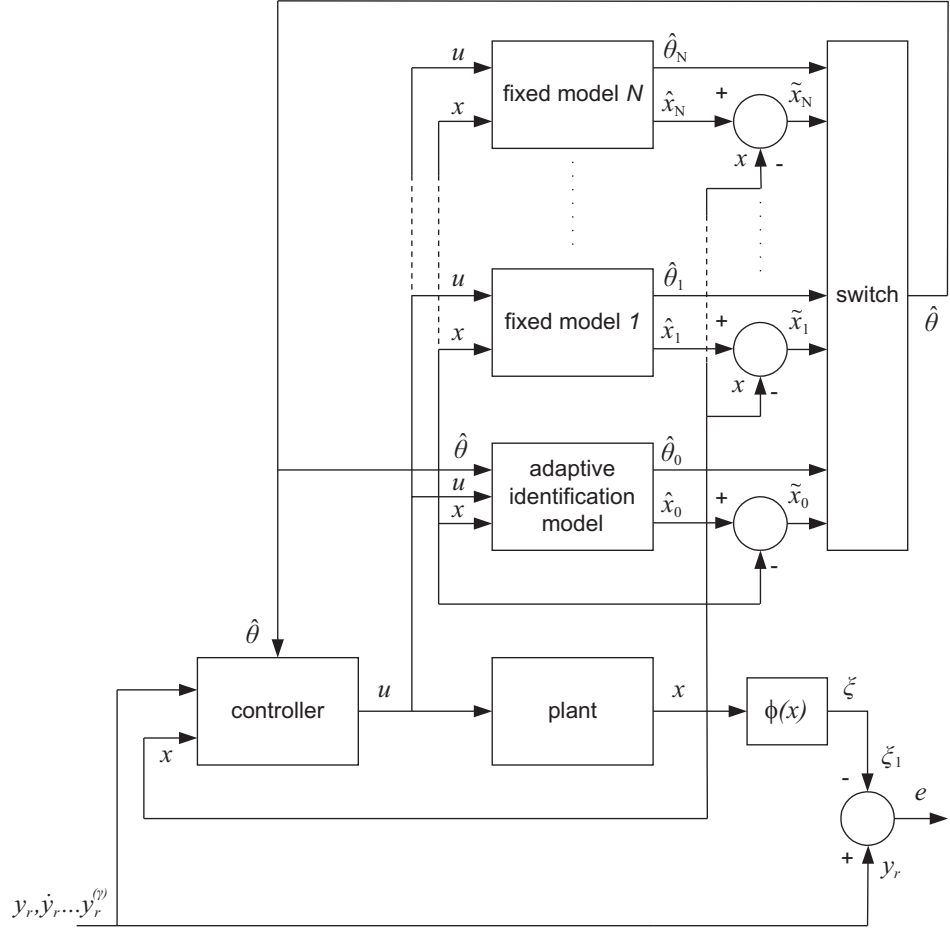


Figure 6.5. Block diagram of the indirect adaptive controller with N fixed models and one adaptive model.

purposes. The system dynamics is given as,

$$\left. \begin{aligned} \dot{x}_1 &= x_2 + p x_1(10 + \sin x_1) \\ \dot{x}_2 &= u \\ y &= x_1 \end{aligned} \right\} \quad (6.33)$$

where $p \in \mathbb{R}$ is the unknown parameter. Relative degree of the system is two. The linearizing transformation is obtained using the descriptions given in Section 3.2.3 as,

$$\begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 + p x_1(10 + \sin x_1) \end{bmatrix} \quad (6.34)$$

The linearizing control law is given as,

$$u = \hat{v} - \hat{p}(10 + \sin x_1 + x_1 \cos x_1) (x_2 + \hat{p}x_1(10 + \sin x_1)) \quad (6.35)$$

The tracking control is obtained as,

$$v = \ddot{y}_r + \lambda_1(\dot{y}_r - \dot{y}) + \lambda_2(y_r - y) \quad (6.36)$$

where y_r is the reference output. However, $\dot{y} = L_f h$ is not measurable, since $L_f h = x_2 + px_1(10 + \sin x_1)$ where p is the unknown parameter. Therefore we use the estimate of p and calculate $\hat{v}$ in (6.35) as,

$$\hat{v} = \ddot{y}_r + \lambda_1[\dot{y}_r - \underbrace{(x_2 + \hat{p}x_1(10 + \sin x_1))}_{L_f h}] + \lambda_2(y_r - y) \quad (6.37)$$

$\lambda_1 = 30$ and $\lambda_2 = 200$ were used in the simulation. For system identification, the system dynamics is written in the regressor form as,

$$\dot{x} = w^T \theta = \begin{bmatrix} x_2 & x_1(10 + \sin x_1) \\ u & 0 \end{bmatrix} \begin{bmatrix} 1 \\ p \end{bmatrix} \quad (6.38)$$

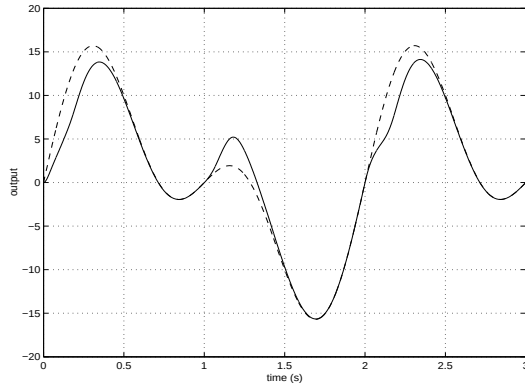
Filtered forms of x and w , namely x_f and w_f , were obtained using (5.14) with $\sigma = 0.06$. The parameter update law is given as $\dot{\hat{\theta}} = -\alpha w_f^T \tilde{x}$ where $\tilde{x} = w_f \tilde{\theta}$ and α was set to 10. The reference trajectory was generated as $y_r = 10 \sin \pi t + 8 \sin 2\pi t$.

Initial estimate of the unknown parameter “ p ” in the system was chosen as $\hat{p}(0) = 2$. The following scenario is assumed. Actual parameter value for p was set to 1 for $0 < t < 1$ sec. At $t = 1$ sec., the parameter p was suddenly changed to 2, and at $t = 2$ sec. the parameter was again changed to 1 to observe the adaptation performance of the control system. We used 6 fixed models, $\hat{\theta}_j \in \{0.9, 1.2, 1.4, 1.6, 1.8, 2.1\}$, for simulation of the proposed scheme. The integration time in (6.11) was $\delta = 0.01$ seconds. The switching threshold for ΔV was $\kappa = 0.3$.

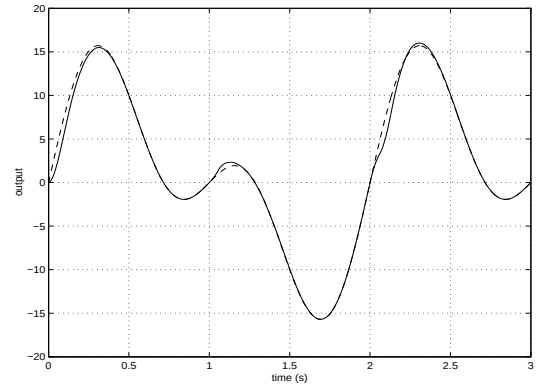
The nonlinear system in (6.33) was first simulated using indirect adaptive control scheme with single identification model. Figure 6.6a displays the output of the system. The parameter changes at $t = 1$ sec. and $t = 2$ sec. cause the system output noticeably deviate from the reference trajectory. Figure 6.6e displays the parameter adaptation for single model indirect adaptive control.

The same scenario was then simulated using the proposed scheme, where multiple fixed identification models are used. In Figure 6.6b, the output of the system is shown. It is clearly seen that the parameter changes at $t = 1$ sec. and $t = 2$ sec. cause only relatively small deviations from the reference. Parameter adaptation under multiple models scheme is shown in Figure 6.6f. Effects of switching are noticed as the jumps in the value of parameter estimate $\hat{p}$.

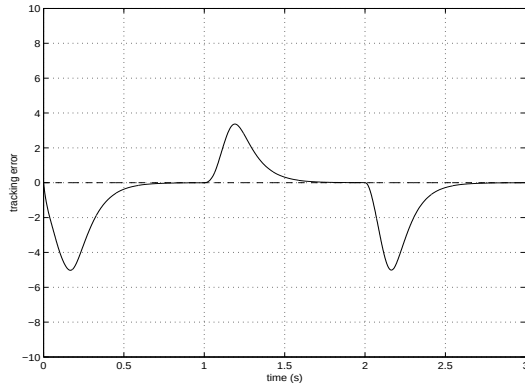
Figure 6.6c and 6.6d provide the tracking errors of single model and multiple model schemes, respectively. It is observed that the proposed scheme leads to much smaller transient errors. Consequently, the tracking errors converge to zero faster for the proposed scheme. This verifies the theoretical results obtained in Section 6.3. Figure 6.7 displays the switching between the identification models. The index “0” corresponds to the adaptive model in Figure 6.7. We see that the multiple model scheme switched only three times to one of the fixed models. The first switching occurs at the very beginning. This in turn greatly reduces the initial tracking error. The other two switchings take place following the parameter changes at $t = 1$ sec. and $t = 2$ sec. The value of the threshold κ has important effect on the time and number of switchings. In our simulation, we have experimentally tuned κ , so that number of switchings is reduced while maintaining a performance increase in transient errors.



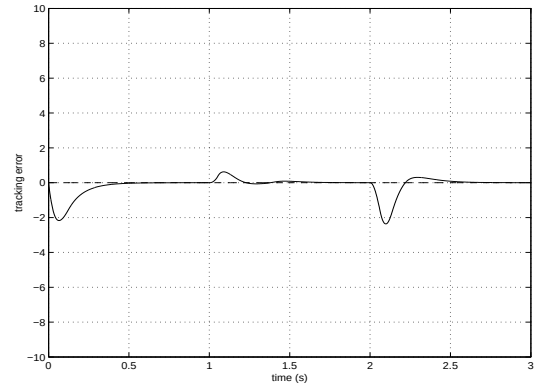
(a) Output y (solid) and y_r (dashed)
(single model)



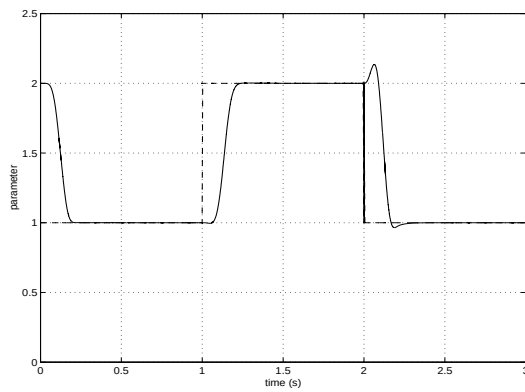
(b) Output y (solid) and y_r (dashed)
(multiple models)



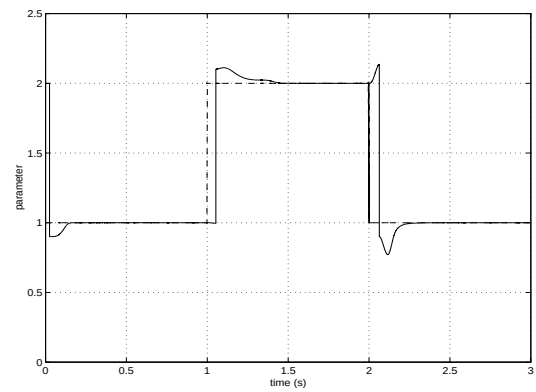
(c) Tracking error e
(single model)



(d) Tracking error e
(multiple models)



(e) Parameter p (dashed), $\hat{p}$ (solid)
(single model)



(f) Parameter p (dashed), $\hat{p}$ (solid)
(multiple models)

Figure 6.6. Simulation results for single model and multiple model cases of the indirect adaptive control scheme given in Section 6.3.1.

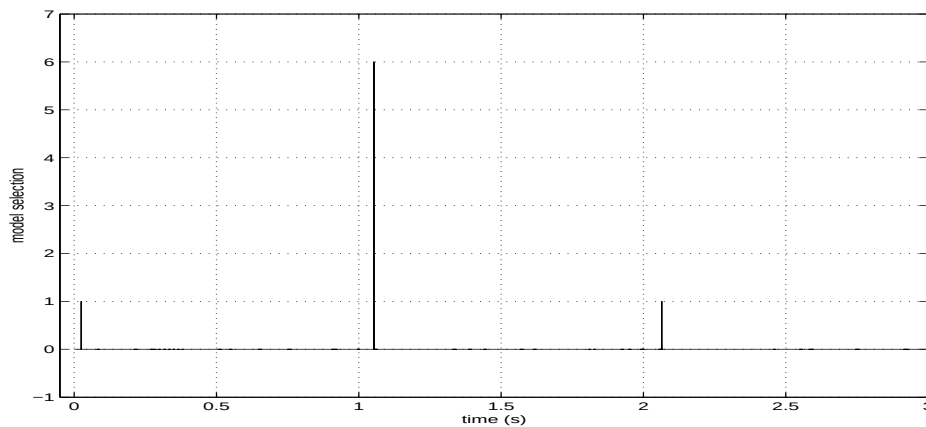


Figure 6.7. Switching model index (0: adaptive model, 1-6: fixed models) of the indirect adaptive control scheme of Section 6.3.1.

6.4. Summary

Transient performance of the adaptive controllers for nonlinear systems are usually inadequate when the parameter estimates are far from their actual values. The idea of using multiple identification models and switching was proposed to improve the transient performance. In a multiple model system, there are a set of fixed and/or adaptive models, one resettable adaptive model for the actual system and a switching logic. The switching logic is based on the minimization of a cost function or a maximum negative jump in the Lyapunov function for the identification dynamics and selects the best model to be used in the controller at each resetting time. The parameter estimates of the adaptive model are reset to parameter values of the selected model. The adaptation continues with this re-initialization until another switching takes place. This method provides faster parameter convergence, and as a result, improves the transient response.

This chapter presented a novel methodology for the indirect adaptive control of nonlinear systems utilizing multiple identification models to improve the transient performance. The indirect adaptive controller derived in Chapter 5 was extended to a multiple model based scheme utilizing N fixed models and one adaptive model. The theoretical results showing the improvement in the transient responses were verified by computer simulations.

7. DIRECT ADAPTIVE METHODS USING MULTIPLE MODELS

The multiple models and switching methodology for transient improvement in adaptive systems was first developed for the indirect adaptive systems. In an indirect adaptive scheme, the parameters of the system are identified separately and explicitly. The controller parameters are then calculated using these estimates of the system parameters. The indirect scheme was studied in Section 5.3, and its multiple models extension was given in Chapter 6 for a class of nonlinear systems.

In this chapter, we will investigate direct adaptive controllers utilizing multiple identification models with the objective of transient performance enhancement. First, we will modify the MRAC system given in Section 5.4 by incorporating multiple models for better transient performance. Then we will apply the multiple models and switching methodology to the direct adaptive scheme derived in Section 5.5.

7.1. Direct Adaptive Control Using Multiple Models under Matching Conditions

A model reference adaptive controller (MRAC) for a nonlinear system satisfying the matching conditions had been derived in Section 5.4. Now we modify this adaptive control scheme by incorporating N fixed models and a switching logic to improve its transient performance [30].

Using (6.3) and the Lyapunov function for the closed loop system defined in Section 5.4, the change in the Lyapunov function at the switching instants is given as,

$$\Delta V = V(t^+) - V(t) = V(e(t^+), \tilde{\theta}(t^+)) - V(e(t), \tilde{\theta}(t)) \quad (7.1)$$

where $V(t^+)$ is used in place of $V(t + T_{min})$ for brevity of notation. Using the Lyapunov

function defined in (5.135), we get,

$$\Delta V = \underbrace{e^T(t^+)Pe(t^+) + (\hat{\theta}(t^+) - \theta)^T \Gamma^{-1}(\hat{\theta}(t^+) - \theta)}_{V(t^+)} - \underbrace{e^T(t)Pe(t) - (\hat{\theta}(t) - \theta)^T \Gamma^{-1}(\hat{\theta}(t) - \theta)}_{-V(t)} \quad (7.2)$$

Expanding the terms in the parentheses we get,

$$\begin{aligned} \Delta V = & (e^T(t^+)Pe(t^+) - e^T(t)Pe(t)) + \left(\hat{\theta}^T(t^+)\Gamma^{-1}\hat{\theta}(t^+) - 2\hat{\theta}^T(t^+)\Gamma^{-1}\theta + \theta^T\Gamma^{-1}\theta \right) \\ & + \left(\hat{\theta}^T(t)\Gamma^{-1}\hat{\theta}(t) - 2\hat{\theta}^T(t)\Gamma^{-1}\theta + \theta^T\Gamma^{-1}\theta \right) \end{aligned} \quad (7.3)$$

Adding and subtracting the terms $2\hat{\theta}^T(t^+)\Gamma^{-1}\hat{\theta}(t)$ and $\hat{\theta}^T(t^+)\Gamma^{-1}\hat{\theta}(t^+)$ to the above equation to complete the quadratic forms, we obtain,

$$\begin{aligned} \Delta V = & (e^T(t^+)Pe(t^+) - e^T(t)Pe(t)) - \left(\hat{\theta}^T(t^+)\Gamma^{-1}\hat{\theta}(t^+) - 2\hat{\theta}^T(t^+)\Gamma^{-1}\hat{\theta}(t) + \hat{\theta}^T(t)\Gamma^{-1}\hat{\theta}(t) \right) \\ & + 2 \left(\hat{\theta}^T(t^+)\Gamma^{-1}\hat{\theta}(t^+) - \hat{\theta}^T(t)\Gamma^{-1}\hat{\theta}(t) - \hat{\theta}^T(t^+)\Gamma^{-1}\theta + \hat{\theta}^T(t)\Gamma^{-1}\theta \right) \end{aligned} \quad (7.4)$$

which leads to,

$$\begin{aligned} \Delta V = & (e^T(t^+)Pe(t^+) - e^T(t)Pe(t)) - \left((\hat{\theta}(t^+) - \hat{\theta}(t))^T \Gamma^{-1}(\hat{\theta}(t^+) - \hat{\theta}(t)) \right) \\ & + 2 \left((\hat{\theta}(t^+) - \hat{\theta}(t))^T \Gamma^{-1}(\hat{\theta}(t^+) - \theta) \right) \end{aligned} \quad (7.5)$$

We will use (7.5) to compute ΔV_j for each fixed model with the associated parameter vector $\hat{\theta}_j$. Since θ is unknown in the above algorithm, we have to invoke an estimation algorithm. The original system dynamics defined in (5.114) can be written in the regressor form as,

$$\dot{x} = w^T(x, u)\theta \quad (7.6)$$

where the regressor $w \in \mathbb{R}^{n \times (p+1)}$ is,

$$w = \begin{bmatrix} f(x) + g(x)u \\ q_1(x) \\ \vdots \\ q_p(x) \end{bmatrix} \quad (7.7)$$

and the parameter vector is,

$$\theta = \begin{bmatrix} 1 \\ \theta_1 \\ \vdots \\ \theta_p \end{bmatrix} \quad (7.8)$$

We assume N fixed models for the given system based on a given set of parameter vectors $\hat{\theta}_j \in \mathbb{R}^{p+1}$, $j = 1, \dots, N$, where each parameter estimate vector $\hat{\theta}_j$ is defined as,

$$\hat{\theta}_j = \begin{bmatrix} 1 \\ \hat{\theta}_{j_1} \\ \vdots \\ \hat{\theta}_{j_p} \end{bmatrix} \quad (7.9)$$

Then these N model estimates lead to the following dynamic models,

$$\dot{\hat{x}}_j = w^T(x, u)\hat{\theta}_j \quad , \quad j = 1, \dots, N \quad (7.10)$$

where $\hat{x}_j$ is the state estimate for the j^{th} model. Denoting the filtered forms of x and w as x_f and w_f respectively, we apply the identification method given in Section 5.1. Assuming a small positive constant σ for the filters, we can build the state estimates based on the given fixed models as,

$$\hat{x}_j = w_f^T \hat{\theta}_j + x_f \quad (7.11)$$

Then we define the observation error for each model as,

$$\tilde{x}_j = \hat{x}_j - x \quad (7.12)$$

Here, we note that observation error $\tilde{x}_j$ is readily available for each model since “ x ” is also available based on our assumption. Using (7.11) and available state information x , we get N error models as,

$$\tilde{x}_j(t) = w_f^T(x, u)\tilde{\theta}_j = w_f^T(x, u)(\hat{\theta}_j - \theta) \quad , \quad j = 1, \dots, N \quad (7.13)$$

In order to utilize ΔV_j in (7.5) to enforce the switching condition, we need to compute the term $(\hat{\theta}_j - \theta)$ for each model at the switching instants. We multiply both sides of (7.13) by $w_f(t)$ and integrate over a time period δ , duration of which should be determined in the implementation. We get,

$$\underbrace{\int_{t-\delta}^t w_f(\tau)\tilde{x}_j(\tau)d\tau}_{\Upsilon_j(t)} = \underbrace{\int_{t-\delta}^t w_f(\tau)w_f^T(\tau)d\tau}_{\Psi(t)} \cdot (\hat{\theta}_j - \theta) \quad (7.14)$$

where we can take the parameter error term $(\hat{\theta}_j - \theta)$ out of the integral, since the parameter errors are constant for each fixed model. Assuming we have persistency of excitation such that Ψ is nonsingular, we can write (7.14) as,

$$(\hat{\theta}_j - \theta) = \Psi^{-1}(t)\Upsilon_j(t) \quad (7.15)$$

Plugging (7.15) in (7.5), we get,

$$\begin{aligned} \Delta V_j = & (e^T(t^+)Pe(t^+) - e^T(t)Pe(t)) - \left((\hat{\theta}(t^+) - \hat{\theta}(t))^T \Gamma^{-1}(\hat{\theta}(t^+) - \hat{\theta}(t)) \right) \\ & + 2 \left((\hat{\theta}(t^+) - \hat{\theta}(t))^T \Gamma^{-1} \Psi^{-1}(t) \Upsilon_j(t) \right) \quad , \quad j = 1, \dots, N \end{aligned} \quad (7.16)$$

Theorem 7.1 *For the nonlinear system given in (5.114), the control input (defined in (5.126) and (5.133) along with the parameter update equation in (5.138)) results in a*

stable system having

$$\lim_{t \rightarrow \infty} |x(t) - \phi^{-1}(\xi_r)| = 0 \quad (7.17)$$

with improved transient performance, when the parameter estimate vector $\hat{\theta}(t)$ in (5.138) is reset to $\hat{\theta}_{j^*}$ at the switching instants where $j^* \in \{1, \dots, N\}$ denotes the chosen model index and is defined with the switching logic as,

$$j^* = \{j : \{\Delta V_j = V_j(t + T_{min}) - V(t) < 0\} \bigcap \{\min \Delta V_j\}\} \quad , \quad j = 1, \dots, N \quad (7.18)$$

Proof: The control input and the parameter update equation is the same for the closed loop system, hence we have the Lyapunov equation as defined in (5.135),

$$V(e, \tilde{\theta}) = e^T P e + \tilde{\theta}^T \Gamma^{-1} \tilde{\theta} \quad (7.19)$$

Since the parameter estimate vector $\hat{\theta}(t)$ can take values from a fixed parameter set $\{\hat{\theta}_j, j = 1, \dots, N\}$ at the switching instants, we compute the change in the Lyapunov function as,

$$\Delta V_j = e^T(t^+) P e(t^+) + \tilde{\theta}_j^T(t^+) \Gamma^{-1} \tilde{\theta}_j(t^+) - e^T(t) P e(t) - \tilde{\theta}^T(t) \Gamma^{-1} \tilde{\theta}(t) \quad (7.20)$$

which is equivalent to,

$$\begin{aligned} \Delta V_j = & (e^T(t^+) P e(t^+) - e^T(t) P e(t)) \\ & - \left((\hat{\theta}(t^+) - \hat{\theta}(t))^T \Gamma^{-1} \left((\hat{\theta}(t^+) - \hat{\theta}(t)) - 2\Psi^{-1}(t) \Upsilon_j(t) \right) \right) \end{aligned} \quad (7.21)$$

where $\Upsilon_j(t)$ and $\Psi(t)$ are as defined in (7.14).

