# Adaptive Control of a First-Order System Providing Linear-Like Behaviour and Asymptotic Tracking

by

Hassaan Ali Qazi

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2021

## Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

## Abstract

Adaptive control is an approach used to deal with systems having uncertain and/or time-varying parameters. In this thesis, we consider the problem of designing an adaptive controller for a discrete-time first-order plant. Recently, Shahab et.al. considered this problem and proposed an approach which provides linear-like behaviour: exponential stability and a convolution bound on the input-output behaviour, together with robustness to slow time-variations and unmodelled dynamics. However, asymptotic tracking of a general reference signal was not provided.

Here, we extend the aforementioned work with the aim to achieve asymptotic tracking while retaining linear-like closed-loop behaviour. We replace this uncertainty set with a pair of convex sets, one for each sign of the input gain, which enables us to use two parameter estimators – one for each convex set. We design these estimators using the modified version of the original projection algorithm. For each estimator, there is the corresponding one-step-ahead control law. A dynamic performance signal based switching rule is then adopted that decides which controller should be used at each time step. It is shown that the proposed approach preserves linear-like behaviour. In addition to that, we also have shown asymptotic trajectory tracking for two different circumstances: when the reference signal is asymptotically strongly persistently exciting of order two, and for a fairly general reference signal but the plant is unstable. Numerical simulations are presented to demonstrate the efficacy of the proposed approach.

**Acknowledgements**

I would like to thank my supervisors, Professor Daniel E. Miller and Professor Christopher Nielsen for providing me an opportunity to work on this research problem. I am truly grateful for their patience and guidance.

I would also like to thank Professor Andrew Heunis and Professor Stephen L. Smith for giving their valuable feedback and comments.

Furthermore, I would like to acknowledge the support of the Professors, whom I have taken courses with, for taking their time out to share their knowledge. I would also like to thank all my friends especially *Rollen D'Souza* for providing me with his unreserved and wholehearted support.

Finally, I am deeply grateful to my parents for their unconditional love and appreciation.

## Dedication

This is dedicated to my family.

# Table of Contents


**List of Figures** **viii**

**List of Tables** **ix**

**1 Introduction** **1**

1.1 Background and Motivation . . . . . 1

1.2 The Objective . . . . . 3

1.3 Notation . . . . . 4

**2 The Setup** **5**

2.1 The Plant . . . . . 5

2.2 Parameter Estimation . . . . . 7

2.3 Revised Estimation Algorithm . . . . . 8

2.4 Switching Controller . . . . . 12

**3 Stability** **17**

3.1 Closed-Loop Stability . . . . . 17

3.2 Tolerance to Time-Variations and Unmodelled Dynamics . . . . . 33

**4 Tracking Problem** **40**

4.1 Preliminary Results . . . . . 40

4.2 Tracking With a Persistence of Excitation Assumption . . . . . 45

4.3 Tracking for a Fairly General Reference Signal . . . . . 53

**5 Numerical Simulations** **63**

5.1 Simulation Parameters . . . . . 63

5.2 Scenario 1: Tracking . . . . . 64

5.3 Scenario 2: Noise Gain . . . . . 65

5.4 Representative Examples . . . . . 66

**6 Summary and Future Work** **72**

6.1 Summary of the Results . . . . . 72

6.2 Future Work . . . . . 73

**References** **74**

# List of Figures


2.1 The Block diagram of the closed-loop system; enclosed inside the dashed boxes are the multiple estimators/controllers (blue), and the switching mechanism (red). . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 15

5.1 Plots for continuous parameter, time-varying reference and no noise. Column (a) shows our proposed algorithm whereas column (b) shows the one presented in [34]. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 69

5.2 Plots for discontinuous parameter, time-varying reference and no noise. Column (a) shows our proposed algorithm whereas column (b) shows the one presented in [34]. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 69

5.3 Plots for time-varying reference and no noise with $\bar{\lambda} = 0.3$. Column (a) shows the continuous parameter case whereas column (b) shows the discontinuous parameter case. . . . . . . . . . . . . . . . . . . . . . . . . . . . 70

5.4 Plot for continuous parameter, time-varying noise and no reference. Column (a) shows our proposed algorithm whereas column (b) shows the one presented in [34]. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 70

5.5 Plots for discontinuous parameter, time-varying noise and no reference. Column (a) shows our proposed algorithm whereas column (b) shows the one presented in [34]. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 71

5.6 Plots for time-varying noise and no reference with $\bar{\lambda} = 0.3$. Column (a) shows the continuous parameter case whereas column (b) shows the discontinuous parameter case. . . . . . . . . . . . . . . . . . . . . . . . . . . . 71

# List of Tables

5.1 Mean and range of tracking gain $W_1(\omega)$ when $n = 0$ and $r$ is time-varying . 65

5.2 Number of switches when $n = 0$ and $r$ is time-varying . . . . . . . . . . . . 66

5.3 Mean and range of noise gain $W_2(\omega)$ when $n$ is time-varying and $r = 0$ . . 67

5.4 Number of switches when $n$ is time-varying and $r = 0$ . . . . . . . . . . . . 68

# Chapter 1

# Introduction

We begin this chapter by discussing the background of the control problem that we consider and the motivation behind our proposed approach. Then we give the overview of our setup and the control law adopted, followed by an outline of this thesis. In the last section, we define some notation that is used throughout the thesis.