Now assuming we have persistency of excitation for the reference input leading to existence of $\Psi^{-1}(t)$, we compute ΔV_j for all j . Based on the Theorem 6.1, the

performance improving and stability preserving condition is given as,

$$\Delta V_j < 0 \quad , \quad j = 1, \dots, N \quad (7.22)$$

The model index j^* chosen by the switching logic in (7.18) gives us the largest negative ΔV_j . This condition actually causes a maximum negative jump in the Lyapunov function in (5.135) which also leads to,

$$\dot{V}(t) \leq -e^T Q e \quad (7.23)$$

Due to (7.23), we have $e \in \mathcal{L}_2 \cap \mathcal{L}_\infty$ and using similar arguments as in the proof of Theorem 5.4, we get $e \rightarrow 0$ as $t \rightarrow \infty$ which leads to $\lim_{t \rightarrow \infty} |x(t) - \phi^{-1}(\xi_r)| = 0$ as well. This result basically demonstrates that while overall stability is preserved with switching, there is transient performance improvement for the closed loop system due to (7.23) $\square$

Remark 7.1 *Since we require non-singularity for $\Psi(t)$, this is simply equivalent to persistent excitation for the reference input which also leads to the asymptotic convergence of $\tilde{\theta}$. The resetting condition defined in (7.18) simply speeds up convergence of $\hat{\theta}$ to θ and hence leading to lower transient errors.*

Remark 7.2 *If none of the N fixed models results in a negative jump in the Lyapunov function (5.135), then there is no switching and adaptation continues with the adaptive model.*

Hence we have shown the stability and transient performance improvement of direct adaptive control system under matching conditions using multiple models and switching. In the following section, we apply the multiple models scheme for the direct adaptive control of a more general class of nonlinear systems.

7.2. Direct Adaptive Control Using Multiple Models without Matching Conditions

In Section 5.5, a direct adaptive control scheme was derived for a class of SISO nonlinear systems in the form of (5.142), following a similar methodology as in [8]. In this section, we extend this control scheme to include multiple models with the same parametrization as (5.142), for faster parameter convergence.

In the case of multiple identification models, the best estimates are selected based on a selection criterion, and the adaptive model is reset to these best values. This idea was formulated for the direct adaptive schemes, where matching parametric uncertainties are assumed [30]. Now we further extend this methodology to a more general class of nonlinear systems without requiring the strict matching conditions [31].

To do this, we first design an identifier as explained in Section 5.1.2, which runs in parallel with the system and the direct controller. We first write the system defined in (5.142) in the regressor form as,

$$\dot{x} = w^T(x, u)\theta \quad (7.24)$$

with w being the regressor matrix and $\theta \in \mathcal{S} \subset \mathbb{R}^p$ is the parameter vector. Let x_f and w_f be the filtered forms of x and w in (7.24), respectively, defined as,

$$\left. \begin{aligned} \sigma \dot{x}_f &= -x_f + x \\ \sigma \dot{w}_f^T &= -w_f^T + \sigma w^T(x, u) \end{aligned} \right\} \quad (7.25)$$

where $\sigma \in \mathbb{R}_+$ is the time constant of a 1st-order filter which is used for eliminating the derivative term in (7.24) and obtaining a linear error equation. The state x can be reconstructed as,

$$x = w_f^T \theta + x_f + [x(0) - w_f^T(0)\theta - x_f(0)]e^{-t/\sigma} \quad (7.26)$$

It is straightforward to verify the equivalence of this representation by taking the time derivative of (7.26) to obtain (7.24). Assuming that σ is chosen small enough, the steady state estimate for the system is obtained as,

$$\hat{x} = w_f^T \hat{\theta} + x_f \quad (7.27)$$

where $\hat{\theta}$ is the parameter vector obtained by the identifier in (5.171). For the adaptive system given in Section 5.5, although Θ is the over-parameterized vector of the system parameters in θ , one can easily extract $\hat{\theta}$ from $\hat{\Theta}$. For the N fixed models, using the filtered forms x_f and w_f , we can write,

$$\hat{x}_j = w_f^T \hat{\theta}_j + x_f \quad , \quad j = 1, \dots, N \quad (7.28)$$

where $\hat{\theta}_j \in \mathcal{S} \subset \mathbb{R}^p$ is the fixed parameter vector of the j^{th} -model. Now we define the state prediction error for the adaptive model as,

$$\tilde{x} \stackrel{def}{=} \hat{x} - x \quad (7.29)$$

where $\hat{x}$ is obtained using (7.27) and x is measurable. For the fixed models, we calculate the state prediction error similarly as,

$$\tilde{x}_j = \hat{x}_j - x \quad , \quad j = 1, \dots, N \quad (7.30)$$

We will use the above state prediction errors in the switching mechanism of the multiple models scheme. We use the permissible switching scheme as defined in 6.1. If a switching occurs at time t , the parameter estimate vector $\hat{\theta}(t)$ is instantaneously reset to $\hat{\theta}_j$, where j denotes the index of the selected model. The multi-linear parameter vector $\hat{\Theta}$ is then obtained easily using the elements of $\hat{\theta}_j$.

Since our control methodology is based on the certainty equivalence principle, better parameter identification will guarantee better tracking performance. Hence, we modify the identification scheme defined in Section 5.5 such that the parameter update

equation in (5.171) re-initializes itself with a fixed model which would minimize a cost function defined as,

$$J_j(t) \stackrel{\text{def}}{=} \tilde{x}_j(t)^T G_1 \tilde{x}_j(t) + \int_0^t e^{-\lambda(t-\tau)} \tilde{x}_j(\tau)^T G_2 \tilde{x}_j(\tau) d\tau \quad , \quad j = 0, \dots, N \quad (7.31)$$

where $G_1, G_2 \in \mathbb{R}^{n \times n}$ are positive (semi-)definite weight matrices and $\lambda \geq 0$ is a scalar forgetting factor. We use the definition for switching logic as given in Definition 6.2. Under this switching logic, the parameter estimate vector $\hat{\theta}$, the multi-linear parameter estimate vector $\hat{\Theta}$, and the parameter error vector $\tilde{\Theta}$, will be defined as $\hat{\theta}_{z(t)}$, $\hat{\Theta}_{z(t)}$, and $\tilde{\Theta}_{z(t)}$, respectively, where $z(t)$ is defined as $z(t) : \mathbb{R}_+ \mapsto 0, \dots, N$ such that if $t \in [T_i, T_{i+1})$ for some $i < \infty$, then $z(t) = z(T_i)$. Here, T_i is the permissible switching sequence defined in Definition 6.1.

Remark 7.3 *Note that in the Lyapunov-function based switching, change in the Lyapunov function ($\Delta V(t) = V(t^+) - V(t)$) is calculated at each switching instant. Since we use performance-index based switching in this section, we do not need to indicate the time argument in $z(t)$. Hence, for brevity of notation, we will simply use z to mean $z(t)$, and we will use $\hat{\theta}_z$, $\hat{\Theta}_z$, etc. in the rest of this section.*

Using the switched forms of parameter estimate vectors $\hat{\theta}_z$ and $\hat{\Theta}_z$, the applied control signal becomes,

$$u = \frac{1}{\hat{\Theta}_z^T \mathcal{G}(x)} (-\hat{\Theta}_z^T \mathcal{F}(x) + \hat{v}_z) \quad (7.32)$$

with

$$\hat{v}_z = y_r^{(\gamma)} - \alpha_1(\hat{y}_z^{(\gamma-1)} - y_r^{(\gamma-1)}) - \dots - \alpha_{\gamma-1}(\dot{\hat{y}}_z - \dot{y}_r) - \alpha_\gamma(\hat{y} - y_r) \quad (7.33)$$

where $\dot{\hat{y}}_z, \dots, \hat{y}_z^{(\gamma-1)}$ are calculated using $\hat{\Theta}_z$. Substituting this control signal in the

system dynamics, we obtain the closed loop error dynamics as,

$$e^{(\gamma)} + \alpha_1 e^{(\gamma-1)} + \dots + \alpha_\gamma e = \Theta^T \mathcal{F}(x) + \frac{\Theta^T \mathcal{G}(x)}{\hat{\Theta}_z^T \mathcal{G}(x)} (-\hat{\Theta}_z^T \mathcal{F}(x) + \hat{v}_z) - v \quad (7.34)$$

We can write,

$$e^{(\gamma)} + \alpha_1 e^{(\gamma-1)} + \dots + \alpha_\gamma e = \underbrace{-\tilde{\Theta}_z^T \mathcal{F}(x) - \tilde{\Theta}_z^T \mathcal{G}(x) \frac{-\hat{\Theta}_z^T \mathcal{F}(x) + \hat{v}_z}{\hat{\Theta}_z^T \mathcal{G}(x)}}_{\tilde{\Theta}_z^T W_{1z}} + \underbrace{\hat{v}_z - v}_{\tilde{\Theta}_z^T W_{2z}} \quad (7.35)$$

Setting $W_z = W_{1z} + W_{2z}$, the error dynamics in (7.35) can be expressed as,

$$e = M(s) \cdot \tilde{\Theta}_z^T W_z \quad (7.36)$$

Comparing (7.35) to (5.160), it is seen that W_z has the same form as W , with the difference that all parameter estimates are replaced by the switched estimates. The augmented error given in (5.167) can now be written as,

$$\varepsilon = e + \hat{\Theta}_z^T M(s) W_z - M(s) \hat{\Theta}_z^T W_z \quad (7.37)$$

Since the parameter update law in (5.171) depends on the augmented error and the regressor vector, we have to re-define the update law and show its stability in the following theorem.

Theorem 7.2 *Consider the error dynamics in (7.36) with the filter defined in (5.165), the augmented error equation in (7.37), and the parameter update law*

$$\dot{\hat{\Theta}} = \frac{-\varepsilon \mu_z}{1 + \mu_z^T \mu_z} \quad (7.38)$$

where $\mu_z = M(s) W_z$. Then, $\tilde{\Theta} \in \mathcal{L}_\infty$, $\dot{\hat{\Theta}} \in \mathcal{L}_2 \cap \mathcal{L}_\infty$, and,

$$|\tilde{\Theta}^T \mu| \leq \lambda(1 + \|\mu\|) \quad , \quad \forall t \quad (7.39)$$

for some $\lambda \in \mathcal{L}_2 \cap \mathcal{L}_\infty$, where $\|\cdot\|$ denotes the ∞ -norm.

Proof: We choose the Lyapunov function as,

$$V(\tilde{\Theta}) = \tilde{\Theta}^T \tilde{\Theta} \quad (7.40)$$

Differentiating (7.40) and using (7.38), we obtain,

$$\dot{V} = \frac{-2\varepsilon^2}{1 + \mu_z^T \mu_z} \leq 0 \quad (7.41)$$

which shows $\tilde{\Theta} \in \mathcal{L}_\infty$. Integrating (7.41), we have,

$$\int_0^t \dot{V}(\tilde{\Theta}) d\tau = \int_0^t \frac{-2\varepsilon^2}{1 + \mu_z^T \mu_z} d\tau \quad (7.42)$$

$$V(\tilde{\Theta}(t)) - V(\tilde{\Theta}(0)) = -2 \int_0^t \frac{\varepsilon^2}{1 + \mu_z^T \mu_z} d\tau \quad (7.43)$$

From (7.40) and boundedness of $\tilde{\Theta}$, the left-hand side of (7.43) is bounded. This leads to $\sqrt{\frac{\varepsilon^2}{1 + \mu_z^T \mu_z}} \in \mathcal{L}_2$. From (7.36) and (7.37), we have $\varepsilon = \tilde{\Theta}_z^T M(s) W_z$. Using this in (7.38), and knowing that $\dot{\tilde{\Theta}} = \dot{\tilde{\Theta}}$, we obtain,

$$\dot{\tilde{\Theta}} = \frac{-\mu_z \mu_z^T \tilde{\Theta}}{1 + \mu_z^T \mu_z} \quad (7.44)$$

Since $\frac{x}{1+x} \leq 1$ for all $x \geq 0$, and $\tilde{\Theta}$ is bounded, we have $\dot{\tilde{\Theta}} \in \mathcal{L}_\infty$. We can also write,

$$|\dot{\tilde{\Theta}}|^2 = \frac{\varepsilon^2}{1 + \mu_z^T \mu_z} \cdot \frac{\mu_z^T \mu_z}{1 + \mu_z^T \mu_z} \quad (7.45)$$

Due to (7.43), $\frac{\varepsilon^2}{1 + \mu_z^T \mu_z}$ in (7.45) is integrable, and $\frac{\mu_z^T \mu_z}{1 + \mu_z^T \mu_z}$ is bounded. Therefore $\dot{\tilde{\Theta}} \in \mathcal{L}_2$.

Defining a λ such that,

$$\lambda = \frac{\varepsilon}{\sqrt{1 + \mu_z^T \mu_z}} \cdot \frac{\sqrt{1 + \mu_z^T \mu_z}}{1 + \|\mu_z\|} \quad (7.46)$$

one observes that $\frac{\varepsilon}{\sqrt{1 + \mu_z^T \mu_z}} \in \mathcal{L}_2$ and $\frac{\sqrt{1 + \mu_z^T \mu_z}}{1 + \|\mu_z\|} \in \mathcal{L}_\infty$. Hence $\lambda \in \mathcal{L}_2 \cap \mathcal{L}_\infty$ and (7.39) holds. $\square$

Remark 7.4 *Note that in (7.38), ε and μ_z are calculated using $\hat{\Theta}_z$ but the left-hand side of this equation is not $\dot{\hat{\Theta}}_z$, but $\dot{\hat{\Theta}}$. Moreover, this update dynamics is subject to further change due to switching through a permissible switching logic as defined in Definition 6.2.*

Having built-up the identification dynamics under the permissible switching, we have to establish the closed loop stability of the error dynamics defined in (7.36). Before stating a theorem about asymptotic tracking using multiple models, we will review two lemmas (Lemma 7.1 and Lemma 7.2) which are well known in the literature. Then we will present additional three lemmas (Lemma 7.3 - Lemma 7.5) about $\tilde{\Theta}_z$, $\dot{\tilde{\Theta}}_z$ and W_z , which are unique for the multiple model adaptive control.

Lemma 7.1 (BOBI stability) [8] *Let $y = H(s)u$, where H is a proper, rational transfer function. If H is minimum-phase (i.e. all zeros on the left half plane) and u is regular, then,*

$$\|u\|_t \leq k_1 \|y\|_t + k_2 \quad (7.47)$$

for some $k_1, k_2 \geq 0$, where the norm $\|\cdot\|_t$ refers to the truncated ∞ -norm as defined in Definition 2.11.

Lemma 7.2 [8] *Let $y = H(s)u$, where H is a proper, stable transfer function, and*

$\lambda \in \mathcal{L}_2 \cap \mathcal{L}_\infty$. If,

$$\|u\|_t \leq \lambda \|q\|_t + \lambda \quad (7.48)$$

then,

$$\|y\|_t \leq \lambda \|q\|_t + \lambda \quad (7.49)$$

Moreover, if H is strictly proper, then,

$$|y|_t \leq \zeta \|q\|_t + \zeta \quad (7.50)$$

where $\zeta \in \mathcal{L}_2 \cap \mathcal{L}_\infty$ and $\zeta \rightarrow 0$ as $t \rightarrow \infty$.

A more detailed discussion of the above lemmas can be found in [6]. Now we state the lemmas that will be used for the case of multiple models.