## 1.1 Background and Motivation

The notion of adaptive control is for the controller to deal with systems having uncertain and/or time-varying parameters. The classical adaptive controller combines a linear time-invariant (LTI) compensator together with a tuning algorithm that adjusts its parameters. The general proofs that parameter adaptive controller could work first came around in 1980, e.g., see [5, 23, 7, 30, 31]. Such controllers, however, are usually not robust to unmodelled dynamics and time-variations, and do not handle disturbances well – see [32]. Furthermore, they place strict assumptions on *a priori* information about the plant structure.

Due to these shortcomings, in the following two decades, a great deal of effort was made to come up with design modifications; these include the use of deadzones, $\sigma$-modification and signal normalization, e.g., see [16, 17, 38, 13, 11]. The approach that turned out quite powerful is that of using projection onto a convex set of parameters, which provided desirable properties like bounded-noise bounded-state and tolerance to unmodelled dynamics and/or slow time-variations – see [41, 42, 27, 40, 39, 15]. However, even in these cases, typically neither a bounded gain on the noise nor exponential stability is proven. This is in contrast to the desirable properties which arise in LTI controller design for an LTI plant.

In addition to above mentioned classical approaches to adaptive control, non-classical approaches were also developed. In [6, 20], a logic-based switching approach is used to find out the most suitable controller among a pre-defined list of candidates. The drawback of this approach is that the transient behaviour can be quite poor with no bounded gain on the noise. A more sophisticated logic-based switching algorithm, labelled Supervisory Control, emerged, which in certain circumstances provided a bounded gain on the noise – see [24, 25, 9, 10, 26]. The use of multiple parameter estimators also surfaced around the same time [28, 29], which argued that the proposed approach improved the transient behaviour. Also, an approach labelled Adaptive Mixing Control was presented in [2, 3, 14], where tolerance to noise and unmodelled dynamics was achieved by enforcing convexity assumptions. Unfortunately, none of these schemes were able to show linear-like behaviour on the closed-loop system, i.e., exponential stability and convolution bound on the exogenous signals.

In recent papers, such as [19, 22, 21], linear-like closed-loop behaviour along with robustness to unmodelled dynamics and parameter variations is shown; the likes of these results were not seen before. The first-order system is considered in [19] and it is assumed that the plant parameters lie in a compact and convex set. The original projection algorithm is used, to estimate the parameters, in conjunction with the one-step-ahead adaptive controller. Asymptotic tracking, for the noiseless case, and a bound on the average size of the tracking error, for the noisy case, is shown. A $d$-step ahead adaptive control law is proposed in [22] for a minimum phase plant where the set of admissible plant parameters need not be convex. The parameters are estimated using a modified version of the ideal projection algorithm termed a "vigilant estimator". The authors established a bound on the tracking error for both the noisy and noiseless cases. In [21], with convexity assumption on the plant parameters set being slightly weakened, the classical pole-placement adaptive controller is employed to prove stability for non-minimum phase system.

Multiple parameter estimators approaches, yielding the desirable linear-like results, are discussed in [34] and [37] with a relaxed convexity requirement. In these papers, the plant parameters are instead assumed to be in a compact uncertainty set, which is then replaced by multiple convex sets. The key idea presented in [34] – dealing with the first-order case – is to design an estimator and the one-step-ahead controller for each of these convex sets. A simple switching algorithm is presented which captures the prediction errors for each estimator and chooses the estimator index, and the corresponding controller, whose prediction error has the smallest absolute value. The authors proved linear-like closed-loop behaviour, but a proof for asymptotic tracking remained elusive. In [35], the higher-order non-minimum phase system is considered. Here, an estimator and the pole-placement based control law are designed for each of the convex sets and a dynamic switching rule is introduced to choose between the controllers at each time step. Not only linear-like closed-loop behaviour but also tracking for the sum of a finite number of sinusoids of known frequency is shown.

## 1.2 The Objective

The main objective of this research is to extend the work presented in [34] in a way that provides for asymptotic tracking. Here, we consider the first-order discrete-time plant where the sign of the input parameter $b$ is unknown. To achieve this objective, we adopt the following approach:

- We start with covering the compact uncertainty set of plant parameters by a pair of convex sets, one for each sign of the input parameter $b$.
- Then, using the revised version of the original projection algorithm [34], we design an estimator for each of the convex sets. Each of these estimators has the corresponding one-step-ahead controller.
- Finally, we propose a dynamic performance signal based switching algorithm that chooses the best suitable one-step-ahead controller for each time step; the performance signals measure the"accuracy" of estimation over time.

In this thesis, analogous to [34], the closed-loop system is proven to have a uniform exponential decay bound on the effect of the initial condition, and a convolution sum bound on the effect of exogenous signals, i.e., both noise and reference. Furthermore, robustness to a degree of time-variations and unmodelled dynamics is also shown.