Lemma 7.3 *If $\tilde{\Theta} \in \mathcal{L}_\infty$, then $\tilde{\Theta}_z \in \mathcal{L}_\infty$.*

Proof: $\tilde{\Theta}_z$ is obtained from $\tilde{\Theta}$ via switching to certain values belonging to a compact set. Since the switched values are fixed and bounded, the statement holds. $\square$

Lemma 7.4 *If $\dot{\tilde{\Theta}} \in \mathcal{L}_2 \cap \mathcal{L}_\infty$ and the switching stops in finite time, then $\dot{\tilde{\Theta}}_z \in \mathcal{L}_2 \cap \mathcal{L}_\infty$.*

Proof: $\dot{\tilde{\Theta}}_z \in \mathcal{L}_\infty$ is obtained as a consequence of nonzero dwell time within which the change in the parameter error vector belongs to a fixed, compact parameter set. We can write an upper bound for $\|\dot{\tilde{\Theta}}_z\|$ simply using the definition of differentiation. If the

dwell time is denoted by T_{min} , then,

$$\|\dot{\tilde{\Theta}}_z\| \leq \sup_{j,t} \frac{\|\tilde{\Theta}_j(t + T_{min}) - \tilde{\Theta}(t)\|}{T_{min}} \quad (7.51)$$

Expanding the terms we get,

$$\|\dot{\tilde{\Theta}}_z\| \leq \sup_{j,t} \frac{\|\hat{\Theta}_j(t + T_{min}) - \Theta(t + T_{min}) - \hat{\Theta}(t) + \Theta(t)\|}{T_{min}} \quad (7.52)$$

Utilizing the fact that the actual system parameters are unknown but constant, we have $\Theta(t + T_{min}) = \Theta(t)$. We also have $\hat{\Theta}_j(t + T_{min}) = \hat{\Theta}_j(t)$ since the models are fixed. Then we can write (7.52) as,

$$\|\dot{\tilde{\Theta}}_z\| \leq \sup_{j,t} \frac{\|\hat{\Theta}_j(t) - \hat{\Theta}(t)\|}{T_{min}} \quad (7.53)$$

Since $\hat{\Theta}(t)$ is bounded for all t , and $\hat{\Theta}_j(t)$ belongs to a compact set for all t with $j = 1, \dots, N$ being the model index, then we have,

$$\|\dot{\tilde{\Theta}}_z\| \leq K \quad (7.54)$$

where $K < \infty$ is a positive bound. It is easy to note that even if the actual system parameters suddenly change such that $\Theta(t + T_{min}) \neq \Theta(t)$ for a specific time instant, we would still be able to write (7.54) since the actual multi-linear parameter vector Θ is bounded. To show that $\dot{\tilde{\Theta}}_z \in \mathcal{L}_2$, we use the condition of the lemma that the switching stops in finite time. If the switching stops at time T , we have,

$$\tilde{\Theta}_z(t) = \tilde{\Theta}(t) \quad , \quad \forall t \geq T > 0 \quad (7.55)$$

Since $\dot{\tilde{\Theta}} \in \mathcal{L}_2$, using (7.54) and (7.55), we conclude that $\dot{\tilde{\Theta}}_z \in \mathcal{L}_2$. $\square$

Lemma 7.5 Consider $\mu = M(s)W$ and $\mu_z = M(s)W_z$ where W, W_z are the regressor vectors defined in (5.160), (7.35), respectively, and $M(s)$ is a stable transfer function

as defined in (5.165). If,

$$\frac{|\tilde{\Theta}^T \mu|}{1 + \|\mu\|_t} \in \mathcal{L}_2 \cap \mathcal{L}_\infty \quad (7.56)$$

and the switching stops in finite time, then,

$$\frac{|\tilde{\Theta}_z^T \mu_z|}{1 + \|\mu_z\|_t} \in \mathcal{L}_2 \cap \mathcal{L}_\infty \quad (7.57)$$

Proof: W_z is calculated as described in (7.35) using $\hat{\Theta}_z$, which is bounded as stated in Lemma 7.3. Hence it is easy to see the boundedness part of the above statement. That is,

$$\frac{|\tilde{\Theta}_z^T \mu_z|}{1 + \|\mu_z\|_t} \in \mathcal{L}_\infty \quad (7.58)$$

Since the switching stops at finite time T as an *a priori* condition of the lemma, we can write,

$$\tilde{\Theta}_z(t) = \tilde{\Theta}(t) \quad , \quad \forall t \geq T > 0 \quad (7.59)$$

and consequently,

$$W_z(t) = W(t) \quad , \quad \forall t \geq T > 0 \quad (7.60)$$

$$\mu_z(t) = \mu(t) \quad , \quad \forall t \geq T > 0 \quad (7.61)$$

To show $\frac{|\tilde{\Theta}_z^T \mu_z|}{1 + \|\mu_z\|_t} \in \mathcal{L}_2$, we will analyze the signal in two parts. For $T \leq t < \infty$, the two signals in (7.56) and (7.57) become identical and we directly obtain,

$$\frac{|\tilde{\Theta}_z^T \mu_z|}{1 + \|\mu_z\|_t} \in \mathcal{L}_2 \quad , \quad T \leq t < \infty \quad (7.62)$$

For $0 \leq t < T$, we will use the boundedness of $\frac{|\tilde{\Theta}_z^T \mu_z|}{1 + \|\mu_z\|_t}$ given in (7.58), and the boundedness of $\dot{\tilde{\Theta}}_z$ as stated by Lemma 7.4 (which makes the time derivative of $\frac{|\tilde{\Theta}_z^T \mu_z|}{1 + \|\mu_z\|_t}$ bounded for $t < T$). Using this result and the *a priori* condition of the lemma given in (7.56), we can write,

$$\frac{|\tilde{\Theta}_z^T \mu_z|}{1 + \|\mu_z\|_t} \in \mathcal{L}_2 \quad , \quad 0 \leq t < T \quad (7.63)$$

The equations (7.58), (7.62) and (7.63) reflect the desired result given by (7.57). $\square$

Based on the above lemmas, we are now in a position to state the main theorem.

Theorem 7.3 *Consider the nonlinear system given in (5.142) and its normal form representation as in (5.148) with exponentially minimum-phase internal dynamics. With the indirect identification model in (7.25), (7.27), N fixed identification models defined in (7.28), and the switching logic for parameter update dynamics defined in Definition 6.2, if the reference signal y_r is sufficiently rich, y_r and its derivatives $\dot{y}_r, \ddot{y}_r, \dots, y_r^{(\gamma+1)}$ are all bounded, $\hat{\Theta}_z^T \mathcal{G}(x)$ is bounded away from zero, f, g, h are Lipschitz continuous functions of x with Lipschitz continuous Lie derivatives $L_f^k h, L_g L_f^k h$ for $k = 1, \dots, \gamma$, and the regressor vector $W(x, \hat{\theta})$ is regular in $x, \hat{\theta}$, then the certainty equivalence control laws defined in (7.32)-(7.33) and the parameter update law defined in (7.38) which is obtained by error augmentation given in (7.37), result in bounded tracking with asymptotic convergence, that is, $y \rightarrow y_r$ as $t \rightarrow \infty$, and improve the transient response.*

Proof: The closed loop error dynamics given in (7.36) is a linear system where $M(s)$ is strictly proper and stable. For that reason, we will utilize Lemma 7.1 and Lemma 7.2. To do this, we will treat the term $\tilde{\Theta}_z^T W_z$ as the input of the linear system. In order to use Lemma 7.1, we have to establish the regularity of the term $\tilde{\Theta}_z^T W_z$ according to Definition 2.13.

From the linearizing diffeomorphism given in Theorem 3.2,

$$\xi = \begin{bmatrix} y \\ \dot{y} \\ \vdots \\ y^{(\gamma-1)} \end{bmatrix} \quad (7.64)$$

Since $y = e + y_r$, (7.64) can be written as,

$$\xi = M(s) \begin{bmatrix} 1 \\ \vdots \\ s^{\gamma-1} \end{bmatrix} \tilde{\Theta}_z^T W_z + \begin{bmatrix} y_r \\ \vdots \\ y_r^{(\gamma-1)} \end{bmatrix} \quad (7.65)$$

We know from Theorem 7.2 that $\tilde{\Theta}$ is bounded. Hence from Lemma 7.3, $\tilde{\Theta}_z$ is bounded too. $y_r, \dot{y}_r, \dots, y_r^{(\gamma)}$ are bounded by assumption. Considering (5.165), all $s^i M(s)$ products are strictly proper transfer functions for $i = 0, \dots, \gamma - 1$. Let $K \in \mathbb{R}$ be a *general-purpose* bound. Hence we can write,

$$\|\xi\|_t \leq K \|W_z\|_t + K \quad (7.66)$$

Since the zero-dynamics $\dot{\eta} = q(0, \xi)$ is assumed to be exponentially minimum-phase, we have,

$$\|\eta\|_t \leq K \|W_z\|_t + K \quad (7.67)$$

Taking the time derivative of ξ , we get,

$$\dot{\xi} = M(s) \begin{bmatrix} s \\ \vdots \\ s^\gamma \end{bmatrix} \tilde{\Theta}_z^T W_z + \begin{bmatrix} \dot{y}_r \\ \vdots \\ y_r^{(\gamma)} \end{bmatrix} \quad (7.68)$$

All $s^i M(s)$ terms are still proper transfer functions for $i = 1, \dots, \gamma$. $\tilde{\Theta}_z$ and $\dot{y}_r, \dots, y_r^{(\gamma)}$

are bounded. Therefore we have,

$$\|\dot{\xi}\|_t \leq K\|W_z\|_t + K \quad (7.69)$$

Due to the same reason that $s^i M(s)$'s are strictly proper for $i = 0, \dots, \gamma - 1$, we also get,

$$\|\dot{\eta}\|_t \leq K\|W_z\|_t + K \quad (7.70)$$

Since x is a diffeomorphism of (ξ, η) , we have,

$$\|x\|_t \leq K(\|\xi\|_t + \|\eta\|_t) \quad (7.71)$$

Using (7.66) and (7.67) in (7.71), we obtain,

$$\|x\|_t \leq K\|W_z\|_t + K \quad (7.72)$$

Since $\dot{x} = f + g \cdot u$, using Lipschitz continuity of $f, g, L_f^k h$ and $L_g L_f^k h$, we have,

$$\|\dot{x}\|_t \leq K\|x\|_t + K \quad (7.73)$$

and substituting (7.72) in (7.73), we obtain,

$$\|\dot{x}\|_t \leq K\|W_z\|_t + K \quad (7.74)$$

W_z is a function of $x, \hat{\theta}_z$ and $y_r, \dot{y}_r, \dots, y_r^{(\gamma)}$ (its dependence to y_r and its first γ derivatives is through $\hat{v}$). Taking the time derivative of W_z , we get,

$$\dot{W}_z = \frac{\partial W_z}{\partial x} \dot{x} + \frac{\partial W_z}{\partial \hat{\theta}_z} \dot{\hat{\theta}}_z + \frac{\partial W_z}{\partial \hat{v}_z} \dot{\hat{v}}_z \quad (7.75)$$

This leads to,

$$\|\dot{W}_z\|_t \leq \left\| \frac{\partial W_z}{\partial x} \right\|_t \|\dot{x}\|_t + \left\| \frac{\partial W_z}{\partial \hat{\theta}_z} \right\|_t \|\dot{\hat{\theta}}_z\|_t + \left\| \frac{\partial W_z}{\partial \hat{v}_z} \right\|_t \|\dot{\hat{v}}_z\|_t \quad (7.76)$$

W is assumed to have bounded derivatives in x and $\hat{\theta}$. Having this assumption, since $\frac{\partial W}{\partial x} = \frac{\partial W_z}{\partial x}$, W_z has bounded partial derivatives in x ; and since $\hat{\theta}_z$ is bounded, W_z has bounded partial derivatives in $\hat{\theta}_z$. On the other hand $\|\hat{\theta}\|, \|\dot{\hat{\theta}}\|$ are bounded due to Theorem 7.2. Boundedness of the term $\left\| \frac{\partial W_z}{\partial \hat{v}_z} \right\|$ can be deduced from (7.35), as $\hat{\Theta}_z^T \mathcal{G}$ is bounded away from zero by assumption, yielding $\left\| \frac{\partial W_z}{\partial \hat{v}_z} \right\|_t = \left\| \frac{\mathcal{G}}{\hat{\Theta}_z^T \mathcal{G}} \right\|_t \leq K$. Differentiating (7.33), we get,

$$\dot{\hat{v}}_z = y_r^{(\gamma+1)} - \alpha_1(\hat{y}^{(\gamma)} - y_r^{(\gamma)}) - \dots - \alpha_\gamma(\dot{\hat{y}} - \dot{y}_r) \quad (7.77)$$

Since $\dot{y}_r, \dots, y_r^{(\gamma+1)}$ are bounded, and $\hat{y}^{(i)} = L_f^i h$ are Lipschitz continuous for $i = 1, \dots, \gamma$ by assumption, we obtain,

$$\|\dot{\hat{v}}_z\|_t \leq K\|x\|_t + K \quad (7.78)$$

Plugging (7.72) into (7.78), we can write,

$$\|\dot{\hat{v}}_z\|_t \leq K\|W_z\|_t + K \quad (7.79)$$

Using (7.74), (7.79) in (7.76), we obtain,

$$\|\dot{W}_z\|_t \leq K\|W_z\|_t + K \quad (7.80)$$

Hence, regularity of W_z is shown. However, we are trying to show the regularity of $\tilde{\Theta}_z^T W_z$. Differentiating this term, we get,

$$\frac{d}{dt}(\tilde{\Theta}_z^T W_z) = \dot{\tilde{\Theta}}_z^T W_z + \tilde{\Theta}_z^T \dot{W}_z \quad (7.81)$$

From Theorem 7.2 and Lemma 7.4, we have $\dot{\tilde{\Theta}}_z \in \mathcal{L}_\infty$. Using (7.80) and the boundedness of $\tilde{\Theta}_z, \dot{\tilde{\Theta}}_z$, we obtain,

$$\left\| \frac{d}{dt}(\tilde{\Theta}_z^T W_z) \right\|_t \leq K \|W_z\|_t + K \quad (7.82)$$

From (7.65), we have,

$$\|\xi\|_t \leq K \|\tilde{\Theta}_z^T W_z\|_t + K \quad (7.83)$$

and from (7.66) and (7.67), we can write,

$$\|\eta\|_t \leq K \|\xi\|_t + K \quad (7.84)$$

Using (7.83) in (7.84), we get,

$$\|\eta\|_t \leq K \|\tilde{\Theta}_z^T W_z\|_t + K \quad (7.85)$$

and using (7.83) and (7.85) in (7.71), we obtain,

$$\|x\|_t \leq K \|\tilde{\Theta}_z^T W_z\|_t + K \quad (7.86)$$

From (7.72), we can write,

$$K \|W_z\|_t + K \leq \|x\| \quad (7.87)$$

for appropriate values of K . Then, we substitute (7.87) in (7.86) and we obtain,

$$\|W_z\|_t \leq K \|\tilde{\Theta}_z^T W_z\|_t + K \quad (7.88)$$

Using this result in (7.82) leads to the regularity of $\tilde{\Theta}_z^T W_z$, that is,

$$\left\| \frac{d}{dt} (\tilde{\Theta}_z^T W_z) \right\|_t \leq K \|\tilde{\Theta}_z^T W_z\|_t + K \quad (7.89)$$

Now we can apply Lemma 7.1 to (5.166), to obtain,

$$\|\tilde{\Theta}_z^T W_z\|_t \leq K|e| + K \quad (7.90)$$

Writing $M(s)$ as,

$$M(s) = c^T (sI - A)^{-1} b \quad (7.91)$$

and using Lemma 2.6 (swapping lemma), we get,

$$\tilde{\Theta}_z^T M(s) W_z - M(s) \tilde{\Theta}_z W_z = -c^T (sI - A)^{-1} \{[(sI - A)^{-1} b W_z^T] \dot{\tilde{\Theta}}_z\} \quad (7.92)$$

From Theorem 7.2 and Lemma 7.4, we have $\dot{\tilde{\Theta}}_z \in \mathcal{L}_2$. Since $M(s)$ is stable, we know that $(sI - A)^{-1} b$ is stable as well. Then we can write,

$$|[(sI - A)^{-1} b W_z^T] \dot{\tilde{\Theta}}_z| \leq \lambda \|W_z\|_t + \lambda \quad (7.93)$$

In (7.91), $c^T (sI - A)^{-1}$ is strictly proper. Let $\zeta \in \mathcal{L}_2 \cap \mathcal{L}_\infty$ be another generic function such that $\zeta \rightarrow 0$ as $t \rightarrow \infty$. Then, from Lemma 7.2, we obtain,

$$|\tilde{\Theta}_z^T M(s) W_z - M(s) \tilde{\Theta}_z W_z| \leq \zeta \|W_z\|_t + \zeta \quad (7.94)$$

Using this in (7.37) with the triangular inequality, we obtain,

$$|e| \leq |\varepsilon| + \zeta \|W_z\|_t + \zeta \quad (7.95)$$

where e is the tracking error and ε is the augmented error. Due to (7.88), we can write,

$$|e| \leq |\varepsilon| + \zeta \|\tilde{\Theta}_z^T W_z\|_t + \zeta \quad (7.96)$$

Using (7.90) in (7.96), we obtain,

$$|e| \leq |\varepsilon| + \zeta \|e\|_t + \zeta \quad (7.97)$$

Now, from Theorem 7.2 and Lemma 7.5, we have,

$$\frac{|\tilde{\Theta}_z^T \mu_z|}{1 + \|\mu_z\|_t} \in \mathcal{L}_2 \cap \mathcal{L}_\infty \quad (7.98)$$

If we can show that $\frac{\tilde{\Theta}_z^T \mu_z}{1 + \|\mu_z\|_t}$ has bounded derivative, then using Barbalat's lemma, we can obtain $|\tilde{\Theta}_z^T \mu_z| \rightarrow 0$ as $t \rightarrow \infty$. Differentiating this term, we get,

$$\frac{d}{dt} \left(\frac{\tilde{\Theta}_z^T \mu_z}{1 + \|\mu_z\|_t} \right) = \frac{1}{1 + \|\mu_z\|_t} (\dot{\tilde{\Theta}}_z^T \mu_z + \tilde{\Theta}_z^T \dot{\mu}_z) \quad (7.99)$$

and

$$\left\| \frac{d}{dt} \left(\frac{\tilde{\Theta}_z^T \mu_z}{1 + \|\mu_z\|_t} \right) \right\|_t \leq \frac{1}{1 + \|\mu_z\|_t} (\|\dot{\tilde{\Theta}}_z^T\|_t \|\mu_z\|_t + \|\tilde{\Theta}_z^T\|_t \|\dot{\mu}_z\|_t) \quad (7.100)$$

We already have $\|\tilde{\Theta}_z^T\|_t \leq K$ and $\|\dot{\tilde{\Theta}}_z^T\|_t \leq K$. As $\mu_z = M(s)W_z$ and $M(s)$ is stable, μ_z is regular too. Hence,

$$\|\mu_z\|_t \leq K \|W_z\|_t + K \quad (7.101)$$

and

$$\|\dot{\mu}_z\|_t \leq K \|\mu_z\|_t + K \quad (7.102)$$

Using (7.102) in (7.100) leads to,

$$\left\| \frac{d}{dt} \left(\frac{\tilde{\Theta}_z^T \mu_z}{1 + \|\mu_z\|_t} \right) \right\|_t \leq K \quad (7.103)$$

Then using (7.98) and (7.103), we can write,

$$|\tilde{\Theta}_z^T \mu_z| \leq \zeta(1 + \|\mu_z\|_t) \quad (7.104)$$

Equivalently,

$$|\varepsilon| \leq \zeta(1 + \|\mu_z\|_t) \quad (7.105)$$

Using this in (7.97), we obtain,

$$|e| \leq \zeta\|e\|_t + \zeta\|\mu_z\|_t + \zeta \quad (7.106)$$

From (7.88), (7.90) and (7.101), we can write,

$$|e| \leq \zeta\|e\|_t + \zeta \quad (7.107)$$

Since $\zeta \rightarrow 0$ as $t \rightarrow \infty$, it is seen that $e \rightarrow 0$, hence $y \rightarrow y_r$, as $t \rightarrow \infty$.

Hence we have proved that the proposed switching system is asymptotically stable. Now we consider the performance of its transient response. Since our methodology is based on the certainty equivalence principle, we assume that there is a correlation between the identification performance and the tracking error. If we can estimate the unknown parameter vector θ better, then we expect that the output error e will converge to zero faster.

From steady-state part of (7.26) and (7.27), we can write the identification error

for the single model adaptive system as,

$$\tilde{x} = \hat{x} - x = w_f^T(\hat{\theta} - \theta) \quad (7.108)$$

The identification error $\tilde{x}$ reflects the parameter error $\tilde{\theta}$. Similarly, one obtains the identification error for the multiple models as,

$$\tilde{x}_{j^*} = \hat{x}_{j^*} - x = w_f^T(\hat{\theta}_z - \theta) \quad (7.109)$$

where j^* is as defined in Definition 6.2.

The switching criterion in (6.2) is based on the cost function,

$$J_j(t) \stackrel{\text{def}}{=} \tilde{x}_j(t)^T G_1 \tilde{x}_j(t) + \int_0^t e^{-\lambda(t-\tau)} \tilde{x}_j(\tau)^T G_2 \tilde{x}_j(\tau) d\tau \quad , \quad j = 0, \dots, N$$

as given in (7.31). The cost function is a quadratic weighted sum of identification errors for all models. The adaptive parameter estimate vector $\hat{\theta}$ is reset to the fixed parameter vector $\hat{\theta}_{j^*}$ which minimizes the cost function. This guarantees that the identification error for the closed-loop multiple model system will decrease due to the switching. That is,

$$\|\tilde{x}_{j^*}\| \leq \|\tilde{x}\| \quad (7.110)$$

where $\tilde{x}$ is the identification error as given in (7.108) for the single-model adaptive system.

The tracking error of the nonlinear system is linear in terms of the multi-linear parameter vector Θ . On the other hand, the identification error is expressed in (7.108) and (7.109) using the plant parameter vector θ . The multi-linear parameter vector Θ can always be constructed from the components of θ . The explicit form of Θ is given

by the equations (5.144) through (5.147). It is easy to see that,

$$\|\tilde{\theta}_z\| \leq \|\tilde{\theta}\| \Rightarrow \|\tilde{\Theta}_z\| \leq \|\tilde{\Theta}\| \quad (7.111)$$

The $\|\tilde{\theta}_z\| \leq \|\tilde{\theta}\|$ part of the above equation is a conclusion of (7.110) under the certainty equivalence assumption. We will use this result to show the transient performance improvement in the multiple model adaptive system. The closed loop tracking errors are given in (5.166) and (7.36) for the single model and the multiple model adaptive systems, respectively. The transient improvement in the tracking error requires,

$$|M(s)\tilde{\Theta}_z^T W_z| \leq |M(s)\tilde{\Theta}^T W| \quad (7.112)$$

Since $M(s)$ is a stable filter, (7.112) leads to,

$$|\tilde{\Theta}_z^T W_z| \leq |\tilde{\Theta}^T W| \quad (7.113)$$

The indirect identification scheme described in (7.25)-(7.27) leads to the convergence of the identification error, that is, $\tilde{x} \rightarrow 0$. Combining this result and (7.110), we obtain $\tilde{x}_{j^*} \rightarrow 0$. The switching is performed via the minimization of cost function given in (7.31), where the index “0” in J (i.e. J_0) represents the cost function calculated for the adaptive model. Then, $\tilde{x} \rightarrow 0$ and $\tilde{x}_{j^*} \rightarrow 0$ imply $J_0 \rightarrow J_{j^*}$ and the switching will stop in finite time T . In fact, if the forgetting factor λ in (7.31) is a strictly positive constant, then we also have $J_0 \rightarrow 0$ and $J_{j^*} \rightarrow 0$. Based on the fact that the switching stops in finite time, we have $\tilde{\theta}_z \rightarrow \tilde{\theta}$. With a sufficiently rich reference signal, we also have $\tilde{\theta} \rightarrow 0$. Hence, $\tilde{\theta}_z \rightarrow 0$. Since the multi-linear parameter error vector $\tilde{\Theta}_z$ is composed of the components of $\tilde{\theta}_z$, we have $\tilde{\Theta}_z \rightarrow 0$, as well.