To show the tracking performance of the closed-loop system in the absence of noise, we first prove that the tracking error goes to zero if the switching rule stops switching. We then use this result to prove asymptotic trajectory tracking for two different circumstances:

(i) when the reference signal is asymptotically strongly persistently exciting of order 2, and

(ii) when the reference signal is fairly general but the uncertainty set that contains plant parameters satisfies several conditions.

We now outline the rest of the thesis. In Chapter 2, we introduce the setup, the details of the parameter estimation algorithm, and the adaptive control law with the proposed switching rule. Chapter 3 presents the stability results and discusses the homogeneity property of the closed-loop system. Results related to the tracking ability of the proposed approach are shown in Chapter 4. Numerical simulations are presented in Chapter 5 which show the efficacy of our proposed approach. Finally, in Chapter 6, we present a summary of our results and some future research directions.

## 1.3 Notation

Let $\mathbb{R}$ denote the set of real numbers, $\mathbb{R}^+$ the set of non-negative real numbers, $\mathbb{Z}$ the set of integers, $\mathbb{Z}^+$ the set of non-negative integers, and $\mathbb{N}$ the set of natural numbers. The symbol $\mathbb{D}^0$ denotes the open unit disc of the complex plane. Let $\lceil \cdot \rceil$ denote the ceiling function. The Euclidean 2-norm is used for vectors and the corresponding induced norm for matrices, and we denote the norm of a vector or matrix by $\| \cdot \|$. We let $\mathbb{S}(\mathbb{R}^{n\times m})$ denote the set of all $\mathbb{R}^{n\times m}$-valued sequences, and $l_\infty(\mathbb{R}^{n\times m})$ denote the subset of bounded sequences, where we define $l_\infty := l_\infty(\mathbb{R})$. The $\infty$-norm of $u \in l_\infty(\mathbb{R}^{n\times m})$ equals $\|u\|_\infty := \sup_{k\in\mathbb{Z}} \|u(k)\|$. Furthermore, adjoint of a matrix is denoted as $\text{adj}(\cdot)$, whereas, determinant of a matrix is denoted by $\det(\cdot)$. Additionally, $I_p$ denotes the identity matrix of size $p$. If $A$, $B$ are square matrices of the same dimension, then $B \leq A$ means that $A - B$ is positive semidefinite.

If $\Omega \subset \mathbb{R}^p$ is a compact set, we define $||\Omega|| := \sup_{x\in\Omega} ||x||$. If $\Omega \subset \mathbb{R}^p$ is a convex and compact set, the function $\text{Proj}_\Omega : \mathbb{R}^p \to \Omega$ denotes the projection onto $\Omega$; it is well known that $\text{Proj}_\Omega$ is well defined. Lastly, the closed convex hull of $\Omega \subseteq \mathbb{R}^p$ is denoted by $\text{conv}(\Omega)$.

# Chapter 2

# The Setup

In this chapter, we present the plant and the assumptions imposed on it. We review the revised estimation algorithm that we will be using in our proposed controller. Finally, we introduce a dynamic performance signal based switching rule along with an adaptive control law.

## 2.1 The Plant

We consider the first order linear time-invariant discrete-time plant

$$x(t+1) = a\,x(t) + b\,u(t) + n(t), \qquad x(t_0) = x_0, \tag{2.1}$$

where $x(t) \in \mathbb{R}$ is the state variable, $u(t) \in \mathbb{R}$ is the control input, and $n(t) \in \mathbb{R}$ is the noise (or disturbance); define

$$\theta^* := \begin{bmatrix} a \\ b \end{bmatrix}, \qquad \phi(t) := \begin{bmatrix} x(t) \\ u(t) \end{bmatrix},$$

with $\phi(t_0) = \phi_0$.

**Assumption 2.1.** *The set of admissible plant parameters $\mathcal{S}$ is known and compact.*

For practical situations, the compactness assumption imposed on the set $\mathcal{S}$ is reasonable and ensures that we can prove uniform bounds and exponential decay rates on the closed-loop behaviour. The next assumption ensures controllability.

**Assumption 2.2.** *For every $a \in \mathbb{R}$,* $\begin{bmatrix} a \\ 0 \end{bmatrix} \notin \mathcal{S}$.

In previous works, such as [8] and [21], the set of admissible parameters was convex. Here, we have relaxed this requirement as the set $\mathcal{S}$ need not be convex. The usual tactic in this case is to replace the set with its closed convex hull. Regrettably, this trick has not been found useful, as the set may contain uncontrollable models (i.e. $b = 0$) – see [35]. The method adopted here, as presented in [34] and [35], is to replace the compact set of admissible parameters $\mathcal{S}$ by a finite number of convex sets. The following proposition illustrates that we can always obtain a cover with just two closed convex sets.

**Proposition 2.1** ([34, Propositon 1])**.** *For any compact set* $\mathcal{S} \subset \left\{ \begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{R}^2 : b \neq 0 \right\}$, *there exist compact and convex sets* $\mathcal{S}_1$ *and* $\mathcal{S}_2$ *which also lie in* $\left\{ \begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{R}^2 : b \neq 0 \right\}$ *such that* $\mathcal{S} \subset \mathcal{S}_1 \cup \mathcal{S}_2$.

To each set $\mathcal{S}_1$, $\mathcal{S}_2$ from Proposition 2.1 we will assign a parameter estimator and we will switch between estimates from time to time for use in the control law, which we discuss in Section 2.4.