Now we turn back to the transient performance enhancement condition expressed in (7.113). From (7.110), (7.111), and the fact that $\tilde{\Theta}_z \rightarrow 0$, we can write,

$$|\tilde{\Theta}_z^T W_z| \leq |\tilde{\Theta}^T W| \quad , \quad \forall t \geq T > 0 \quad (7.114)$$

where T denotes the time after which no switching occurs. Therefore, under the certainty equivalence principle, it is guaranteed that the tracking error of the multiple model adaptive system will not be larger than the tracking error of the single model adaptive system, after some small time T . If the fixed models are selected properly such that they span the parameter space in which the parameter vector is confined to, then based on the switching logic, the switching takes place and this leads to,

$$|\tilde{\Theta}_z^T W_z| < |\tilde{\Theta}^T W| \quad , \quad \forall t \geq T > 0 \quad (7.115)$$

Based on the above analysis and proper determination of the fixed models, the multiple models and switching yield better transient response in the tracking error than the case of single model direct adaptive controller. $\square$

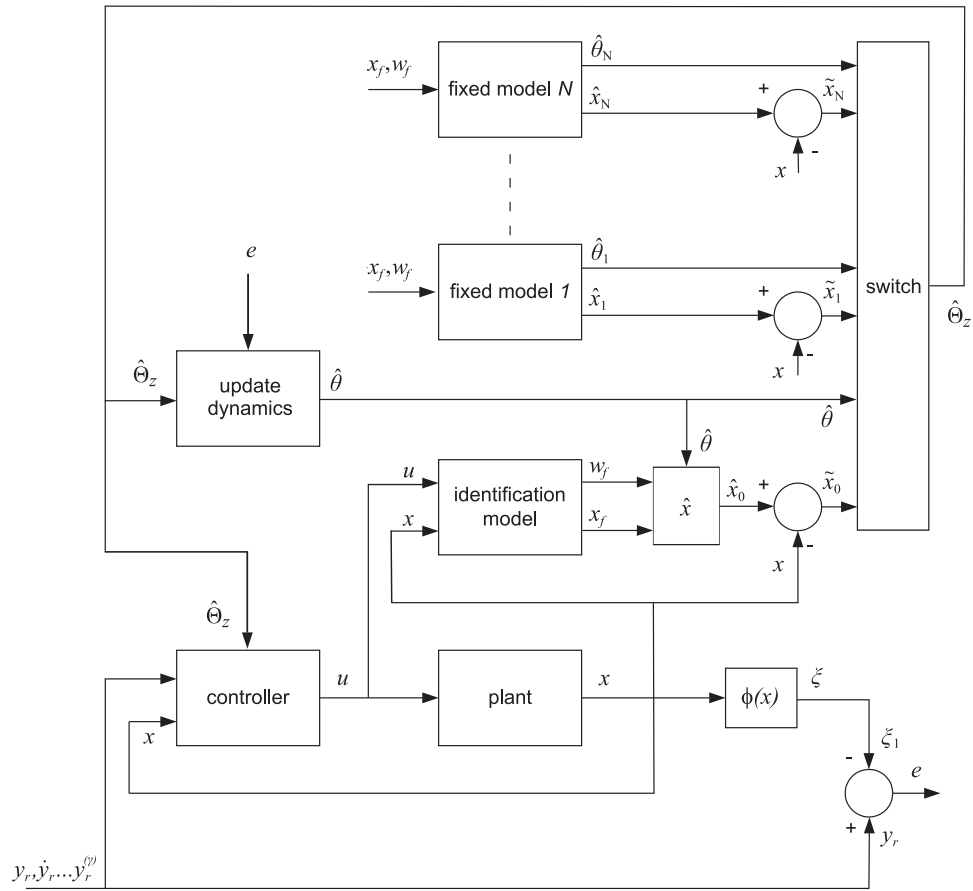


Figure 7.1. Block diagram of the direct adaptive controller with N fixed models and one adaptive model.

The block diagram of the proposed direct adaptive control system with N fixed identification models and one adaptive model is shown in Figure 7.1. Next, we present a simulation study in order to verify the results obtained for the proposed direct adaptive controller with multiple models and switching.

7.2.1. Simulation Study

We use the nonlinear system given in Section 6.3.1 for the simulation of the direct adaptive control scheme using multiple models. The linearizing control law is obtained from (5.154) as,

$$u = \hat{v} - \hat{p}(10 + \sin x_1 + x_1 \cos x_1)(x_2 + \hat{p}x_1(10 + \sin x_1)) \quad (7.116)$$

From (5.153), the tracking control is given as,

$$v = \ddot{y}_r + \alpha_1(\dot{y}_r - \dot{y}) + \alpha_2(y_r - y) \quad (7.117)$$

where y_r is the reference output. However, $\dot{y} = L_f h$ is not measurable, since $L_f h = x_2 + px_1(10 + \sin x_1)$, where “ p ” is the unknown parameter. Therefore we use the estimate of p and calculate $\hat{v}$ in (7.116) as,

$$\hat{v} = \ddot{y}_r + \alpha_1[\dot{y}_r - \underbrace{(x_2 + \hat{p}x_1(10 + \sin x_1))}_{L_{\hat{f}}h}] + \alpha_2(y_r - y) \quad (7.118)$$

The feedback gains are chosen as $\alpha_1 = 20$ and $\alpha_2 = 250$. The closed loop error dynamics is obtained as,

$$\dot{e} = \tilde{\Theta}^T W = \begin{bmatrix} \tilde{p} \\ \tilde{p}^2 \end{bmatrix}^T \begin{bmatrix} x_2(10 + \sin x_1 + x_1 \cos x_1) + \alpha_1 x_1(10 + \sin x_1) \\ (10 + \sin x_1 + x_1 \cos x_1)(10 + \sin x_1) \end{bmatrix} \quad (7.119)$$

Filtered forms of x and w , namely x_f and w_f , were obtained using (7.25) with $\sigma = 0.09$. The reference trajectory was generated as $y_r = 10 \sin \pi t + 8 \sin 2\pi t$. The weighting ma-

trices for the performance index are set to $G_1 = \text{diag}[1, 1]$, $G_2 = \text{diag}[0, 0]$. The integration time was 0.05 seconds and the filter period for the calculation of the augmented error was 0.035 seconds.

Initial estimate of the unknown parameter “ p ” in the system was chosen as $\hat{p}(0) = 1.5$. The following scenario is assumed. Actual parameter value for p was set to 2 for the first 5 seconds of the tracking task. At $t = 5$ sec., the parameter p was suddenly changed to 1 to observe the adaptation performance of the control system. We chose six fixed models, $\hat{\theta}_j \in \{0.9, 1.1, 1.3, 1.7, 1.9, 2.3\}$, for simulation of the proposed scheme.

The nonlinear system in (6.33) was first simulated using the direct adaptive control scheme. Figure 7.2a shows the output of the system. The transient tracking error is very large. The parameter change at $t = 5$ sec. also causes the system output noticeably deviate from the reference trajectory. Figure 7.2e displays the parameter adaptation for the single model direct adaptive control.

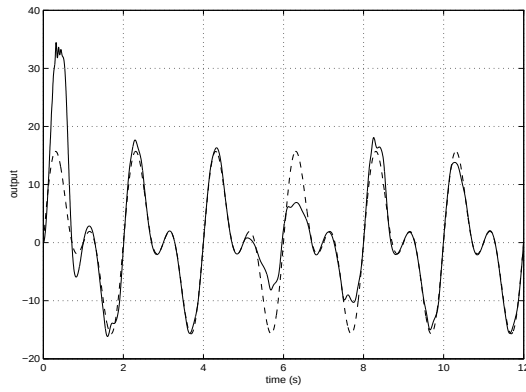
The same system was then simulated using the proposed scheme, in which multiple fixed identification models are incorporated and the over-parametrized parameter vector of the direct scheme was calculated using switched parameters. In Figure 7.2b, the output of the system is shown. It is seen that the initial tracking error is much smaller in this case, and the parameter change at $t = 5$ sec. causes only relatively small deviations from the reference. Parameter adaptation under multiple models scheme is displayed in Figure 7.2f. Effects of switching are noticed as the jumps in the value of parameter estimate $\hat{p}$.

Figure 7.3 displays the switching between the identification models. The index “0” corresponds to the adaptive model. The first switching occurs at the very beginning. This in turn greatly reduces the initial tracking error. The system keeps switching until the adaptive model gives a smaller cost function value than all of the fixed models.

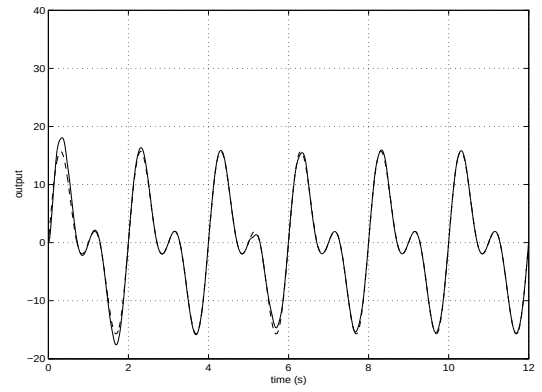
Figure 7.3 displays the switching between the identification models. The index

“0” corresponds to the adaptive model. The first switching occurs at the very beginning. This in turn greatly reduces the initial tracking error. The system keeps switching until the adaptive model gives a lower cost function value than all of the fixed models.

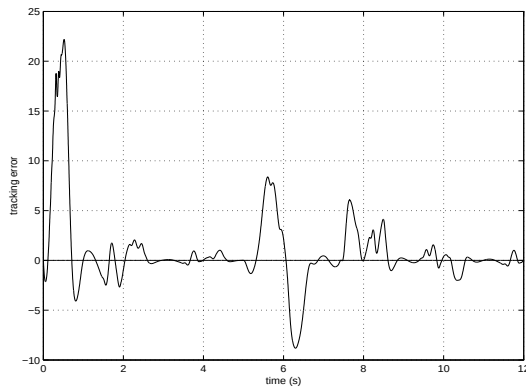
Figure 7.3 displays the switching between the identification models. The index “0” corresponds to the adaptive model. The first switching occurs at the very beginning. This in turn greatly reduces the initial tracking error. The system keeps switching until the adaptive model gives a lower cost function value than all of the fixed models.



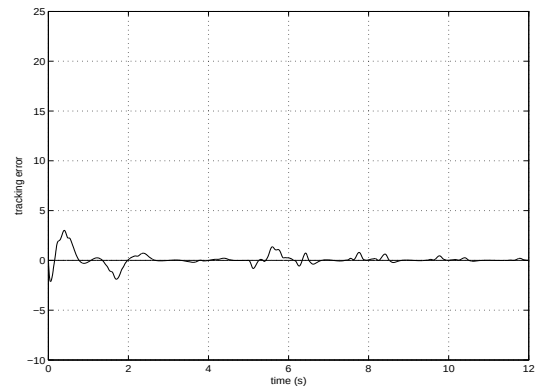
(a) Output y (solid) and y_r (dashed)
(single model)



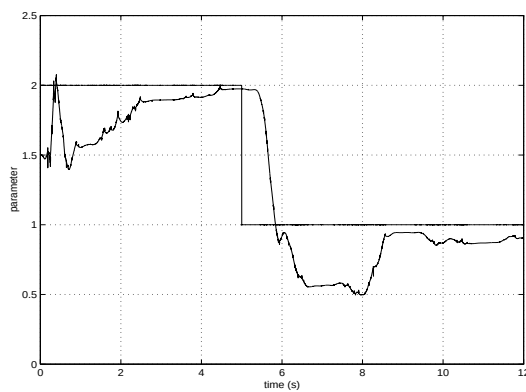
(b) Output y (solid) and y_r (dashed)
(multiple models)



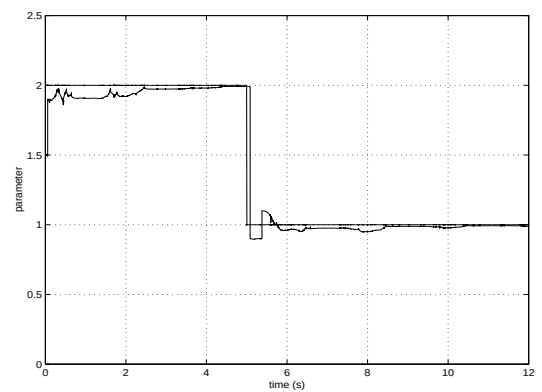
(c) Tracking error e
(single model)



(d) Tracking error e
(multiple models)



(e) Parameter p (dashed), $\hat{p}$ (solid)
(single model)



(f) Parameter p (dashed), $\hat{p}$ (solid)
(multiple models)

Figure 7.2. Simulation results for single model and multiple model cases of the direct adaptive control scheme in Section 7.2.1.

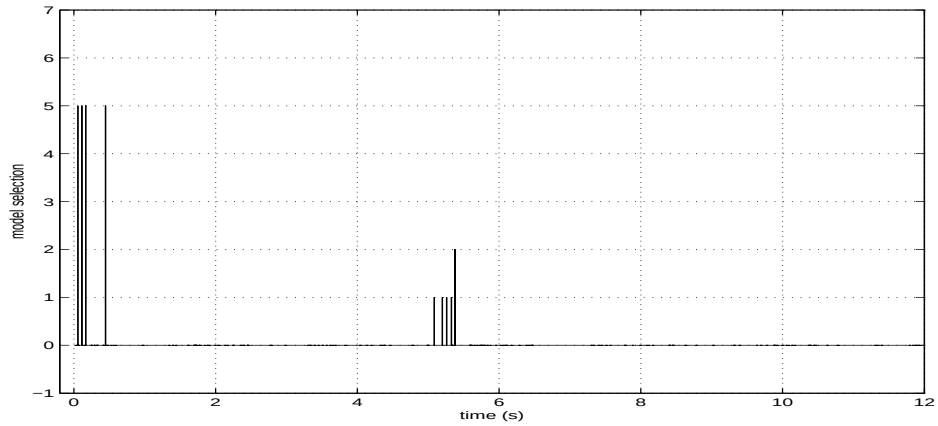


Figure 7.3. Switching between the models (0: adaptive model, 1-6: fixed models) for the direct adaptive control scheme of Section 7.2.1.

7.3. Summary

In this chapter, the multiple models and switching idea was further extended to the direct adaptive control schemes. We presented two MMAC systems based on the multiple identification models and switching. First, a MRAC scheme for a restricted class of nonlinear systems satisfying the matching conditions was modified as a multiple model based scheme. Then, a more general class of nonlinear systems without requiring the matching conditions was considered, and a multiple model based adaptive controller was proposed.

8. COMBINED DIRECT AND INDIRECT ADAPTIVE CONTROL USING MULTIPLE MODELS

Adaptive control schemes are categorized as *direct* and *indirect* approaches. The difference between the two lies mainly on the identification. The indirect approach utilizes a separate identifier which produces estimates of the plant parameters continuously. The *controller parameters* are then calculated using these estimates of the *plant parameters*. In the direct approach, the plant parameters are not identified explicitly. Instead, the controller parameters are identified directly. The parameter update law of the direct approach depends on the tracking error (i.e. output error), while the update law of the indirect scheme is driven by the prediction error (state prediction error, input prediction error, etc.).

The idea of the combined adaptive scheme is based on the observation that, if both the prediction error and the tracking error are used together to update the unknown parameters, the overall estimation might be better, hence leading to better tracking performance. A combined direct and indirect adaptive scheme was first proposed in [56] for linear systems. This approach was later extended to robotic manipulators in [26, 57, 58, 59]. A combined approach for a more general class of nonlinear systems is not yet available in the literature.

In this chapter, we consider the combined direct and indirect adaptive control for a class of nonlinear systems. However, there is a structural limitation on the classes of nonlinear systems for which this combination can be performed. The limitation stems from the *over-parametrization* problem of the direct adaptive scheme. If the over-parametrization is present (as a result of feedback linearization), then the number of parameters to be updated in the direct method is usually larger than those of the indirect method. In fact, in the indirect method, the unknown plant parameters are estimated. Hence there is no over-parametrization. On the other hand, the estimated controller parameters in the direct method are not necessarily the same as the plant

parameters. Moreover, the number of the estimated parameters might be much larger than the number of actual plant parameters due to the over-parametrization.

Because of this limitation, we will analyze the nonlinear systems for which the feedback linearization does not lead to over-parametrization. The systems under study will also be single-input single output (SISO), state-feedback linearizable, and having *matched* uncertainties.

8.1. The Over-parametrization Problem

We consider the same class of nonlinear systems given in Chapter 4 as,

$$\left. \begin{aligned} \dot{x} &= f(x, \theta) + g(x, \theta)u \\ y &= h(x) \end{aligned} \right\} \quad (8.1)$$

where $x \in \mathbb{R}^n$ is the state vector, $f, g \in \mathbb{R}^{n \times p}$ and $h \in \mathbb{R}$ are sufficiently smooth vector fields, $\theta \in \mathcal{S}$ is the unknown parameter vector with $\mathcal{S} \subset \mathbb{R}^p$ being a compact set, $u \in \mathbb{R}$ is the input and $y \in \mathbb{R}$ is the output.

We assume that the nonlinear system defined in (8.1) is linear in its parameters. The vector fields f and g can be written as,

$$f(x, \theta) = \sum_{k=1}^p \theta_k f_k(x) \quad (8.2)$$

$$g(x, \theta) = \sum_{k=1}^p \theta_k g_k(x) \quad (8.3)$$

Without loss of generality, the elements θ_k of the parameter vectors are the same for both vector fields f and g , as given in (8.2) and (8.3).

We put (8.1) into the normal form as,

$$\left. \begin{aligned} \dot{\xi}_1 &= \xi_2 \\ &\vdots \\ \dot{\xi}_\gamma &= L_f^\gamma h(x)(x, \theta) + L_g L_f^{\gamma-1} h(x)(x, \theta) \cdot u \\ \dot{\eta} &= q(\xi, \eta) \\ y &= \xi_1 \end{aligned} \right\} \quad (8.4)$$

where γ is the relative degree of (8.1). Utilizing the linearity-in-the-parameters property, $L_f^\gamma h(x)(x, \theta)$ and $L_g L_f^{\gamma-1} h(x)(x, \theta)$ terms in (8.4) can be expanded in terms of the inner-products of the parameter vector θ and the vector spaces f and g . From (8.2) and (8.3) we have,

$$L_f^\gamma h(x)(x, \theta) = \sum_{i_1}^p \sum_{i_2}^p \cdots \sum_{i_\gamma}^p \frac{\partial}{\partial x} \left(\frac{\partial}{\partial x} (\cdots (\frac{\partial h}{\partial x} f_{i_1}) \cdots) f_{i_{\gamma-1}} \right) f_{i_\gamma} \theta_{i_1} \theta_{i_2} \cdots \theta_{i_\gamma} \quad (8.5)$$

and

$$L_g L_f^{\gamma-1} h(x)(x, \theta) = \sum_j^p \sum_{i_1}^p \cdots \sum_{i_{\gamma-1}}^p \frac{\partial}{\partial x} \left(\frac{\partial}{\partial x} (\cdots (\frac{\partial h}{\partial x} f_{i_1}) \cdots) f_{i_{\gamma-1}} \right) g_j \theta_j \theta_{i_1} \cdots \theta_{i_{\gamma-1}} \quad (8.6)$$

where f_i and g_i are used instead of $f_i(x)$ and $g_i(x)$ for the brevity of notation.

The parameter inner-products that appear in (8.5) and (8.6) lead to the over-parametrization.