**Remark 2.1.** *If a convex set is complicated, it may be difficult (numerically) to project onto it. If we define* $\bar{a} := \max\left\{ |a| : \begin{bmatrix} a \\ b \end{bmatrix} \in \mathcal{S} \right\}, \bar{b} := \max\left\{ |b| : \begin{bmatrix} a \\ b \end{bmatrix} \in \mathcal{S} \right\}$ *and* $\underline{b} := \min\left\{ |b| : \begin{bmatrix} a \\ b \end{bmatrix} \in \mathcal{S} \right\}$, *then Proposition 2.1 also holds with*

$$\mathcal{S}_1 := \left\{ \begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{R}^2 : a \in [-\bar{a}, \bar{a}], b \in [\underline{b}, \bar{b}] \right\}$$

*and*

$$\mathcal{S}_2 := \left\{ \begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{R}^2 : a \in [-\bar{a}, \bar{a}], b \in [-\bar{b}, -\underline{b}] \right\}$$

*which are rectangles.*

In the light of Proposition 2.1, we consider that $\mathcal{S} \subset \mathcal{S}_1 \cup \mathcal{S}_2$ and that each set $\mathcal{S}_i$, $i \in \{1, 2\}$ is known, convex, compact and satisfies $\begin{bmatrix} a \\ 0 \end{bmatrix} \notin \mathcal{S}_i$ for every $a \in \mathbb{R}$.

With $\theta^*$, we let $g(\theta^*)$ denote the index $i$ for which $\theta^* \in \mathcal{S}_i$; since it could be that $\mathcal{S}_1 \cap \mathcal{S}_2 \neq \varnothing$, a precise definition is

$$g(\theta^*) = \min\{i : \theta^* \in \mathcal{S}_i\}. \tag{2.2}$$

To minimize notation, when there is no risk of confusion, we will drop the argument and simply denote this index by $g$ indicating that $\mathcal{S}_g$ is the "good" set because it contains the true parameter vector. We also define the constant $\bar{\mathbf{s}} := \max_i\{||\mathcal{S}_i||\}$.

## 2.2 Parameter Estimation

Given an estimate $\hat{\theta}_i$ of $\theta^*$ at time $t$, we define the associated prediction error by

$$e_i(t+1) := x(t+1) - \phi(t)^\top \hat{\theta}_i(t); \tag{2.3}$$

this is a measure of the error in $\hat{\theta}_i(t)$. The common way to obtain a new estimate is by solving the optimization problem

$$\underset{\theta}{\operatorname{argmin}} \left\{ \|\theta - \hat{\theta}_i(t)\| : x(t+1) = \phi(t)^\top \theta \right\},$$

yielding the ideal projection algorithm

$$\hat{\theta}_i(t+1) = \begin{cases} \hat{\theta}_i(t), & \text{if } \phi(t) = 0, \\ \hat{\theta}_i(t) + \dfrac{\phi(t)}{\phi(t)^\top \phi(t)} e_i(t+1), & \text{otherwise.} \end{cases} \tag{2.4}$$

Of course, if $\phi(t)$ is close to zero, numerical problems can occur, so it is the norm in the literature, e.g. [8], [7], to replace this by the **classical algorithm**; with $0 < \alpha < 2$ and $\beta > 0$, define

$$\hat{\theta}_i(t+1) = \hat{\theta}_i(t) + \frac{\alpha \phi(t)}{\beta + \phi(t)^\top \phi(t)} e_i(t+1). \tag{2.5}$$

As discussed in [18] and[19], when the algorithm (2.5) is used all of the stability results are asymptotic, and exponential stability and a bounded gain on the noise are never proven.

The reason being that the gain on the update law of the estimator is small for small values of $\|\phi(t)\|$. To address this issue, as proposed in [21], one way is to use the ideal projection algorithm (2.4) together with projection onto $\mathcal{S}_i$. Towards this end, we set, for $i \in \{1, 2\}$,

$$\check{\theta}_i(t+1) = \begin{cases} \hat{\theta}_i(t), & \text{if } \phi(t) = 0, \\ \hat{\theta}_i(t) + \dfrac{\phi(t)}{\phi(t)^\top \phi(t)}\, e_i(t+1), & \text{otherwise,} \end{cases} \tag{2.6}$$

which we project onto $\mathcal{S}_i$:

$$\hat{\theta}_i(t+1) = \mathrm{Proj}_{\mathcal{S}_i}\left(\check{\theta}_i(t+1)\right). \tag{2.7}$$