Definition 8.1 Consider the nonlinear system in (8.1) with parametric linearity conditions in (8.2) and (8.3), which can be written in the regressor form as,

$$\dot{x} = w^T(x, u)\theta \quad (8.7)$$

where w is the regressor matrix and θ is the parameter vector of the system. Also consider the normal form transformation of (8.1) as given in (8.4), and the “ ξ ” subsystem

which can be written in the regressor form as,

$$\dot{\xi} = w_n^T(\xi, u)\theta_n \quad (8.8)$$

where w_n is the regressor matrix and θ_n is the parameter vector of this linearized subsystem. The system in (8.1) is said to have **no over-parametrization** if,

$$\theta \equiv \theta_n \quad (8.9)$$

Example 8.1 Consider the following system,

$$\left. \begin{aligned} \dot{x}_1 &= x_1^2 + \theta_1 x_2 + \theta_2 \\ \dot{x}_2 &= x_1 + u \\ y &= x_1 \end{aligned} \right\} \quad (8.10)$$

The relative degree of this system is $\gamma = 2$. It is written in the regressor form as,

$$\dot{x} = \begin{bmatrix} x_1^2 & x_2 & 1 \\ x_1 + u & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ \theta_1 \\ \theta_2 \end{bmatrix} \quad (8.11)$$

Using input-output linearization, we have state transformations $\xi_1 = h = x_1$ and $\xi_2 = L_f h = x_1^2 + \theta_1 x_2 + \theta_2$. Taking the derivatives and putting the linearized system in the regressor form, we obtain,

$$\dot{\xi} = \begin{bmatrix} \xi_2 & 0 & 0 \\ 2\xi_1\xi_2 & \xi_1 + u & 0 \end{bmatrix} \begin{bmatrix} 1 \\ \theta_1 \\ \theta_2 \end{bmatrix} \quad (8.12)$$

We see that the parameter vectors of (8.11) and (8.12) are identical. Notice that θ_2 in (8.12) is multiplied by zeros in the regressor, i.e. θ_2 is missing in the linearized

dynamics. However, the certainty equivalence linearizing control law is found to be,

$$u = \frac{1}{\hat{\theta}_1}(-2\xi_1\xi_2 - \hat{\theta}_1\xi_1 + v) \quad (8.13)$$

When this control law is put into the linearized dynamics, the $\dot{\xi}_2$ becomes,

$$\dot{\xi}_2 = 2\xi_1\xi_2 + \theta_1\xi_1 + \frac{\theta_1}{\hat{\theta}_1}(-2\xi_1\xi_2 - \hat{\theta}_1\xi_1 + v) \quad (8.14)$$

The term $\frac{\theta_1}{\hat{\theta}_1}$ is one of the reasons preventing us from writing error dynamics which would be otherwise linear in $\tilde{\theta}$. The other reason is that the system in (8.10) does not satisfy the matching conditions, hence $\dot{y}$ is not available which is required to calculate v .

The above example demonstrates that even though a nonlinear dynamic system guarantees the “no over-parametrization” according to the Definition 8.1, this is not sufficient to formulate a combined adaptive control scheme. If the closed loop error dynamics of the nonlinear system in (8.1) can be written as,

$$\dot{e} = \tilde{\theta}^T \varphi(x) \quad (8.15)$$

where e is the tracking error vector and $\tilde{\theta} = \hat{\theta} - \theta$ is the parameter error vector, then we can apply the combined direct and indirect adaptive control scheme to the nonlinear system.

Theorem 8.1 *Consider the parametric system in (8.1) and its feedback linearization in (8.4), with linear parameter dependence as stated in (8.2) and (8.3). A combined direct and indirect adaptive control scheme can be applied for this system, if,*

1. $\frac{\partial}{\partial \theta} \left[\frac{\partial}{\partial x} L_f^{\gamma-1} h(x, \theta) \right] = 0$
2. $\frac{\partial}{\partial \theta} \left[L_g L_f^{\gamma-1} h(x, \theta) \right] = 0$
3. The matching conditions given in Definition 5.1 hold.

Proof: The nonlinear system given in (8.1) is put into the normal form in (8.4) via the diffeomorphism ϕ given by,

$$\phi(x, \theta) = \begin{bmatrix} \xi \\ \eta \end{bmatrix} = \begin{bmatrix} h(x) \\ L_f h(x) \\ \vdots \\ L_f^{\gamma-1} h(x) \\ \eta_1 \\ \vdots \\ \eta_{m-\gamma} \end{bmatrix} \quad (8.16)$$

Hence,

$$L_f^\gamma h(x)(x, \theta) = L_f \underbrace{L_f^{\gamma-1} h(x)}_{\xi_\gamma}(x, \theta) = L_f \xi_\gamma \quad (8.17)$$

and

$$L_g \underbrace{L_f^{\gamma-1} h(x)}_{\xi_\gamma}(x, \theta) = L_g \xi_\gamma \quad (8.18)$$

The ξ -subsystem of (8.4) can be written as,

$$\left. \begin{aligned} \dot{\xi}_1 &= \xi_2 \\ &\vdots \\ \dot{\xi}_\gamma &= L_f \xi_\gamma + L_g \xi_\gamma \cdot u \end{aligned} \right\} \quad (8.19)$$

Let us write $\dot{\xi}_\gamma$ in (8.19) using (8.2) and (8.3) as,

$$\dot{\xi}_\gamma = \frac{\partial \xi_\gamma}{\partial x} \sum_{k=1}^p f_k(x) \theta_k + \frac{\partial \xi_\gamma}{\partial x} \sum_{k=1}^p g_k(x) \theta_k \cdot u \quad (8.20)$$

Equivalently, we have,

$$\dot{\xi}_\gamma = \sum_{k=1}^p \frac{\partial \xi_\gamma}{\partial x} f_k(x) \theta_k + \sum_{k=1}^p \frac{\partial \xi_\gamma}{\partial x} g_k(x) \theta_k \cdot u \quad (8.21)$$

Due to the second condition of the theorem, the term $\sum_{k=1}^p \frac{\partial \xi_\gamma}{\partial x} g_k(x) \theta_k$ is free of the parameter vector θ . Thus, we can write,

$$\dot{\xi}_\gamma = \sum_{k=1}^p \frac{\partial \xi_\gamma}{\partial x} f_k(x) \theta_k + \sum_{k=1}^p \frac{\partial \xi_\gamma}{\partial x} g_k(x) \cdot u \quad (8.22)$$

It is observed that if the term $\frac{\partial \xi_\gamma}{\partial x}$ is not a function of the elements θ_k of the parameter vector θ , then no parameter inner-products will exist in the linearization. The first condition of the theorem states that $\frac{\partial \xi_\gamma}{\partial x}$ is free of the parameters.

Now let us suppose that $\frac{\partial \xi_\gamma}{\partial x}$ is a function of the parameters but f is not. In this case there would be no parameter inner-products in (8.22). However this is not possible, because,

$$\xi_\gamma = L_f^{\gamma-1} h = \frac{\partial}{\partial x} (\cdots (\frac{\partial h}{\partial x} f) \cdots) f \quad (8.23)$$

Hence if ξ_γ is a function of the parameter vector, so is f . Therefore the first condition is also necessary. Finally, the matching conditions are required to be able to calculate the nominal tracking control signal v . The certainty equivalence linearizing control is given as,

$$u = \frac{1}{\sum_{k=1}^p \frac{\partial \xi_\gamma}{\partial x} g_k(x)} \left(- \sum_{k=1}^p \frac{\partial \xi_\gamma}{\partial x} f_k(x) \hat{\theta}_k + v \right) \quad (8.24)$$

When plugged into (8.19), we obtain,

$$\left. \begin{aligned} \dot{\xi}_1 &= \xi_2 \\ &\vdots \\ \dot{\xi}_\gamma &= - \sum_{k=1}^p \frac{\partial \xi_\gamma}{\partial x} f_k(x) \hat{\theta}_k + \sum_{k=1}^p \frac{\partial \xi_\gamma}{\partial x} f_k(x) \hat{\theta}_k + v \end{aligned} \right\} \quad (8.25)$$

which leads to,

$$\left. \begin{aligned} \dot{\xi}_1 &= \xi_2 \\ &\vdots \\ \dot{\xi}_\gamma &= -\sum_{k=1}^p \frac{\partial \xi_\gamma}{\partial x} f_k(x)(\hat{\theta}_k - \theta_k) + v \end{aligned} \right\} \quad (8.26)$$

Defining the tracking error as $e = \xi_1 - y_r$, and the nominal tracking control signal v as,

$$v = y_r^{(\gamma)} + \lambda_1(y_r^{(\gamma-1)} - y^{(\gamma-1)}) + \cdots + \lambda_\gamma(y_r - y) \quad (8.27)$$

we get,

$$e^{(\gamma)} + \lambda_1 e^{(\gamma-1)} + \cdots + \lambda_\gamma e = \tilde{\theta}^T \varphi \quad (8.28)$$

which is in the form of (8.15). □

Remark 8.1 *It should be noted that, due to the matching conditions, the derivatives of the output $\dot{y}, \ddot{y}, \dots, y^{(\gamma-1)}$ in (8.27) are available, since $L_f h, L_f^2 h, \dots, L_f^{\gamma-1} h$ do not contain any uncertainties. Therefore, unlike in [8], there is no error term here stemming from calculation of “ v ”.*

Next we derive a combined direct/indirect adaptive controller for a class of nonlinear systems which satisfy the conditions stated in Theorem 8.1.

8.2. Direct and Indirect Adaptive Control of Nonlinear Systems

Theorem 8.1 suggests that nonlinear systems in the form of

$$\left. \begin{aligned} \dot{x} &= f(x) + g(x)u + \sum_{k=1}^p q_k(x)\theta_k \\ y &= h(x) \end{aligned} \right\} \quad (8.29)$$

can be controlled using a combined direct/indirect adaptive scheme. For the system given in (8.29), $x \in \mathbb{R}^n$ is the state vector, $f, g \in \mathbb{R}^n$ are known smooth vector fields with $f(0) = 0$ and $g(0) \neq 0$, and $\theta = \begin{bmatrix} \theta_1 & \theta_2 & \dots & \theta_p \end{bmatrix}^T$ is the constant parameter vector which takes values on a compact set $\mathcal{S} \subset \mathbb{R}^p$ and which confines all the uncertainties for the given system. $q_k(x) \in \mathbb{R}^n$ are known vector fields, $u \in \mathbb{R}$ is the input to the system, $y \in \mathbb{R}$ is the output of the system and $h : \mathbb{R}^n \mapsto \mathbb{R}$ is a known smooth function, with $h(0) = 0$ such that it gives a diffeomorphism which fully linearizes the nominal part of (8.29). The origin is assumed to be the isolated equilibrium point of (8.29) without loss of generality. We also assume that the state vector x is available for feedback, and the relative degree of the system in (8.29) is n . A reference model which does not depend on the parameter vector θ is given and the control objective is to design an adaptive controller such that tracking error between the states of the system and those of the reference model goes to zero asymptotically.

8.2.1. Direct Adaptive Scheme

For the derivation of a direct adaptive scheme, we apply the same steps as in Section 5.4. The resulting error dynamics is obtained as,

$$\dot{e} = A_r e + b_r \Psi^T(\xi) \tilde{\theta} \quad (8.30)$$

where e is the state error vector, $\Psi \in \mathbb{R}^p$ is the regressor vector, and (A_r, b_r) is a controllable pair as defined in (5.132). The Lyapunov analysis yields the parameter update law given in (5.138) as,

$$\dot{\hat{\theta}} = -\Gamma \Psi b_r^T P e \quad (8.31)$$

where $\Gamma \in \mathbb{R}^{p \times p}$ is a positive-definite gain matrix, and $P \in \mathbb{R}^{n \times n}$ is the symmetric, positive-definite solution of the Lyapunov equation in (5.136).

Note that the parameter update dynamics defined in (8.31) is driven by the tracking error vector e and the updated parameters are actually the *controller parameters*

to be directly used in obtaining the control signal.

8.2.2. Indirect Adaptive Control and Nonlinear System Identification

The original system dynamics defined in (8.29) can be written in the form,

$$\dot{x} = w^T(x, u)\Theta \quad (8.32)$$

where $w^T \in \mathbb{R}^{n \times (p+1)}$ is the regressor matrix, and $\Theta = \begin{bmatrix} 1 & \theta_1 & \dots & \theta_p \end{bmatrix}^T$ is the parameter vector. We apply the regressor filtering as given in Section 5.1.2 with the filtered forms of x and w denoted by x_f and w_f , respectively, and obtain the state estimate model as,

$$\hat{x} = w_f^T \hat{\Theta} + x_f \quad (8.33)$$

We can then define the *observation or identification error* $\tilde{x}$ as $\tilde{x} = \hat{x} - x$. Then using (8.33) as the state estimate and the available state information x , we obtain an error model as,

$$\tilde{x}(t) = w_f^T(x, u)\tilde{\Theta} = w_f^T(t)(\hat{\Theta} - \Theta) \quad (8.34)$$

which is also equal to,

$$\tilde{x}(t) = \bar{w}_f^T(x, u)\tilde{\theta} = \bar{w}_f^T(t)(\hat{\theta} - \theta) \quad (8.35)$$

where $\bar{w}_f$ is the filtered regressor with its first column dropped and $\tilde{\theta} = \hat{\theta} - \theta$ is the parameter error vector for the identification model (since the first element of $\tilde{\Theta}$ becomes zero). Note that $\tilde{x}$ is readily available since the state x was already assumed to be known and its estimate is obtained from (8.33) using the filtered regressor set-up.

Now that a linear error model (8.35) is available for identification, one can choose from a list of various identification schemes to estimate the parameter vector θ of the

system dynamics defined in (8.29). Here we choose a recursive least squares estimation (RLS) scheme with the following parameter update mechanism,

$$\left. \begin{aligned} \dot{\tilde{\theta}} &= -\Gamma \bar{w}_f x \\ \dot{\Gamma} &= -\Gamma \bar{w}_f \bar{w}_f^T \Gamma \end{aligned} \right\} \quad (8.36)$$

where Γ is the uniformly positive definite gain matrix, sometimes referred to as the covariance matrix. The update mechanism for Γ can also be written as,

$$\dot{\Gamma}^{-1} = +\bar{w}_f \bar{w}_f^T \quad (8.37)$$

If the time varying gain matrix Γ is set to a constant, then we get the well celebrated gradient update law. The convergence of the above RLS estimation scheme can be seen by picking a Lyapunov function $V(\tilde{\theta}) = \tilde{\theta}^T \Gamma^{-1} \tilde{\theta}$ and taking its time derivative. It gives $\tilde{x} \in \mathcal{L}_\infty$ and $\tilde{\theta} \in \mathcal{L}_\infty$. Moreover, if the regressor $\bar{w}_f$ is sufficiently rich, leading to persistent excitation, then both $\tilde{x}$ and $\tilde{\theta}$ converge to zero asymptotically.

8.3. Combined Direct and Indirect Adaptive Scheme

Combined adaptive scheme is derived based on the observation that the parameter uncertainty is reflected in both the tracking error e , and the identification error $\tilde{x}$. This is clearly seen in the tracking error dynamics defined in (5.134) and the identification error model in (8.35). Therefore we propose to extract the parameter information from both sources to increase parameter convergence performance. The key point here is the fact that the parametrization of the error dynamics and the identification dynamics yield the same parameter error vector $\tilde{\theta}$ as a direct consequence of Definition 8.1 and Theorem 8.1

The control signal u and the auxiliary input v for the combined scheme are the same certainty equivalence based control signals given in Section 5.4 as,

$$u = \frac{1}{L_g L_f^{n-1} h(x)} (-L_f^n h(x) + v) \quad (8.38)$$

and

$$v = - \sum_{i=1}^n a_{i-1} \xi_i - \Psi^T(\xi) \hat{\theta} + v_r \quad (8.39)$$

respectively. Using the same reference dynamics given in (5.131), the resulting error dynamics for the system in (8.29) becomes the same as the tracking error dynamics in (8.30). We propose a new parameter update scheme where tracking error and identification error are used in a combined form as,

$$\dot{\hat{\theta}} = -\Gamma(\Psi b_r^T P e + K \bar{w}_f \tilde{x}) \quad (8.40)$$

where Γ is the symmetric positive definite gain matrix of the RLS algorithm defined in (8.36), $K = \alpha I$ is a positive definite matrix where the parameter $\alpha > 0$ determines the weight of the prediction error in the update law. With this proposed parameter update dynamics, we next show the overall stability of the closed loop adaptive scheme.

Theorem 8.2 *For the dynamic system given in (8.29), the adaptive control scheme with the control input given in (8.38) along with the direct/indirect parameter update dynamics defined in (8.40) results in a stable closed loop system with $e \in \mathcal{L}_2 \cap \mathcal{L}_\infty$, $\tilde{x} \in \mathcal{L}_\infty$, $\tilde{\theta} \in \mathcal{L}_\infty$, and furthermore tracking error e and identification error $\tilde{x}$ converge to zero asymptotically.*

Proof: We choose a Lyapunov function for the closed loop error dynamics as,

$$V = e^T P e + \tilde{\theta}^T \Gamma^{-1} \tilde{\theta} \geq 0 \quad (8.41)$$

where P is the solution to the Lyapunov equation in (5.136). Taking the time derivative of V and using the tracking error dynamics defined in (8.30), we obtain,

$$\dot{V} = e^T \underbrace{(A_r^T P + P A_r)}_{-Q} e + 2e^T P b_r \Psi^T \tilde{\theta} + 2\tilde{\theta}^T \Gamma^{-1} \dot{\tilde{\theta}} + \tilde{\theta}^T \dot{\Gamma}^{-1} \tilde{\theta} \quad (8.42)$$

Plugging (8.40) and (8.37) in (8.42), we get,

$$\dot{V} = -e^T Q e + 2e^T (P b_r \Psi^T) \tilde{\theta} - \underbrace{2\tilde{\theta}^T (\Psi b_r^T P) e}_{2e^T (P b_r \Psi^T) \tilde{\theta}} - 2\tilde{\theta}^T K \bar{w}_f \tilde{x} + \tilde{\theta}^T \bar{w}_f \bar{w}_f^T \tilde{\theta} \quad (8.43)$$

where the second and the third term cancel each other. Using $K = \alpha I$, we get,

$$\dot{V} = -e^T Q e - 2\alpha \tilde{\theta}^T \bar{w}_f \tilde{x} + \tilde{\theta}^T \bar{w}_f \bar{w}_f^T \tilde{\theta} \quad (8.44)$$

Note that $\tilde{\theta}^T \bar{w}_f = \tilde{x}^T$ from the identification error model in (8.35), hence we obtain,

$$\dot{V} = -e^T Q e - \underbrace{(2\alpha - 1)}_{\beta} \tilde{x}^T \tilde{x} \quad (8.45)$$

If we set the gain $\alpha \geq 0.5$, we have $\beta \geq 0$ and we get,

$$\dot{V} = -e^T Q e - \beta \tilde{x}^T \tilde{x} \leq 0 \quad (8.46)$$

This leads to $e \in \mathcal{L}_2 \cap \mathcal{L}_\infty$, $\tilde{\theta} \in \mathcal{L}_\infty$, and $\tilde{x} \in \mathcal{L}_2$. Next we have to evaluate $\ddot{V}$ in order to show the convergence of e and $\tilde{x}$. Taking the time derivative of (8.46), we get,

$$\ddot{V} = -2e^T Q \dot{e} - 2\beta \tilde{x}^T \dot{\tilde{x}} \quad (8.47)$$

Since $e = \xi - z$, $e \in \mathcal{L}_\infty$ leads to $\xi \in \mathcal{L}_\infty$ and $\Psi(\xi) \in \mathcal{L}_\infty$. We already have $\tilde{\theta} \in \mathcal{L}_\infty$, hence we get $\dot{e} \in \mathcal{L}_\infty$ from the tracking error dynamics in (8.30).

We also get boundedness of the original system state $x \in \mathcal{L}_\infty$, as $\xi \in \mathcal{L}_\infty$ and ϕ is a diffeomorphism ($\xi = \phi(x)$). Hence the regressor $w(x) \in \mathcal{L}_\infty$, and accordingly the filtered regressor $\bar{w}_f(x) \in \mathcal{L}_\infty$ (due to the filter is exponentially stable). This leads to $\tilde{x} \in \mathcal{L}_\infty$ as well from the error equation (8.35), since $\tilde{\theta} \in \mathcal{L}_\infty$. To show the boundedness of $\dot{\tilde{x}}$, we differentiate (8.35) to obtain $\dot{\tilde{x}} = \bar{w}_f^T \dot{\tilde{\theta}} + \dot{\bar{w}}_f^T \tilde{\theta}$. We have $\dot{\tilde{\theta}} \in \mathcal{L}_\infty$ from update dynamics in (8.40) and $\dot{\bar{w}}_f^T \in \mathcal{L}_\infty$ from the smooth and bounded regressor matrix $\bar{w}_f$ and bounded state x . Hence we get $\dot{\tilde{x}} \in \mathcal{L}_\infty$. Since all signals in (8.47) are bounded,

we have $\ddot{V} \in \mathcal{L}_\infty$ which leads to uniform continuity of $\dot{V}$. Using the boundedness and uniform continuity of $\dot{V}$ with Barbalat's lemma, we get $\dot{V} \rightarrow 0$. This leads to $e \rightarrow 0$ and $\tilde{x} \rightarrow 0$, which also leads to $\lim_{t \rightarrow \infty} |x(t) - \phi^{-1}(z)| = 0$. $\square$

In the above result, the Lyapunov function is the same as the one used in the direct adaptive controller. However, its derivative now includes an additional negative term which indicates faster convergence. The combined scheme leads to the convergence of both the tracking error e and identification error $\tilde{x}$. Convergence of the parameter error vector $\tilde{\theta}$ requires the persistency of excitation.

Since our adaptive control scheme is based on the certainty equivalence principle, the asymptotic convergence of the identification error would also lead to smaller peaks in the transient response of the tracking error. This has been verified by simulations and averaging effect of the proposed combined scheme has been discussed in [32]. Now we use this combined scheme for the case of multiple identification models and switching to further improve the transient response.