## 2.3 Revised Estimation Algorithm

One may be concerned that the original problem of dividing by a number close to zero, which motivates the use of classical algorithm, remains. To deal with this issue, as presented in [21], we follow a middle ground. A simple analysis of (2.3) reveals that

$$e_i(t+1) = \phi(t)^\top \left(\theta^* - \hat{\theta}_i(t)\right) + n(t), \tag{2.8}$$

which means that

$$|e_i(t+1)| \leq 2\,\bar{\mathbf{s}} \times ||\phi(t)|| + |n(t)|. \tag{2.9}$$

Therefore, if

$$|e_i(t+1)| > 2\,\bar{\mathbf{s}} \times ||\phi(t)||,$$

then the update to $\hat{\theta}_i(t)$ will be greater than $2\,\bar{\mathbf{s}}$, which means that there is little information content in $e_i(t+1)$ – it is dominated by the disturbance. With this as motivation, and with $\delta \in (0, \infty)$, we replace (2.6) with

$$\check{\theta}_i(t+1) = \begin{cases} \hat{\theta}_i(t) + \dfrac{\phi(t)}{||\phi(t)||^2}\, e_i(t+1), & \text{if } \dfrac{|e_i(t+1)|}{||\phi(t)||} < (2\,\bar{\mathbf{s}} + \delta), \\ \hat{\theta}_i(t), & \text{otherwise.} \end{cases} \tag{2.10}$$

The algorithm (2.10) assures that the update term is bounded above by $2\,\bar{\mathbf{s}} + \delta$, which should alleviate concerns about having infinite gain. In order to write the update rule (2.10) more concisely, we define $\rho_\delta\ : \mathbb{R}^2 \times \mathbb{R} \to \{0, 1\}$ and $\rho_i(t) := \rho_\delta(\phi(t), e_i(t+1))$, where

$$\rho_\delta(\phi(t), e_i(t+1)) := \begin{cases} 1, & \text{if } |e_i(t+1)| < (2\,\bar{\mathbf{s}} + \delta)\, ||\phi(t)||, \\ 0, & \text{otherwise.} \end{cases} \tag{2.11}$$

Using the above notation, the estimation algorithm (2.7), (2.10) is written

$$\check{\theta}_i(t+1) = \hat{\theta}_i(t) + \rho_i(t)\frac{\phi(t)}{||\phi(t)||^2}\, e_i(t+1), \tag{2.12}$$

$$\hat{\theta}_i(t+1) = \mathrm{Proj}_{\mathcal{S}_i}\Big(\check{\theta}_i(t+1)\Big). \tag{2.13}$$

In the case of $\phi(t) = 0$, we adopt the understanding that $0 \div 0 := 0$.

The next proposition presents a bound on $\phi$ when we turn off the update on the estimator, i.e., when $\rho_i(t) = 0$ – which only happens when the estimator is inundated with noise.

**Proposition 2.2.** *For every $t_0 \in \mathbb{Z}$, $t \geq t_0$, $x_0 \in \mathbb{R}$, $\theta^* \in \mathcal{S}$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$, $i \in \{1, 2\}$, $n \in l_\infty$ and $\delta \in (0, \infty)$, if $\rho_i(t) = 0$, then*

$$||\phi(t)|| \leq \frac{1}{\delta}\, |n(t)|. \tag{2.14}$$

*Proof.* Let $t_0 \in \mathbb{Z}$, $x_0 \in \mathbb{R}$, $\theta^* \in \mathcal{S}$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$, $i \in \{1, 2\}$, $n \in l_\infty$ and $\delta \in (0, \infty)$ be arbitrary. For every $t \geq t_0$, if $\rho_i(t) = 0$, then from (2.11) we have

$$|e_i(t+1)| \geq (2\,\bar{\mathbf{s}} + \delta)||\phi(t)||. \tag{2.15}$$

Combining (2.9) and (2.15) yields,

$$2\,\bar{\mathbf{s}} \times ||\phi(t)|| + |n(t)| \geq (2\,\bar{\mathbf{s}} + \delta)||\phi(t)||,$$

which, rearranging, gives (2.14). ■

**Remark 2.2.** *If the disturbance $n(t) = 0$, then the estimation algorithm (2.11), (2.12), (2.13) enjoys a nice scaling property. In this case, if $\phi(t) \neq 0$ then $\rho_i(t) = 1$ and (2.12) becomes*

$$\check{\theta}_i(t+1) = \hat{\theta}_i(t) + \frac{\phi(t)\phi(t)^\top}{\phi(t)^\top\phi(t)}\big(\theta^* - \hat{\theta}_i(t)\big).$$

*Thus, if $\phi(t)$ is replaced by $\gamma\phi(t)$ with $\gamma \neq 0$, then the update $\check{\theta}(t+1)$ (and its projection $\hat{\theta}(t+1)$) remains unchanged. Notice that the classical algorithm (2.5) does not enjoy this property. This provides a clue that the estimation algorithm (2.11), (2.12), (2.13) may provide closed-loop properties not provided by (2.5).*

Define for each $i \in \{1, 2\}$ the parameter estimation error $\tilde{\theta}_i(t) \coloneqq \hat{\theta}_i(t) - \theta^*$. The following proposition lists properties of the estimation algorithm (2.11), (2.12), (2.13).

**Proposition 2.3** ([35, Proposition 2])**.** *For every $t_0 \in \mathbb{Z}$, $x_0 \in \mathbb{R}$, $\theta^* \in \mathcal{S}$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$, $i \in \{1, 2\}$, $n \in l_\infty$, when the estimation algorithm (2.11), (2.12), (2.13) is applied to the plant (2.1), the following holds:*

*(a) For* ***every estimator****, we have*

$$||\hat{\theta}_i(t) - \hat{\theta}_i(t_0)|| \leq \sum_{j=t_0}^{t-1} \rho_i(j) \times \frac{|e_i(j+1)|}{||\phi(j)||}, \quad t > t_0. \tag{2.16}$$

*(b) For the* ***correct estimator****, we have*

$$||\tilde{\theta}_g(t)||^2 \leq ||\tilde{\theta}_g(t_0)||^2 + \sum_{j=t_0}^{t-1} \rho_g(j) \left( -\frac{1}{2}\frac{|e_g(j+1)|^2}{||\phi(j)||^2} + 2\,\frac{|n(j)|^2}{||\phi(j)||^2} \right), \quad t > t_0. \tag{2.17}$$

*Proof.* Let $t_0 \in \mathbb{Z}$, $x_0 \in \mathbb{R}$, $\theta^* \in \mathcal{S}$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$, $i \in \{1,2\}$, $n \in l_\infty$ be arbitrary. For the estimation algorithm (2.11), (2.12), (2.13), since projection onto a convex set does not make the parameter estimate worse; for $t \geq t_0$, it follows from (2.12), (2.13) that

$$\begin{aligned} ||\hat{\theta}_i(t+1) - \hat{\theta}_i(t)|| &\leq ||\check{\theta}_i(t+1) - \hat{\theta}_i(t)|| \\ &= \left\| \rho_i(t)\, \frac{\phi(t)\, e_i(t+1)}{||\phi(t)||^2} \right\| \\ &= \rho_i(t)\, \frac{||\phi(t)||\, |e_i(t+1)|}{||\phi(t)||^2} \\ &= \rho_i(t)\, \frac{|e_i(t+1)|}{||\phi(t)||}. \end{aligned}$$

We conclude that (2.16) follows by iteration.

To prove (2.17), first define $\tilde{\check{\theta}}_g(t) := \check{\theta}_g(t) - \theta^*$. When $\rho_g(t) = 0$, we have $\hat{\theta}_g(t+1) = \hat{\theta}_g(t)$ by (2.12), (2.13), which implies that

$$||\tilde{\theta}_g(t+1)||^2 = ||\tilde{\theta}_g(t)||^2. \tag{2.18}$$

On the other hand, when $\rho_g(t) = 1$, by (2.12), (2.13) we have

$$\begin{aligned} \tilde{\check{\theta}}_g(t+1) &= \tilde{\theta}_g(t) + \frac{\phi(t)}{||\phi(t)||^2}\, e_g(t+1) \\ \Rightarrow ||\tilde{\check{\theta}}_g(t+1)||^2 &= ||\tilde{\theta}_g(t)||^2 + \frac{|e_g(t+1)|^2}{||\phi(t)||^2} + 2\, \frac{\tilde{\theta}_g(t)^\top \phi(t)\, e_g(t+1)}{||\phi(t)||^2}. \end{aligned} \tag{2.19}$$

Combining (2.8) and (2.19) we obtain

$$\begin{aligned} ||\tilde{\check{\theta}}_g(t+1)||^2 &= ||\tilde{\theta}_g(t)||^2 + \frac{|e_g(t+1)|^2}{||\phi(t)||^2} + 2\, \frac{\big(n(t) - e_g(t+1)\big)\, e_g(t+1)}{||\phi(t)||^2} \\ &= ||\tilde{\theta}_g(t)||^2 - \frac{|e_g(t+1)|^2}{||\phi(t)||^2} + 2\, \frac{n(t)\, e_g(t+1)}{||\phi(t)||^2} \\ &\leq ||\tilde{\theta}_g(t)||^2 - \frac{1}{2}\, \frac{|e_g(t+1)|^2}{||\phi(t)||^2} + 2\, \frac{|n(t)|^2}{||\phi(t)||^2} \end{aligned}$$

(the last step uses the fact that for $a, b \geq 0$, we have $-a^2 + 2ab \leq -\frac{1}{2}a^2 + 2b^2$). We can rewrite the above, again using the fact that the projection does not worsen the parameter estimate, as

$$||\tilde{\theta}_g(t+1)||^2 \leq ||\tilde{\theta}_g(t)||^2 - \frac{1}{2}\frac{|e_g(t+1)|^2}{||\phi(t)||^2} + 2\frac{|n(t)|^2}{||\phi(t)||^2}. \tag{2.20}$$

By combining the bound (2.20) for the case of $\rho_g(t) = 1$ with (2.18) for the case of $\rho_g(t) = 0$, and iterating over $t$, we obtain (2.17) as required. ■

## 2.4 Switching Controller

It is natural to parametrize $\hat{\theta}_i(t)$ as

$$\hat{\theta}_i(t) =: \begin{bmatrix} \hat{a}_i(t) \\ \hat{b}_i(t) \end{bmatrix}.$$

Let $r$ be an exogenous reference signal that we want the plant state to track asymptotically. We assume that the value of $r$ is known one step ahead, i.e., $r(t+1)$ is known at time $t$. If we invoke the Certainty Equivalence Principle there is a natural choice for the one-step-ahead adaptive control law associated with the $i$th estimator:

$$u(t) = -\frac{\hat{a}_i(t)}{\hat{b}_i(t)}\,x(t) + \frac{1}{\hat{b}_i(t)}\,r(t+1), \tag{2.21}$$

which enforces that $r(t+1) = \phi(t)^\top \hat{\theta}_i(t)$. Here, of course, we do not know which set $\mathcal{S}_i$ contains the true parameter $\theta^*$.