8.4. Combined Direct and Indirect Adaptive Control Using Multiple Models

In this section we extend the combined adaptive scheme developed in Section 8.3 to a multiple models based adaptive control system. First, we define N fixed estimation models, in the form of (8.35) as,

$$\tilde{x}_j = \overline{w}_f^T(x, u) \tilde{\theta}_j \quad (8.48)$$

where $\tilde{x}_j$ is the state estimation error, and $\tilde{\theta}_j = \hat{\theta}_j - \theta \in \mathcal{S} \subset \mathbb{R}^p$ is the fixed parameter error vector of the j^{th} -model, with $j = 1, \dots, N$. We employ the permissible switching sequence as defined in Definition 6.1. If a switching occurs at time t , the parameter estimate vector $\hat{\theta}(t)$ is instantaneously reset to $\hat{\theta}_j$, where j denotes the index of the selected model.

We use the same Lyapunov function for the closed loop error dynamics (defined by the equations (8.30) and (8.40)), as,

$$V(t) = e^T P e + \tilde{\theta}^T \Gamma^{-1} \tilde{\theta} \quad (8.49)$$

Assuming that there is a parameter reset at the switching instants, the change in the Lyapunov function is given by,

$$\Delta V = V(t + T_{min}) - V(t) \quad (8.50)$$

where $T_{min} > 0$ is the constant dwell time. For simplicity, we will denote $(t + T_{min})$ by (t^+) . Since we have N fixed models with the associated identification error equations defined in (8.48), the change in the Lyapunov function for the N fixed models is defined as,

$$\Delta V_j = V_j(t^+) - V(t) \quad , \quad j = 1, \dots, N \quad (8.51)$$

where $V_j(t^+)$ is the Lyapunov function computed at the resetting instants with $\hat{\theta}_j(t^+)$ as,

$$V_j(t^+) = e^T(t^+) P e(t^+) + \tilde{\theta}_j^T(t^+) \Gamma^{-1} \tilde{\theta}_j(t^+) \quad (8.52)$$

where $\tilde{\theta}_j = \hat{\theta}_j - \theta$. The switching logic is now based on the following condition,

$$\Delta V_j = V_j(t^+) - V(t) < 0 \quad , \quad j = 1, \dots, N \quad (8.53)$$

Note that $V_j(t^+)$ is the same as $V(t)$ with only one difference that the former is computed using $\hat{\theta}_j(t^+)$ and the latter is computed using $\hat{\theta}(t)$. Substituting (8.49) and (8.52)

in (8.53), we have,

$$\Delta V_j = e^T(t^+)Pe(t^+) - e^T(t)Pe(t) + (\hat{\theta}_j(t^+) - \theta)^T \Gamma^{-1}(\hat{\theta}_j(t^+) - \theta) - (\hat{\theta}(t) - \theta)^T \Gamma^{-1}(\hat{\theta}(t) - \theta) \quad (8.54)$$

If $\Delta V_j \leq -\kappa$ for some $\kappa > 0$ is satisfied, then the parameter estimates of the adaptive model are reset to $\hat{\theta}_j$, and adaptation is restarted from that point. If more than one model makes $\Delta V_j \leq -\kappa$, then the model providing the largest κ value is selected. If no models gives a negative jump in the Lyapunov function, then no switching takes place, and adaptation continues without any resets.

Since $\hat{\theta}_j$'s are fixed, we have $\hat{\theta}_j(t^+) = \hat{\theta}_j(t) = \hat{\theta}_j$, and we can rewrite (8.54) as,

$$\Delta V_j = e^T(t^+)Pe(t^+) - e^T(t)Pe(t) + (\hat{\theta}_j - \theta)^T \Gamma^{-1}(\hat{\theta}_j - \theta) - (\hat{\theta}(t) - \theta)^T \Gamma^{-1}(\hat{\theta}(t) - \theta) \quad (8.55)$$

Following a similar derivation as in Section 7.1, we can also write (8.55) as,

$$\begin{aligned} \Delta V_j = & e^T(t^+)Pe(t^+) - e^T(t)Pe(t) - (\hat{\theta}_j - \hat{\theta}(t))^T \Gamma^{-1}(\hat{\theta}_j - \hat{\theta}(t)) \\ & + 2(\hat{\theta}_j - \hat{\theta}(t))^T \Gamma^{-1}(\hat{\theta}_j - \theta) \end{aligned} \quad (8.56)$$

where θ is the actual parameter value and is not available for computation. Here we use a similar approach discussed in Section 7.1 to evaluate (8.56). Using the identification error given in (8.48), we obtain,

$$\underbrace{\int_{t-\delta}^t \bar{w}_f(\tau) \tilde{x}_j(\tau) d\tau}_{\Upsilon_j(t)} = \underbrace{\int_{t-\delta}^t \bar{w}_f(\tau) w_f^T(\tau) d\tau}_{\Phi(t)} \cdot (\hat{\theta}_j - \theta) \quad (8.57)$$

where $\delta > 0$ is a constant time window. Pulling $(\hat{\theta}_j - \theta)$ from (8.57), ΔV_j can now be computed as,

$$\begin{aligned} \Delta V_j = & e^T(t^+)Pe(t^+) - e^T(t)Pe(t) - (\hat{\theta}_j - \hat{\theta}(t))^T \Gamma^{-1}(\hat{\theta}_j - \hat{\theta}(t)) \\ & + 2(\hat{\theta}_j - \hat{\theta}(t))^T \Gamma^{-1} \Phi^{-1}(t) \Upsilon_j(t) \end{aligned} \quad (8.58)$$

Now, having ΔV_j available for all N fixed models, we state the switching logic in the following definition.

Definition 8.2 *With N fixed identification models defined with the error equations, (as in (8.48)),*

$$\tilde{x}_j = \bar{w}_f^T \tilde{\theta}_j \quad (8.59)$$

where $\tilde{\theta}_j = \hat{\theta}_j - \theta$ with $\hat{\theta}_j$, $j = 1, \dots, N$, being the fixed parameter estimates, and with a permissible switching time T_{min} given in Definition 6.1, the switching logic which would lead to transient performance improvement given as,

$$j^* = \begin{cases} j : \min \Delta V_j & , \quad \Delta V_j = V_j(t + T_{min}) - V(t) < 0 \\ 0 & , \quad \Delta V_j = V_j(t + T_{min}) - V(t) \geq 0 \end{cases} \quad , \quad j = 1, \dots, N \quad (8.60)$$

where j^* gives the index for the chosen identification model. $V(t)$ is the Lyapunov function given in (8.49), $V_j(t + T_{min})$ is the Lyapunov function given in (8.52), which is evaluated with fixed parameter estimates at the switching instants.

Under this switching rule, the parameter estimate vector $\hat{\theta}$ and the parameter error vector $\tilde{\theta}$, will take the form, respectively, as $\hat{\theta}_{z(t)}$ and $\tilde{\theta}_{z(t)}$, where $z(t)$ is defined as $z(t) : \mathbb{R}_+ \mapsto 0, \dots, N$ such that if $t \in [T_i, T_{i+1})$ for some $i < \infty$, then $z(t) = z(T_i)$. Here, T_i is the permissible switching sequence defined in Definition 6.1 and the index “0” corresponds to the adaptive model. $\hat{\theta}_{z(t)}$ is now a piecewise continuous signal with possible jumps at the switching instants.

Using the switched form of parameter estimate vector $\hat{\theta}_{z(t)}$, the control signal keeps the same form as in (8.38), but the tracking control input v now becomes,

$$v = - \sum_{i=1}^n a_{i-1} \xi_i - \Psi^T(\xi) \hat{\theta}_{z(t)} + v_r \quad (8.61)$$

Substituting this control signal in the system dynamics, we obtain the closed loop error dynamics as,

$$\dot{e} = A_r e + b_r \Psi^T(\xi) \tilde{\theta}_{z(t)} \quad (8.62)$$

along with the update equation given in (8.40) and the switching logic defined in Definition 8.2.

Theorem 8.3 *Consider the combined direct and indirect adaptive scheme with the control signal given in (8.38), (8.61), and the adaptation dynamics given in (8.40). With N fixed identification models defined in (8.48) and the switching logic for parameter update dynamics defined in Definition 8.2, the closed loop error dynamics in (8.62) is asymptotically stable, such that $y \rightarrow y_r$ as $t \rightarrow \infty$, with improved transient performance.*

Proof: The Lyapunov function for the closed loop error dynamics is given in (8.49) as,

$$V(t) = e^T P e + \tilde{\theta}^T \Gamma^{-1} \tilde{\theta} \quad (8.63)$$

Since the parameter estimate vector $\hat{\theta}(t)$ can take values from a fixed parameter set $\{\hat{\theta}_j, j = 1, \dots, N\}$ at the switching instants, we compute the change in the Lyapunov function as given in (8.58).

Based on Branicky's theorem for multiple Lyapunov functions as given in Theorem 6.1, if,

$$V_j(x(T_{i+1})) < V_j(x(T_i))$$

holds for all j where T_i is the permissible switching sequence given in Definition 6.1, then the switching system is stable in the sense of Lyapunov. The performance im-

proving and stability preserving condition is given as,

$$\Delta V_j < 0 \quad , \quad j = 1, \dots, N \quad (8.64)$$

The model index j^* chosen by the switching logic in (8.60) gives us the largest negative ΔV_j . This condition actually causes a maximum negative jump in the Lyapunov function in (8.49). Taking the similar steps as in the proof of Theorem 8.2, we will get,

$$\dot{V} = -e^T Q e - \underbrace{(2\alpha - 1)}_{\beta} \tilde{x}^T \tilde{x} \quad (8.65)$$

Setting the gain $\alpha \geq 0.5$, we have $\beta = 2\alpha - 1 \geq 0$, and we obtain,

$$\dot{V} = -e^T Q e - \beta \tilde{x}^T \tilde{x} \leq 0 \quad (8.66)$$

This leads to $e \in \mathcal{L}_2 \cap \mathcal{L}_\infty$ and $\tilde{\theta} \in \mathcal{L}_\infty$. Using similar arguments as in the proof of Theorem 8.2, we get $e \rightarrow 0$ as $t \rightarrow \infty$. $\square$

Remark 8.2 *Since we require non-singularity for $\Phi(t)$ for the computation of ΔV_j given in (8.58), this is simply equivalent to persistent excitation for the reference input which also leads to the asymptotic convergence of $\tilde{\theta}$. The resetting condition defined in (8.60) simply speeds up convergence of $\hat{\theta}$ to θ and hence leading to lower transient errors.*

Remark 8.3 *If none of the N fixed models results in a negative jump in the Lyapunov function given in (8.49), then the resetting condition in (8.60) gives $j^* = 0$, which corresponds to the adaptive model. In this case, there is no switching and the adaptation continues with the adaptive model.*

8.5. Summary

The indirect adaptive schemes use separate identifiers for parameter adaptation. These identifiers are driven by the identification error. On the other hand, in the direct adaptive schemes, control parameters are updated using the output (tracking) error. The motivation for the combined direct and indirect scheme is based on the fact that, using both the identification error and the tracking error may result in a faster adaptation, which would lead to smaller transient errors.

The main drawback of this scheme is that it cannot be applied to all nonlinear systems due to the over-parametrization problem. To be able to use the combined scheme, the parameter vector for the direct and the indirect schemes must be identical. In this chapter, first, *a priori* conditions for the applicability of combined schemes for a general class of nonlinear systems are derived. Then a combined adaptive controller is proposed for a specific class of nonlinear systems along with its stability properties.

To improve the transient error performance, multiple models scheme was applied to this combined adaptive controller. Stability and performance enhancement of the combined scheme using multiple models was also shown.

9. CONCLUSIONS

Adaptive control of nonlinear systems was considered in this study. The main objective of the thesis was to improve the transient performance of the adaptive controllers for nonlinear systems. Today, a major drawback of adaptive control systems is the unpredictable transient behavior and slow adaptation. Especially when there are large changes in the system parameters, unacceptably large transient errors might be present. This problem is even more crucial in the control of nonlinear plants. High adaptation gains cannot be used, as high gains increase the noise sensitivity and may give rise to instability. In this thesis, we propose the use of multiple identification models and switching to improve the transient response, based on the principle that the better we know the actual system parameters, the better response we can get from our controllers.

The possible values of the parameters of a dynamic system is usually known in advance. The idea in the multiple model approach is to include these parameter values as fixed reference points in the controllers. During the execution of the adaptive controller, at each determined time interval, a previously defined cost function for the adaptive model and for all the fixed models is evaluated. This cost function is based on the parameter estimates. If any one of the fixed models minimizes this cost function, then the adaptation is re-initialized from the parameters of this corresponding fixed model. This can be visualized in the parameter space as jumping from one point to another point whenever it is decided that the target point is closer to where the actual system parameters lie. This in turn leads to faster convergence of the parameter vector to its true values and hence enhanced transient response of the controller. Another switching method, which is based on multiple Lyapunov functions of the closed loop system and maximum negative jump criterion in these Lyapunov functions is also discussed.

Previous results on the adaptive control using multiple models were limited to linear systems and only a few nonlinear systems in special forms. The systems studied

in this thesis cover relatively a large class of SISO, minimum-phase nonlinear systems. We developed a multiple model control methodology for the control-affine nonlinear systems without requiring the restrictive matching conditions. On the other hand, currently available methods on multiple models based adaptive control utilize indirect adaptive schemes. In this thesis, we presented direct adaptive control mechanisms utilizing multiple identification models along with stability proofs, for matched and unmatched uncertainties.

The system identification errors are essential for switching in the multiple models schemes. However, as the main focus being better adaptation, it is logical to use both direct and the indirect approaches in a combined fashion. The direct approach makes use of the tracking error, whereas the indirect approach uses the identification errors for parameter adaptation. The obstacle for such a combined methodology is the over-parametrization problem. We developed the conditions under which the parametrization of the direct and the indirect approaches becomes identical for a class of nonlinear systems. Under these conditions, we proposed a combined direct and indirect adaptive control scheme and further extended this approach for the case of multiple models and switching.

It is shown both theoretically and experimentally by computer simulations that the proposed methodologies which utilize multiple models and switching yield remarkable improvement in the transient response. We believe this thesis presents an important contribution for the development of multiple models and switching concepts for the adaptive control of a large class of nonlinear systems.

The future work in this field may be on further extending the theory for a more general class of nonlinear systems without requiring the assumptions used in this thesis. Some areas for further research may include the following:

- The full state measurement requirement may be relaxed, or output-feedback linearization may be considered,
- Nonlinearly parameterized systems may be considered and multiple artificial neu-

ral networks may be used as adaptive models for system modeling,

- An intelligent learning mechanism can be incorporated such that new models are generated/deleted through experience.

APPENDIX A: EXAMPLES

In this section, we provide some examples about the matching conditions, linearization, and adaptive control of nonlinear systems. A typical nonlinear single-input single-output system is described by,

$$\dot{x} = f_0(x) + \sum_{i=1}^p \theta_i f_i(x) + \left(g_0(x) + \sum_{i=1}^p \theta_i g_i(x) \right) u \quad (\text{A.1})$$

where $f_0(x) \in \mathbb{R}^n$ and $g_0(x) \in \mathbb{R}^n$ are the *nominal* parts of the system, θ_i 's are the unknown parameters, and p is the number of parameters.

A.1. Simple Pendulum ($n = 2$, $\gamma = 2$)

Consider the pendulum system described in Example 5.4.

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -\frac{g}{l} \sin x_1 - \frac{k}{m} x_2 + \frac{1}{ml^2} T \end{aligned} \quad (\text{A.2})$$

The order of this system is $n = 2$. We can write (A.2) as,

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -\theta_1 \sin x_1 - \theta_2 x_2 + \theta_3 u \end{aligned} \quad (\text{A.3})$$

This system is in the form of,

$$\dot{x} = f_0(x) + \sum_{i=1}^3 \theta_i f_i(x) + g_0(x)u + \left(\sum_{i=1}^3 \theta_i g_i(x) \right) u \quad (\text{A.4})$$

where $g_0 = 0$ in this case. The matching conditions state that $f_i \in \mathcal{G}^0$ and $g_i \in \mathcal{G}^0$ for $i = 1, 2, 3$, where,

$$\mathcal{G}^0 = \text{span}\{g_0\}$$

Writing (A.4) in its explicit form, we obtain,

$$\dot{x} = \begin{bmatrix} x_2 \\ 0 \end{bmatrix} + \theta_1 \begin{bmatrix} 0 \\ -\sin x_1 \end{bmatrix} + \theta_2 \begin{bmatrix} 0 \\ -x_2 \end{bmatrix} + \theta_3 \begin{bmatrix} 0 \\ 1 \end{bmatrix} u \quad (\text{A.5})$$

It is seen that,

$$\begin{bmatrix} 0 \\ -\sin x_1 \end{bmatrix} \notin \text{span} \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}$$

and

$$\begin{bmatrix} 0 \\ -x_2 \end{bmatrix} \notin \text{span} \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}$$

Hence the matching conditions are not satisfied. Another check for the matching conditions could be performed via the diffeomorphism which puts the system into normal form. First we have to find the relative degree of the system. Let us assume that the angle of the pendulum is its output, i.e. $y = x_1$. Taking time derivative of the output, we get,

$$\dot{y} = \dot{x}_1 = x_2 \quad (\text{A.6})$$

Note that the control input u does not appear in (A.6). So we differentiate again and we obtain,

$$\ddot{y} = \dot{x}_2 = -\theta_1 \sin x_1 - \theta_2 x_2 + \theta_3 u \quad (\text{A.7})$$

in which u appears. Therefore the relative degree of the system is two, i.e. $\gamma = 2$. No internal dynamics is present since $\gamma = n$. For (A.3), the linearizing diffeomorphism is,

$$\phi(x) = \xi = \begin{bmatrix} h(x) \\ L_f h(x) \end{bmatrix} \quad (\text{A.8})$$

$$\phi(x) = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad (\text{A.9})$$

The inverse diffeomorphism is,

$$\phi^{-1}(\xi) = x = \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} \quad (\text{A.10})$$

Note that the diffeomorphism in this case simply changes the state variables as $\xi = x$, since the system in (A.3) is already in the normal form. Using (A.9), we obtain,

$$\begin{aligned} \dot{\xi}_1 &= \xi_2 \\ \dot{\xi}_2 &= -\theta_1 \sin \xi_1 - \theta_2 \xi_2 + \theta_3 u \\ y &= \xi_1 \end{aligned} \quad (\text{A.11})$$

in which we see that the control input u is multiplied by some uncertain term. In other words, $L_g L_f h \neq 0$. We can check this as,

$$L_f h = x_1 \begin{bmatrix} \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} \end{bmatrix} \begin{bmatrix} x_2 \\ -\theta_1 \sin x_1 - \theta_2 x_2 \end{bmatrix} = x_2 \quad (\text{A.12})$$

$$L_g L_f h = x_2 \begin{bmatrix} \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} \end{bmatrix} \begin{bmatrix} 0 \\ \theta_3 \end{bmatrix} = \theta_3 \quad (\text{A.13})$$

Let us now assume that θ_3 is known as $\theta_3 = 1$ for simplicity. The knowledge of θ_3 corresponds to the knowledge of mass and length of the pendulum. The system model can now be written as,

$$\dot{x} = \begin{bmatrix} x_2 \\ 0 \end{bmatrix} + \theta_1 \begin{bmatrix} 0 \\ -\sin x_1 \end{bmatrix} + \theta_2 \begin{bmatrix} 0 \\ -x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u \quad (\text{A.14})$$

leading to,

$$\begin{bmatrix} 0 \\ -\sin x_1 \end{bmatrix} \in \text{span} \left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\} \quad \text{and} \quad \begin{bmatrix} 0 \\ -x_2 \end{bmatrix} \in \text{span} \left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$$

Now the matching conditions are satisfied, and the uncertainties can be cancelled by appropriate control action. Let us now assume that the angle of the pendulum is supposed to follow a desired trajectory. Applying the linearizing control law to (A.14), we have,

$$u = \frac{1}{L_g L_f^{r-1} h(x)} (-L_f^r h(x) + v) \quad (\text{A.15})$$

The linearized closed loop system can be written as,

$$\begin{aligned} \dot{\xi}_1 &= \xi_2 \\ \dot{\xi}_2 &= v \\ y &= \xi_1 \end{aligned} \quad (\text{A.16})$$

where v is the transformed linear input. However, (A.15) and (A.9) cannot be calculated exactly due to the uncertainties in them. Therefore we use the *certainty equivalence principle* to write the control law in (A.15) as,

$$u = \frac{1}{L_{\hat{g}} L_{\hat{f}}^{r-1} h(x)} (-L_{\hat{f}}^r h(x) + v) \quad (\text{A.17})$$

which is,

$$u = \hat{\theta}_1 \sin x_1 + \hat{\theta}_2 x_2 + v \quad (\text{A.18})$$

Substituting (A.18) in (A.14), the linearized system is obtained as,

$$\begin{aligned}\dot{\xi}_1 &= \xi_2 \\ \dot{\xi}_2 &= v + w\tilde{\theta} \\ y &= \xi_1\end{aligned}\tag{A.19}$$

where $\tilde{\theta} = \theta - \hat{\theta}$ is the parameter error vector and $w = \begin{bmatrix} -\sin x_1 & -x_2 \end{bmatrix}$ is the regressor. The nominal tracking control input v is designed for asymptotic tracking as,

$$v = \ddot{y}_d - \alpha_1(\xi_1 - y_d) - \alpha_2(\xi_2 - \dot{y}_d)\tag{A.20}$$

where y_d is the desired output. Defining the tracking error as $e = y - y_d$, we obtain the error dynamics as,

$$\dot{e} = \begin{bmatrix} 0 & 1 \\ -\alpha_1 & -\alpha_2 \end{bmatrix} e + W(x, \hat{\theta}, u)\tilde{\theta}\tag{A.21}$$

where

$$W(x, \hat{\theta}, u) = \begin{bmatrix} 0 & 0 \\ -\sin x_1 & -x_2 \end{bmatrix}\tag{A.22}$$

Using this regressor, we can determine a parameter update rule as,

$$\dot{\hat{\theta}} = -\dot{\tilde{\theta}} = \Gamma W^T(x, \hat{\theta}, u)Pe\tag{A.23}$$

where $\Gamma > 0$ is the adaptation gain matrix, and P and Q are positive-definite matrices satisfying the Lyapunov equation,

$$A^T P + P A = -Q\tag{A.24}$$

A.2. Simple Pendulum ($n = 2, \gamma = 1$)

Let us assume that, angular velocity of the pendulum is the system output. That is,

$$\begin{aligned}\dot{x}_1 &= x_2 \\ \dot{x}_2 &= -\theta_1 \sin x_1 - \theta_2 x_2 + u \\ y &= x_2\end{aligned}\tag{A.25}$$

Taking time derivative of the output, we obtain,

$$\dot{y} = \dot{x}_2 = -\theta_1 \sin x_1 - \theta_2 x_2 + u\tag{A.26}$$

in which the control input u does appear. Hence the relative degree of the system is unity, i.e. $\gamma = 1$. In this case, the input-output linearizing control law is given as,

$$u = \frac{1}{L_g h(x)} (-L_f h(x) + v)\tag{A.27}$$

This leads to,

$$u = \hat{\theta}_1 \sin x_1 + \hat{\theta}_2 x_2 + v\tag{A.28}$$

We see that the linearizing control law is the same as $\gamma = 2$ case. However, the error dynamics is simpler, as,

$$\dot{e} = -\alpha e + \begin{bmatrix} -\sin x_1 & -x_2 \end{bmatrix} \begin{bmatrix} \tilde{\theta}_1 \\ \tilde{\theta}_2 \end{bmatrix}\tag{A.29}$$

for some $\alpha > 0$.