We will define a switching signal $\sigma : \mathbb{Z} \to \{1, 2\}$ which decides which parameter estimates to use at each point in time and we set

$$u(t) = -\frac{\hat{a}_{\sigma(t)}(t)}{\hat{b}_{\sigma(t)}(t)}\,x(t) + \frac{1}{\hat{b}_{\sigma(t)}(t)}\,r(t+1). \tag{2.22}$$

The tracking error $\varepsilon$ at time $t$ equals

$$\varepsilon(t) \coloneqq x(t) - r(t). \tag{2.23}$$

Let us analyze the relationship between the tracking error and prediction error when the control law (2.22) is applied:

$$\begin{aligned}\varepsilon(t+1) &= x(t+1) - r(t+1) \\ &= x(t+1) - \phi(t)^\top \hat{\theta}_{\sigma(t)}(t) \\ &= e_{\sigma(t)}(t+1), \qquad t \geq t_0,\end{aligned} \tag{2.24}$$

which shows that $\varepsilon$ at time $t+1$ is affected by the value of $\sigma$ at time $t$.

In recent works, such as [34], the following switching rule was adopted

$$\sigma(t) = \operatorname*{argmin}_i |e_i(t)|, \quad t > t_0. \tag{2.25}$$

This simple rule compares the prediction errors for each estimator and chooses the estimator index whose prediction error has the smallest absolute value. Since it is memoryless, the switching decision is based on the current values of the prediction errors. While the simplicity of the rule (2.25) is appealing, a downside is that the proof of asymptotic tracking using this switching algorithm remains elusive [21].

**Remark 2.3.** *We suspect that, in general, it will not be possible to prove that the switching algorithm of [34] can be used to prove asymptotic tracking. To see why we think this, suppose that $\mathcal{S}_i$ are as chosen in Remark 2.1 and $\underline{b} > 0$, which means that $\mathcal{S}_1 \cap \mathcal{S}_2 = \varnothing$. Suppose, for simplicity, that the plant parameter lies in $\mathcal{S}_1$, and we apply the controller of [34]. The prediction error $e_2$ is given by*

$$\begin{aligned}e_2(t+1) &= x(t+1) - \phi(t)^\top \hat{\theta}_2(t) \\ &= \big(a - \hat{a}_2(t)\big)\, x(t) + \big(b - \hat{b}_2(t)\big)\, u(t).\end{aligned}$$

*If $\sigma(t) = 1$, then $u(t)$ is being generated by the correct estimator $\hat{\theta}_1(t)$; it is easy to verify that*

$$e_2(t+1) = \left(a - \hat{a}_2(t) - \frac{\hat{a}_1(t)\,\big(b - \hat{b}_2(t)\big)}{\hat{b}_1(t)}\right) x(t) + \frac{b - \hat{b}_2(t)}{\hat{b}_1(t)}\, r(t+1).$$

*It could certainly be that $r(t+1)$ is such that*

$$e_2(t+1) = 0,$$

*in which case*

$$\sigma(t+1) = 2,$$

*unless, by luck, $e_1(t+1) = 0$ as well. Of course, it is conceivable that $r(t)$ is such that this happens an infinite number of times, in which case switching never stops. Since, in general, we would not expect the incorrect estimator to yield tracking, we would not expect that asymptotic tracking would occur in this case. If each time the wrong estimator is selected, the absolute value of the tracking error jumps by at least a fixed constant, then the lim sup of the tracking error is positive and asymptotic tracking can not be achieved; this scenario is difficult to rule out.*

Motivated by this observation, we introduce dynamic performance signals $J_i$, $i \in \{1, 2\}$, which accumulate scaled prediction error over time:

$$J_i(t+1) = \bar{\lambda}\, J_i(t) + \rho_i(t)\, \frac{|e_i(t+1)|}{||\phi(t)||}, \quad J_i(t_0) = 0, \tag{2.26}$$

where $\bar{\lambda} \in (0, 1)$. We adopt a switching rule that selects the estimator index associated to the minimum performance signal $J_i(t)$.

$$\sigma(t) = \underset{i}{\operatorname{argmin}}\, J_i(t), \qquad t > t_0, \quad \sigma(t_0) = \sigma_0 \in \{1, 2\}. \tag{2.27}$$

In the case when $J_1(t) = J_2(t)$ we (somewhat arbitrarily) select $\sigma(t)$ to be 1.

**Remark 2.4.** *When we set $\bar{\lambda} = 0$, our performance signal based switching rule (2.26), (2.27) recovers (2.25) which was presented in [34].*

Before discussing closed-loop stability, we first show that the simple logic in (2.26), (2.27) yields a useful closed-loop property, which is analogous to [21, Lemma 2].