A.3. Fictitious System ($n = 2, \gamma = 2$)

Consider the following nonlinear system,

$$\begin{aligned}\dot{x}_1 &= \theta_1 x_2^2 \\ \dot{x}_2 &= \theta_2 u \\ y &= x_1\end{aligned}\tag{A.30}$$

Due to the term “ x_2^2 ”, this system is not Lipschitz in the whole $\mathbb{R}^2$. Therefore we suppose that the system is defined in a bounded subset $\mathbb{X} \subset \mathbb{R}^2$ such that it satisfies Lipschitz continuity condition defined in Definition 2.6 at each point of $\mathbb{X}$. We would like to put this system in normal form. To do this, we first find the relative degree of the system taking time derivative of the output as,

$$\dot{y} = \dot{x}_1 = \theta_1 x_2^2\tag{A.31}$$

Since the input u does not appear in the derivative, we differentiate the output once more and we get,

$$\ddot{y} = 2\theta_1 x_2 \dot{x}_2 = 2\theta_1 \theta_2 x_2 u\tag{A.32}$$

The relative degree of (A.30) is two. Using the diffeomorphism,

$$\phi(x) = \xi = \begin{bmatrix} h(x) \\ L_f h(x) \end{bmatrix}\tag{A.33}$$

The normal form equations become,

$$\begin{aligned}\dot{\xi}_1 &= \xi_2 \\ \dot{\xi}_2 &= L_f^2 h + L_g L_f h \cdot u \\ y &= \xi_1\end{aligned}\tag{A.34}$$

We compute the Lie derivatives as follows,

$$L_f h = x_1 \left[\frac{\partial}{\partial x_1} \quad \frac{\partial}{\partial x_2} \right] \begin{bmatrix} \theta_1 x_2^2 \\ 0 \end{bmatrix} = \theta_1 x_2^2 \quad (\text{A.35})$$

$$L_f^2 h = \theta_1 x_2^2 \left[\frac{\partial}{\partial x_1} \quad \frac{\partial}{\partial x_2} \right] \begin{bmatrix} \theta_1 x_2^2 \\ 0 \end{bmatrix} = 0 \quad (\text{A.36})$$

$$L_g L_f h = \theta_1 x_2^2 \left[\frac{\partial}{\partial x_1} \quad \frac{\partial}{\partial x_2} \right] \theta_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 2\theta_1 \theta_2 x_2 \quad (\text{A.37})$$

The diffeomorphism is obtained as,

$$\phi(x) = \xi = \begin{bmatrix} x_1 \\ \theta_1 x_2^2 \end{bmatrix} \quad (\text{A.38})$$

and the inverse-diffeomorphism is given by,

$$\phi^{-1}(\xi) = x = \begin{bmatrix} \xi_1 \\ \sqrt{\frac{\xi_2}{\theta_1}} \end{bmatrix} \quad (\text{A.39})$$

Now we can write (A.30) in the normal form as,

$$\begin{aligned} \dot{\xi}_1 &= \xi_2 \\ \dot{\xi}_2 &= 2\theta_1 \theta_2 \sqrt{\frac{\xi_2}{\theta_1}} \cdot u \\ y &= \xi_1 \end{aligned} \quad (\text{A.40})$$

Inspecting (A.40), we see that matching conditions are not satisfied, due to the uncertainty in the factor of u . Therefore we apply the methodology described in Section 5.5.

We first write (A.30) as,

$$\dot{x} = \theta_1 \begin{bmatrix} x_2^2 \\ 0 \end{bmatrix} + \theta_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} u \quad (\text{A.41})$$

Then we rewrite (A.32) in general form as,

$$\ddot{y} = L_f^2 h + L_g L_f h \cdot u \quad (\text{A.42})$$

We define the tracking control input “ v ” as,

$$v = \ddot{y}_d - \alpha_1(\dot{y} - \dot{y}_d) - \alpha_2(y - y_d) \quad (\text{A.43})$$

where y_d is desired output. The difficulty here is that $\dot{y} = L_f h$ is not measurable, since $L_f h = \theta_1 x_2^2$ and θ_1 is an unknown parameter. The control law,

$$u = \frac{1}{L_g L_f h} (-L_f^2 h + v) \quad (\text{A.44})$$

cannot be calculated either, due to the unknown parameters in it. Therefore we apply the certainty equivalence control law,

$$\hat{u} = \frac{1}{L_{\hat{g}} L_{\hat{f}} h} (-L_{\hat{f}}^2 h + \hat{v}) \quad (\text{A.45})$$

where

$$L_{\hat{f}}^2 h = 0 \quad (\text{A.46})$$

$$L_{\hat{g}} L_{\hat{f}} h = 2\hat{\theta}_1 \hat{\theta}_2 x_2 \quad (\text{A.47})$$

and

$$\hat{v} = \ddot{y}_d - \alpha_1(\hat{\theta}_1 x_2^2 - \dot{y}_d) - \alpha_2(x_1 - y_d) \quad (\text{A.48})$$

Substituting (A.45) in (A.42) (using the certainty equivalence), we get

$$\ddot{y} = L_f^2 h + L_g L_f h \cdot \frac{1}{L_{\hat{g}} L_{\hat{f}} h} \left(-L_{\hat{f}}^2 h + \hat{v} \right) \quad (\text{A.49})$$

and subtracting v from both sides of (A.49), we obtain,

$$\ddot{y} - \ddot{y}_d + \alpha_1(\dot{y} - \dot{y}_d) + \alpha_2(y - y_d) = L_f^2 h + L_g L_f h \frac{1}{L_{\hat{g}} L_{\hat{f}} h} \left(-L_{\hat{f}}^2 h + \hat{v} \right) - v \quad (\text{A.50})$$

which leads to the error dynamics given by,

$$\ddot{e} + \alpha_1 \dot{e} + \alpha_2 e = L_f^2 h + L_g L_f h \frac{1}{L_{\hat{g}} L_{\hat{f}} h} \left(-L_{\hat{f}}^2 h + \hat{v} \right) - v \quad (\text{A.51})$$

Substituting the computed terms into (A.51), we obtain,

$$\ddot{e} + \alpha_1 \dot{e} + \alpha_2 e = \frac{\theta_1 \theta_2}{\hat{\theta}_1 \hat{\theta}_2} \hat{v} - v \quad (\text{A.52})$$

Comparing (A.43) and (A.48), we can write,

$$v = \hat{v} - \alpha_1(\theta_1 x_2^2 - \hat{\theta}_1 x_2^2) \quad (\text{A.53})$$

Substituting (A.53) in (A.52), we get,

$$\ddot{e} + \alpha_1 \dot{e} + \alpha_2 e = \frac{\theta_1 \theta_2}{\hat{\theta}_1 \hat{\theta}_2} \hat{v} - \hat{v} + \alpha_1(\theta_1 x_2^2 - \hat{\theta}_1 x_2^2) \quad (\text{A.54})$$

which can be written in regressor form as,

$$\ddot{e} + \alpha_1 \dot{e} + \alpha_2 e = \underbrace{\begin{bmatrix} \theta_1 - \hat{\theta}_1 & \theta_1 \theta_2 - \hat{\theta}_1 \hat{\theta}_2 \end{bmatrix}}_{\tilde{\Theta}^T} \underbrace{\begin{bmatrix} \alpha_1 x_2^2 \\ \frac{\hat{v}}{\hat{\theta}_1 \hat{\theta}_2} \end{bmatrix}}_W \quad (\text{A.55})$$

$$\ddot{e} + \alpha_1 \dot{e} + \alpha_2 e = \tilde{\Theta}^T W \quad (\text{A.56})$$

We define,

$$M(s) = \frac{1}{s^2 + \alpha_1 s + \alpha_2} \quad (\text{A.57})$$

so that we can write (A.56) as,

$$e = M(s) \cdot \tilde{\Theta}^T W \quad (\text{A.58})$$

We now define the augmented error as,

$$\varepsilon = e + \left(\hat{\Theta}^T M(s) W - M(s) \hat{\Theta}^T W \right) \quad (\text{A.59})$$

In (A.59), all terms are available. Denoting $\xi = M(s)W$, a normalized parameter update law is obtained as,

$$\dot{\hat{\Theta}} = \frac{-\varepsilon \xi}{1 + \xi^T \xi} \quad (\text{A.60})$$

following the discussions given in Chapter 5.

A.4. Motor Controlled Nonlinear Valve ($n = 3, \gamma = 3$)

The system dynamics of the motor controlled nonlinear valve which is shown in Figure A.1 is given as,

$$\begin{aligned}
 \dot{x}_1 &= x_2 \\
 \dot{x}_2 &= -2x_2 - x_1 + Ax_3^2 \\
 \dot{x}_3 &= -Bx_1 + r \\
 y &= x_1
 \end{aligned} \tag{A.61}$$

where A and B are unknown parameters, and r is the control input.

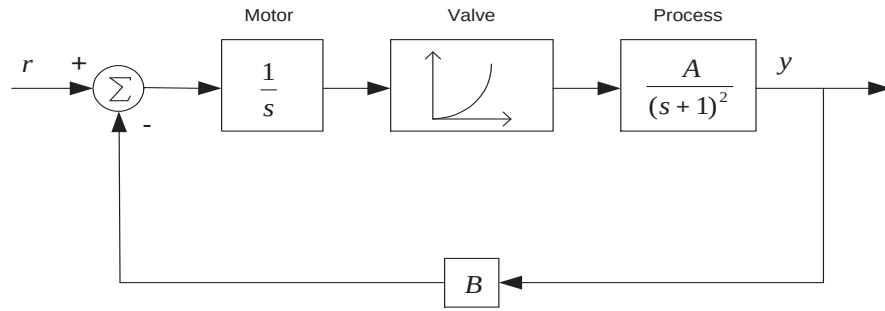


Figure A.1. Block diagram of the nonlinear valve system

To check the matching conditions, the system dynamics can be written as,

$$\dot{x} = \begin{bmatrix} x_2 \\ -2x_2 - x_1 \\ 0 \end{bmatrix} + A \begin{bmatrix} 0 \\ x_3^2 \\ 0 \end{bmatrix} + B \begin{bmatrix} 0 \\ 0 \\ -x_1 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u \tag{A.62}$$

where $u = r$. We see that,

$$\begin{bmatrix} 0 \\ 0 \\ -x_1 \end{bmatrix} \in \text{span} \left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

but,

$$\begin{bmatrix} 0 \\ x_3^2 \\ 0 \end{bmatrix} \notin \text{span} \left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Therefore, the matching conditions are not satisfied and the uncertainty due to the unknown parameter A cannot be compensated by the control u .

A.4.1. Case 1

Let us first assume that A is known and $A = 1$. Taking the time derivative of the output, we get,

$$\dot{y} = \dot{x}_1 = x_2 \quad (\text{A.63})$$

The control input u does not appear in (A.63). Differentiating again, we obtain,

$$\ddot{y} = \dot{x}_2 = -2x_2 - x_1 + x_3^2 \quad (\text{A.64})$$

The input u does not yet appear in (A.64). Continuing differentiation, we get,

$$y^{(3)} = -2\dot{x}_2 - \dot{x}_1 + 2x_3(-Bx_1 + u) \quad (\text{A.65})$$

It is seen that the relative degree of the system is three ($\gamma = 3$). The system dynamics is third order ($n = 3$), thus $\gamma = n$ and linearization yields no internal dynamics. In other words, the system in (A.61) is full-state linearizable. The linearizing diffeomorphism is given by,

$$\phi(x) = \xi = \begin{bmatrix} h(x) \\ L_f h(x) \\ L_f^2 h(x) \end{bmatrix} \quad (\text{A.66})$$

$$\phi(x) = \begin{bmatrix} x_1 \\ x_2 \\ -2x_2 - x_1 + x_3^2 \end{bmatrix} \quad (\text{A.67})$$

and the inverse diffeomorphism is,

$$\phi^{-1}(\xi) = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \sqrt{\xi_1 + 2\xi_2 + \xi_3} \end{bmatrix} \quad (\text{A.68})$$

Using (A.67), the nonlinear system in (A.62) (with $A = 1$) will be expressed in the normal form. We compute,

$$L_f h = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_2 \\ -2x_2 - x_1 + x_3^2 \\ -Bx_1 \end{bmatrix} = x_2 \quad (\text{A.69})$$

$$L_f^2 h = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_2 \\ -2x_2 - x_1 + x_3^2 \\ -Bx_1 \end{bmatrix} = -2x_2 - x_1 + x_3^2 \quad (\text{A.70})$$

$$L_f^3 h = \begin{bmatrix} -1 & -2 & 2x_3 \end{bmatrix} \begin{bmatrix} x_2 \\ -2x_2 - x_1 + x_3^2 \\ -Bx_1 \end{bmatrix} = 3x_2 + 2x_1 - 2x_3^2 - 2Bx_1x_3 \quad (\text{A.71})$$

and

$$L_g L_f^2 h = \begin{bmatrix} -1 & -2 & 2x_3 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = 2x_3 \quad (\text{A.72})$$

The normal form is obtained as,

$$\begin{aligned}
 \dot{\xi}_1 &= \xi_2 \\
 \dot{\xi}_2 &= \xi_3 \\
 \dot{\xi}_3 &= -\xi_2 - 2\xi_3 - 2B\xi_1\sqrt{\xi_1 + 2\xi_2 + \xi_3} + 2\sqrt{\xi_1 + 2\xi_2 + \xi_3}u \\
 y &= \xi_1
 \end{aligned} \tag{A.73}$$

The linearizing control law is,

$$u = \frac{1}{2\sqrt{\xi_1 + 2\xi_2 + \xi_3}} \left(\xi_2 + 2\xi_3 + 2B\xi_1\sqrt{\xi_1 + 2\xi_2 + \xi_3} + v \right) \tag{A.74}$$

However, the parameter B is unknown in (A.74). Therefore we apply the certainty equivalence control law using $\hat{B}$ instead of B in (A.74) as,

$$u = \frac{1}{2\sqrt{\xi_1 + 2\xi_2 + \xi_3}} \left(\xi_2 + 2\xi_3 + 2\hat{B}\xi_1\sqrt{\xi_1 + 2\xi_2 + \xi_3} + v \right) \tag{A.75}$$

When (A.75) is substituted in (A.73), it leads to,

$$\begin{aligned}
 \dot{\xi}_1 &= \xi_2 \\
 \dot{\xi}_2 &= \xi_3 \\
 \dot{\xi}_3 &= v + w\tilde{p} \\
 y &= \xi_1
 \end{aligned} \tag{A.76}$$

where $\tilde{p} = B - \hat{B}$ is the parameter error and $w = -2\xi_1\sqrt{\xi_1 + 2\xi_2 + \xi_3}$ is the regressor. Then we design v for asymptotic tracking as,

$$v = y_d^{(3)} - \alpha_1(\xi_1 - y_d) - \alpha_2(\xi_2 - \dot{y}_d) - \alpha_3(\xi_3 - \ddot{y}_d) \tag{A.77}$$

where y_d is the desired output. Defining the error as $e = y - y_d$, we obtain the error dynamics as,

$$\dot{e} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -\alpha_1 & -\alpha_2 & -\alpha_3 \end{bmatrix} e + W(\xi, \hat{\theta}, u) \tilde{\theta} \quad (\text{A.78})$$

where

$$W(\xi, \hat{\theta}, u) = \begin{bmatrix} 0 \\ 0 \\ -2\xi_1 \sqrt{\xi_1 + 2\xi_2 + \xi_3} \end{bmatrix} \quad (\text{A.79})$$

Using this regressor, we can determine a parameter update rule as,

$$\dot{\hat{\theta}} = -\dot{\tilde{\theta}} = \Gamma W^T(x, \hat{\theta}, u) P e \quad (\text{A.80})$$

where $\Gamma > 0$ is the adaptation gain matrix, and P and Q are positive-definite matrices satisfying the Lyapunov equation,

$$A^T P + P A = -Q \quad (\text{A.81})$$

Note that (A.67), (A.75) and (A.77) can be computed using only measured states and the parameter estimate, because the matching conditions are satisfied in this example.

A.4.2. Case 2

Let us now assume the parameter A is unknown. The matching conditions are not satisfied in this case. There are two options to follow for this case.

1. Design via the so called *extended* matching conditions,
2. Design via the direct adaptive control methodology described in Section 5.5

A.4.2.1. The Extended Matching Conditions. We first check the extended matching conditions. The extended matching conditions state that $f_i \in \mathcal{G}^1$ and $g_i \in \mathcal{G}^0$ for $i = 1, 2, 3$, where,

$$\mathcal{G}^0 = \text{span}\{g_0\}$$

and

$$\mathcal{G}^1 = \text{span}\{g_0, \text{ad}_{f_0}g_0\}$$

$$\text{ad}_{f_0}g_0 = \frac{\partial g_0}{\partial x}f_0 - \frac{\partial f_0}{\partial x}g_0 \quad (\text{A.82})$$

$$= \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3} \end{bmatrix} \begin{bmatrix} x_2 \\ -2x_2 - x_1 \\ 0 \end{bmatrix} - \begin{bmatrix} x_2 \\ -2x_2 - x_1 \\ 0 \end{bmatrix} \begin{bmatrix} \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad (\text{A.83})$$

$$= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_2 \\ -2x_2 - x_1 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 & 1 & 0 \\ -1 & -2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad (\text{A.84})$$

Hence, we get,

$$\text{ad}_{f_0}g_0 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad (\text{A.85})$$

We observe that,

$$\begin{bmatrix} 0 \\ x_3^2 \\ 0 \end{bmatrix} \notin \text{span} \left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right\}$$

That is, the extended matching conditions are not satisfied either.

A.4.2.2. The Direct Adaptive Control Method. In this method, Lipschitz continuity of all nonlinearities are required. Therefore we suppose the system is defined in a domain $\mathbb{X} \subset \mathbb{R}^3$ such that the Lipschitz continuity of the term “ x_3^2 ” is guaranteed. We rewrite the system dynamics as,

$$\dot{x} = 1 \cdot \begin{bmatrix} x_2 \\ -2x_2 - x_1 \\ 0 \end{bmatrix} + A \begin{bmatrix} 0 \\ x_3^2 \\ 0 \end{bmatrix} + B \begin{bmatrix} 0 \\ 0 \\ -x_1 \end{bmatrix} + 1 \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u \quad (\text{A.86})$$

which is the form of (A.1). There is a slight difference in the definition/notation. In this method, the nominal parts in (A.1) are not used, but instead, they are included as factors of a known parameter. Therefore, the parameter vector is $\theta = \begin{bmatrix} 1 \\ A \\ B \end{bmatrix}$. Since the relative degree is three, the third derivative of the output is,

$$y^{(3)} = L_f^3 h + L_g L_f^2 h \cdot u \quad (\text{A.87})$$

and the control law is given as,

$$\hat{u} = \frac{1}{L_{\hat{g}} L_{\hat{f}}^2 h} \left(-L_{\hat{f}}^3 h + \hat{v} \right) \quad (\text{A.88})$$

Using the certainty equivalence, we substitute (A.88) in (A.87) and obtain,

$$y^{(3)} = L_f^3 h + L_g L_f^2 h \frac{1}{L_{\hat{g}} L_f^2 h} \left(-L_{\hat{f}}^3 h + \hat{v} \right) \quad (\text{A.89})$$

where

$$\hat{v} = y_d^{(3)} - \alpha_1(L_{\hat{f}}^2 h - \ddot{y}_d) - \alpha_2(L_{\hat{f}} h - \dot{y}_d) - \alpha_3(y - y_d) \quad (\text{A.90})$$

Subtracting v from both sides of (A.89), we obtain,

$$y^{(3)} - y_d^{(3)} + \alpha_1(L_{\hat{f}}^2 h - \ddot{y}_d) + \alpha_2(L_{\hat{f}} h - \dot{y}_d) + \alpha_3(y - y_d) = L_f^3 h + L_g L_f^2 h \frac{1}{L_{\hat{g}} L_f^2 h} \left(-L_{\hat{f}}^3 h + \hat{v} \right) - v \quad (\text{A.91})$$

$$e^{(3)} + \alpha_1 \ddot{e} + \alpha_2 \dot{e} + \alpha_3 e = L_f^3 h + L_g L_f^2 h \frac{1}{L_{\hat{g}} L_f^2 h} \left(-L_{\hat{f}}^3 h + \hat{v} \right) - v \quad (\text{A.92})$$

We have,

$$L_f h = x_2 \quad (\text{A.93})$$

$$L_f^2 h = -x_1 - 2x_2 + Ax_3^2 \quad (\text{A.94})$$

$$L_f^3 h = -2x_1 + x_2 - 2Ax_3^2 - 2ABx_1x_3 \quad (\text{A.95})$$

$$L_g L_f^2 h = 2Ax_3 \quad (\text{A.96})$$

We have also,

$$v = \hat{v} - \alpha_1(L_f^2 h - L_{\hat{f}}^2 h) - \alpha_2(L_f h - L_{\hat{f}} h) \quad (\text{A.97})$$

Substituting (A.95)-(A.97) into (A.92) and rewriting (A.92) in regressor form, we get,

$$e^{(3)} + \alpha_1 \ddot{e} + \alpha_2 \dot{e} + \alpha_3 e = \tilde{\Theta}^T W \quad (\text{A.98})$$

A.5. Field-Controlled DC Motor ($n = 2, \gamma = 1$)

A state-space model of a field-controlled DC motor is given as,

$$\begin{aligned} \dot{x}_1 &= -\theta_1 x_1 - \theta_2 x_2 u + \theta_3 \\ \dot{x}_2 &= -\theta_4 x_2 + \theta_5 x_1 u \\ y &= x_2 \end{aligned} \quad (\text{A.99})$$

This system can also be written as,

$$\dot{x} = \theta_1 \begin{bmatrix} -x_1 \\ 0 \end{bmatrix} + \theta_3 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \theta_4 \begin{bmatrix} 0 \\ -x_2 \end{bmatrix} + \left(\theta_2 \begin{bmatrix} -x_2 \\ 0 \end{bmatrix} + \theta_5 \begin{bmatrix} 0 \\ x_1 \end{bmatrix} \right) u \quad (\text{A.100})$$

We see that $f_i \notin \text{span}\{g_0\}$ and $g_i \notin \text{span}\{g_0\}$, since $\text{span}\{g_0\} = \emptyset$. Therefore, the strict matching conditions are not satisfied. We apply the direct adaptive control method described in Section 5.5. Taking time derivative of y , we obtain,

$$\dot{y} = \dot{x}_2 = -\theta_4 x_2 + \theta_5 x_1 u \quad (\text{A.101})$$

The relative degree of the system is unity. The linearizing control law is given by,

$$u = \frac{1}{L_{\hat{g}} h} (-L_{\hat{f}} h + v) \quad (\text{A.102})$$

with

$$L_{\hat{f}} h = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} -\hat{\theta}_1 x_1 + \hat{\theta}_3 \\ -\hat{\theta}_4 x_2 \end{bmatrix} = -\hat{\theta}_4 x_2 \quad (\text{A.103})$$

and

$$L_{\hat{g}}h = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} -\hat{\theta}_2 x_2 \\ \hat{\theta}_5 x_1 \end{bmatrix} = \hat{\theta}_5 x_1 \quad (\text{A.104})$$

Computing the regressor vectors $W_1 = \begin{bmatrix} L_{f_1}h \\ L_{f_2}h \\ L_{f_3}h \end{bmatrix}$ and $W_2 = \begin{bmatrix} L_{g_1}h \\ L_{g_2}h \end{bmatrix} \frac{-L_f h + v}{L_{\hat{g}}h}$, we get,

$$W_1 = \begin{bmatrix} 0 \\ 0 \\ -x_2 \end{bmatrix} \quad (\text{A.105})$$

and

$$W_2 = \begin{bmatrix} 0 \\ (\hat{\theta}_4 x_2 + v)/\hat{\theta}_5 \end{bmatrix} \quad (\text{A.106})$$

Then we can write,

$$\dot{y} = v + \tilde{\Theta}^T \begin{bmatrix} W_1 \\ W_2 \end{bmatrix} \quad (\text{A.107})$$

The control law for tracking is given as,

$$v = \dot{y}_d - \alpha(y - y_d) \quad (\text{A.108})$$

Defining the output error $e = y - y_d$, we have,

$$\dot{e} + \alpha e = \tilde{\Theta}^T W \quad (\text{A.109})$$

where W is the concatenation of W_1 and W_2 . The parameter update law is given by,

$$\dot{\hat{\theta}} = -eW \quad (\text{A.110})$$

In our example,

$$\dot{\hat{\theta}} = -e \begin{bmatrix} 0 \\ 0 \\ -x_2 \\ 0 \\ (\hat{\theta}_4 x_2 + v)/\hat{\theta}_5 \end{bmatrix} \quad (\text{A.111})$$

and

$$u = \frac{1}{\hat{\theta}_5 x_1} (\hat{\theta}_4 x_2 + \dot{y}_d - \alpha(x_2 - y_d)) \quad (\text{A.112})$$

where $\alpha > 0$ is a design parameter.

A.6. Field-Controlled DC Motor ($n = 2, \gamma = 1$)

Let us assume that the parameters θ_2 and θ_5 in (A.99) are known and equal to unity. Renaming the parameters, we can write the system model as,

$$\begin{aligned} \dot{x}_1 &= -p_1 x_1 - x_2 u + p_2 \\ \dot{x}_2 &= -p_3 x_2 + x_1 u \\ y &= x_2 \end{aligned} \quad (\text{A.113})$$

which can also be written as,

$$\dot{x} = p_1 \begin{bmatrix} -x_1 \\ 0 \end{bmatrix} + p_2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + p_3 \begin{bmatrix} 0 \\ -x_2 \end{bmatrix} + \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix} u \quad (\text{A.114})$$

We know from Section A.5 that the relative degree of the system is unity. Since the system degree is two and $\text{span}\{g_0\} = \mathbb{R}^2$, we observe that the matching conditions are satisfied for this case. However, we have to check for the stability of the internal dynamics, because $\gamma < n$. It is seen that the system in (A.113) is already in the normal form, and could be expressed as,

$$\begin{aligned}\dot{\xi} &= -p_3\xi + \eta u \\ \dot{\eta} &= -p_1\eta + p_2 - \xi u \\ y &= \xi\end{aligned}\tag{A.115}$$

Setting $\xi = 0$, we get,

$$\dot{\eta} = -p_1\eta + p_2\tag{A.116}$$

and the solution of the zero-dynamics is obtained as

$$\eta(t) = p_2 e^{-p_1 \eta t}\tag{A.117}$$

We see that the zero dynamics is exponentially stable provided $p_1 > 0$ for all time. The control law is given as,

$$u = \frac{1}{\eta}(\hat{p}_3\xi + v)\tag{A.118}$$

Another condition for the tracking control is that $\eta = x_1$ must be away from zero. The tracking control input v is given by,

$$v = \dot{y}_d - \alpha(\xi - y_d)\tag{A.119}$$

Defining $e = y - y_d$ and plugging (A.118) and (A.119) in (A.115), we get,

$$\dot{\xi} = v - (p_3 - \hat{p}_3)\xi\tag{A.120}$$

This leads to the error dynamics given by,

$$\dot{e} = -\alpha e - \xi \tilde{p}_3 \quad (\text{A.121})$$

which suggests a parameter update law as,

$$\dot{\tilde{p}}_3 = \dot{\hat{p}}_3 = -\Gamma \xi e \quad (\text{A.122})$$

where $\Gamma > 0$ is a gain constant.

In this example, we see that we cannot identify the parameters p_1 and p_2 as they affect only the internal dynamics. From control point of view, the positivity of p_1 is vital because it determines stability of the zero dynamics. p_2 has no critical effect on system stability.

REFERENCES

1. Morse, A. S., "Global Stability of Parameter-Adaptive Control Systems", *IEEE Transactions on Automatic Control*, Vol. 25, No. 3, pp. 433-439, 1980.
2. Narendra, K. S., Y. H. Lin and L. S. Valavani, "Stable Adaptive Controller Design, Part II: Proof of Stability", *IEEE Transactions on Automatic Control*, Vol. 25, No. 3, pp. 440-448, 1980.
3. Goodwin, G. C., P. J. Ramadge and P. E. Caines, "Discrete-Time Multivariable Adaptive Control", *IEEE Transactions on Automatic Control*, Vol. 25, No. 3, pp. 449-456, 1980.
4. Narendra, K. S. and A. M. Annaswamy, *Stable Adaptive Systems*, Prentice-Hall, New Jersey, 1989.
5. Astrom, K. J. and B. Wittenmark, *Adaptive Control*, Addison-Wesley, Reading, 1995.
6. Sastry, S. and M. Bodson, *Adaptive Control: Stability, Convergence, and Robustness*, Prentice-Hall, New Jersey, 1989.
7. Nam, K. and A. Arapostathis, "A Model Reference Adaptive Control Scheme for Pure-Feedback Nonlinear Systems", *IEEE Transactions on Automatic Control*, Vol. 33, No. 9, pp. 803-811, 1988.
8. Sastry, S. and A. Isidori, "Adaptive Control of Linearizable Systems", *IEEE Transactions on Automatic Control*, Vol. 34, No. 11, pp. 1123-1131, 1989.
9. Teel, A., R. Kadiyala, P. Kokotovic and S. Sastry, "Indirect Techniques for Adaptive Input-Output Linearization of Non-linear Systems", *Int. Journal of Control*, Vol. 53, No. 1, pp. 193-222, 1991.

10. Campion, G. and G. Bastin, "Indirect Adaptive State Feedback Control of Linearly Parametrized Nonlinear Systems", *Int. Journal of Adaptive Control and Signal Processing*, Vol. 4, pp. 345-358, 1990.
11. Taylor, D. G., P. V. Kokotovic, R. Marino and I. Kanellakopoulos, "Adaptive Regulation of Nonlinear Systems with Unmodeled Dynamics", *IEEE Transactions on Automatic Control*, Vol. 34, No. 4, pp. 405-412, 1989.
12. Marino, R., P. Tomei, I. Kanellakopoulos and P. V. Kokotovic, "Adaptive Tracking for a Class of Feedback Linearizable Systems", *IEEE Transactions on Automatic Control*, Vol. 39, No. 6, pp. 1314-1319, 1994.
13. Kanellakopoulos, I. and P. V. Kokotovic and R. Marino, "Robustness of Adaptive Nonlinear Control under an Extended Matching Condition", *IFAC Symposium on Nonlinear System Design*, Capri, Italy, June 1989, pp. 192-197.
14. Kanellakopoulos, I., P. V. Kokotovic and A. S. Morse, "Systematic Design of Adaptive Controllers for Feedback Linearizable Systems", *IEEE Transactions on Automatic Control*, Vol. 36, No. 11, pp. 1241-1253, 1991.
15. Martennson, B., *Adaptive Stabilization*, Ph.D. Thesis, Lund Institute of Technology, 1986.
16. Fu, M. and B. R. Barmish, "Adaptive Stabilization of Linear Systems via Switching Control", *IEEE Transactions on Automatic Control*, Vol. 31, No. 12, pp. 1097-1103, 1986.
17. Morse, A. S., D. Q. Mayne and G. C. Goodwin, "Applications of Hysteresis Switching in Parameter Adaptive Control", *IEEE Transactions on Automatic Control*, Vol. 37, No. 9, pp. 1343-1354, 1992.
18. Morse, A. S., "Supervisory Control of Families of Linear Set-Point Controllers: Part I", *IEEE Transactions on Automatic Control*, Vol. 41, No. 10, pp. 1413-1431,

1996.

19. Hespanha, J., D. Liberzon and A. S. Morse, "Overcoming the Limitations of Adaptive Control by Means of Logic Based Switching", *Systems and Control Letters*, Vol. 49, pp. 49-65, 2003.
20. Narendra, K. S. and J. Balakrishnan, "Improving Transient Response of Adaptive Control Systems Using Multiple Models and Switching", *IEEE Transactions on Automatic Control*, Vol. 39, No. 9, pp. 1861-1866, 1994.
21. Narendra, K. S. and J. Balakrishnan, "Adaptive Control Using Multiple Models", *IEEE Transactions on Automatic Control*, Vol. 42, No. 2, pp. 171-187, 1997.
22. Autenrieth, T. and E. Rogers, "Performance Enhancements for a Class of Multiple Model Adaptive Schemes", *Int. Journal of Adaptive Control and Signal Processing*, Vol. 13, pp. 105-127, 1999.
23. Narendra, K. S., J. Balakrishnan and M. K. Ciliz, "Adaptation and Learning Using Multiple Models, Switching and Tuning", *IEEE Control Systems Magazine*, Vol. 15, No. 3, pp. 37-51, 1995.
24. Ciliz, K. and K. S. Narendra, "Adaptive Control of Robotic Manipulators Using Multiple Models and Switching", *Int. Journal of Robotics Research*, Vol. 15, pp. 592-610, 1996.
25. Narendra, K. S., O. A. Driollet, M. Feiler and K. George, "Adaptive Control Using Multiple Models, Switching and Tuning", *Int. Journal of Adaptive Control and Signal Processing*, Vol. 17, pp. 87-102, 2003.
26. Ciliz, M. K., "Combined Direct and Indirect Adaptive Control of Robot Manipulators Using Multiple Models", *Advanced Robotics*, Vol. 20, No. 4, pp. 483-497, 2006.
27. Kalkkuhl, J., T. A. Johansen and J. Lüdemann, "Improved Transient Performance

- of Nonlinear Adaptive Backstepping Using Estimator Resetting Based on Multiple Models”, *IEEE Transactions on Automatic Control*, Vol. 47, No. 1, pp. 136-140, 2002.
28. Freidovich, L. B. and H. K. Khalil, “Logic-Based Switching for Robust Control of Minimum-Phase Nonlinear Systems”, *Systems and Control Letters*, Vol. 54, No. 8, pp. 713-727, 2005.
 29. Cezayirli, A. and M. K. Ciliz, “Indirect Adaptive Control of Nonlinear Systems Using Multiple Identification Models and Switching”, 2006, (submitted paper).
 30. Ciliz, M. K. and A. Cezayirli, “Increased Transient Performance for the Adaptive Control of Feedback Linearizable Systems Using Multiple Models”, *Int. Journal of Control*, Vol. 79, No. 10, pp. 1205-1215, 2006.
 31. Cezayirli, A. and M. K. Ciliz, “Transient Performance Enhancement of Direct Adaptive Control of Nonlinear Systems Using Multiple Models and Switching”, 2006, (submitted paper).
 32. Ciliz, M. K. and A. Cezayirli, “Combined Direct and Indirect Adaptive Control of Feedback Linearizable Nonlinear Systems”, 2006, (submitted paper).
 33. Isidori, A., *Nonlinear Control Systems*, Springer-Verlag, London, 1995.
 34. Khalil, H., *Nonlinear Systems*, Prentice-Hall, New Jersey, 2000.
 35. Sastry, S., *Nonlinear Systems Analysis, Stability, and Control*, Springer-Verlag, New York, 1999.
 36. Nijmeijer, H. and A. van der Schaft, *Nonlinear Dynamical Control Systems*, Springer-Verlag, New York, 1990.
 37. Krstic, A., I. Kanellakopoulos and P. Kokotovic, *Nonlinear and Adaptive Control Design*, Wiley, New York, 1995.

38. Ioannou, P. A. and J. Sun, *Robust Adaptive Control*, Prentice-Hall, New Jersey, 1996.
39. Tao, G., *Adaptive Control Design and Analysis*, Wiley, New York, 2003.
40. Ogata, K., *Modern Control Engineering*, Prentice-Hall, New Jersey, 2002.
41. Slotine, J. E. and W. Li, *Applied Nonlinear Control*, Prentice-Hall, New Jersey, 1991.
42. Marino, R. and P. Tomei, *Nonlinear Control Design: Geometric, Adaptive and Robust*, Prentice-Hall, London, 1995.
43. Arslan, G. and T. Basar, "Disturbance Attenuating Controller Design for Strict-Feedback Systems with Structurally Unknown Dynamics", *Automatica*, Vol. 37, No. 8, pp. 1175-1188, 2001.
44. Krstic, M. and P. V. Kokotovic, "Adaptive Nonlinear Output-Feedback Schemes with Marino-Tomei Controller", *IEEE Transactions on Automatic Control*, Vol. 41, No. 2, pp. 274-280, 1996.
45. Khalil, H., "Adaptive Output Feedback Control of Nonlinear Systems Represented by Input-Output Models", *IEEE Transactions on Automatic Control*, Vol. 41, No. 2, pp. 177-188, 1996.
46. Aloliwi, B. and H. Khalil, "Robust Adaptive Output Feedback Control of Nonlinear Systems Without Persistence of Excitation", *Automatica*, Vol. 33, No. 11, pp. 2025-2032, 1997.
47. Seto, D. and A. M. Annaswamy and J. Baillieul, "Adaptive Control of Nonlinear Systems with a Triangular Structure", *IEEE Transactions on Automatic Control*, Vol. 39, No. 7, pp. 1411-1428, 1994.
48. Kanellakopoulos, I., P. V. Kokotovic and R. Marino, "An Extended Direct Scheme

- for Robust Adaptive Nonlinear Control”, *Automatica*, Vol. 27, No. 2, pp. 247-255, 1991.
49. Pomet, J. B. and L. Praly, “Adaptive Nonlinear Regulation: Estimation from the Lyapunov Equation”, *IEEE Transactions on Automatic Control*, Vol. 37, No. 6, pp. 729-740, 1992.
 50. Anderson, B. D. O., T. S. Brinsmead, F. De Bruyne, J. Hespanha, D. Liberzon and A. S. Morse, “Multiple Model Adaptive Control. Part I: Finite Controller Coverings”, *Int. Journal of Robust and Nonlinear Control*, Vol. 10, pp. 909-929, 2000.
 51. Hespanha, J., D. Liberzon, A. S. Morse, B. D. O. Anderson, T. S. Brinsmead and F. De Bruyne, “Multiple Model Adaptive Control. Part II: switching”, *Int. Journal of Robust and Nonlinear Control*, Vol. 11, pp. 479-496, 2001.
 52. Narendra, K. S. and K. George, “Adaptive Control of Simple Nonlinear Systems Using Multiple Models”, *Proc. of the American Control Conference*, Anchorage, USA, May 2002, pp. 1779-1784.
 53. Cezayirli, A. and K. Ciliz, “Multiple Model Based Adaptive Control of a DC Motor under Load Changes”, *Proc. IEEE International Conference on Mechatronics*, Istanbul, Turkey, June 2004, pp. 328-333.
 54. Branicky, M. S., “Multiple Lyapunov Functions and Other Analysis Tools for Switched and Hybrid Systems”, *IEEE Transactions on Automatic Control*, Vol. 43, No. 4, pp. 475-482, 1998.
 55. Cezayirli, A. and K. Ciliz, “Adaptive Tracking for Nonlinear Plants Using Multiple Identification Models and State-Feedback”, *The 32nd IEEE Annual Conference on Industrial Electronics - IECON'06*, Paris, France, November 2006.
 56. Duarte, M. A. and K. S. Narendra, “Combined Direct and Indirect Approach to

- Adaptive Control”, *IEEE Transactions on Automatic Control*, Vol. 34, No. 10, pp. 1071-1075, 1989.
57. Yu, H., L. D. Seneviratne and S. W. E. Earles, “Robust Adaptive Control for Robot Manipulators Using a Combined Method”, *Proc. of IEEE Int. Conf. on Robotics and Automation*, Atlanta, USA, May 1993, pp. 612-617.
58. Yu, H. and S. Lloyd, “Combined Direct and Indirect Adaptive Control of Constrained Robots”, *Int. Journal of Control*, Vol. 68, No. 5, pp. 955-970, 1997.
59. Slotine, J. and W. Li, “Composite Adaptive Control of Robot Manipulators”, *Automatica*, Vol. 25, No. 4, pp. 509-519, 1989.