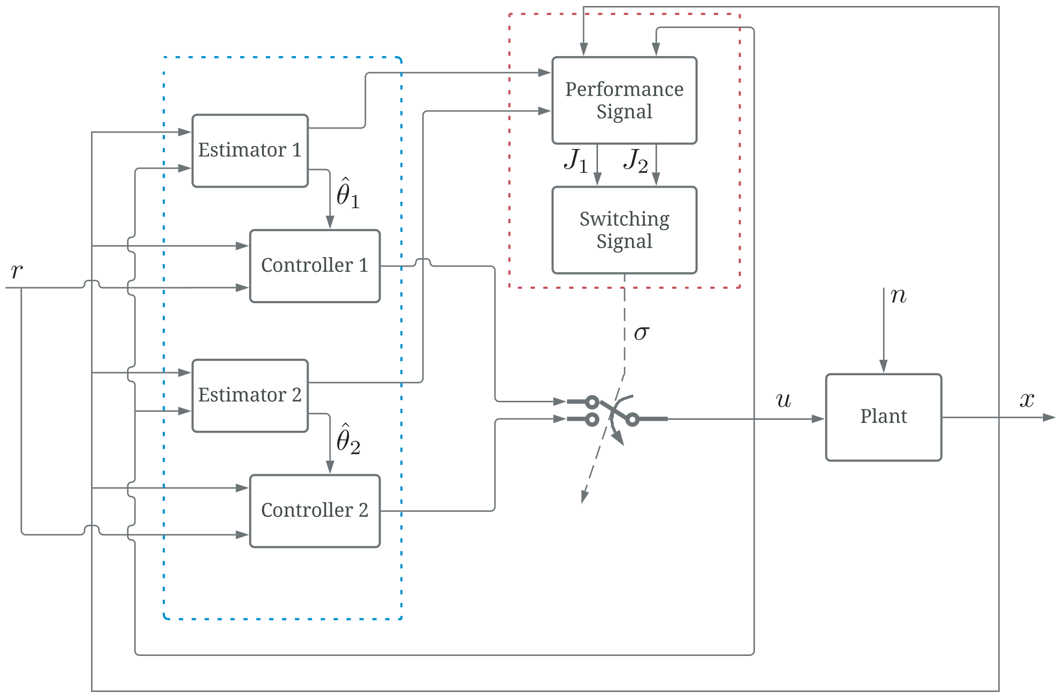



Figure 2.1: The Block diagram of the closed-loop system; enclosed inside the dashed boxes are the multiple estimators/controllers (blue), and the switching mechanism (red).

**Lemma 2.1.** *Suppose that the adaptive controller consisting of the parameter estimator (2.11), (2.12), (2.13), the control law (2.22), and the switching rule (2.26), (2.27) is applied to the plant (2.1). Then, for every* $t_0 \in \mathbb{Z}$, $x_0 \in \mathbb{R}$, $\theta^* \in \mathcal{S}$, $\sigma_0 \in \{1, 2\}$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$, $i \in \{1, 2\}$, $r$, $n \in l_\infty$ *and* $t > t_0$ *we have that either*

(a) $J_{\sigma(t-1)}(t) \leq J_i(t)$ or

(b) $J_{\sigma(t)}(t+1) \leq J_i(t+1)$.

*Proof.* Fix $t_0 \in \mathbb{Z}$, $x_0 \in \mathbb{R}$, $\theta^* \in \mathcal{S}$, $\sigma_0 \in \{1, 2\}$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$, $i \in \{1, 2\}$, $r$, $n \in l_\infty$ and $t > t_0$. Let $\bar{\imath}$ be the element of $\{1, 2\}$ which is not $i$.

Suppose (b) fails to hold then, it must be that $\sigma(t) = \bar{\imath}$. From (2.27), it follows that

$J_{\bar{\imath}}(t) \leq J_i(t)$. But $\sigma(t-1) \in \{i, \bar{\imath}\}$ and therefore $J_{\sigma(t-1)}(t) \leq J_i(t)$, i.e., (a) holds. ■

# Chapter 3

# Stability

Building upon the setup of Chapter 2, in this chapter we prove a linear-like convolution bound on the closed-loop system along with a uniform exponential decay rate. Next, we show that the closed-loop system enjoys the homogeneity property. In the last section, we prove robustness of the proposed controller to time-variations and unmodelled dynamics.

## 3.1 Closed-Loop Stability

In the main result of this chapter, we show that the closed-loop system exhibits linear-like behaviour.

**Theorem 3.1.** *Suppose that the adaptive controller consisting of the parameter estimator (2.11), (2.12), (2.13), the control law (2.22), and the switching rule (2.26), (2.27) is applied to the plant (2.1). For every* $\bar{\lambda} \in (0,1)$ *and* $\lambda \in (0,1)$*, there exists a constant* $\bar{\gamma} \geq 1$ *such that for every* $t_0 \in \mathbb{Z}$, $x_0 \in \mathbb{R}$, $\theta^* \in \mathcal{S}$, $\sigma_0 \in \{1,2\}$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$, $i \in \{1,2\}$, $r$, $n \in l_\infty$*, the closed-loop system satisfies*

$$||\phi(t)|| \leq \bar{\gamma}\, \lambda^{t-t_0}\, |x_0| + \sum_{j=t_0}^{t-1} \bar{\gamma}\, \lambda^{t-1-j} \Big( |n(j)| + |r(j+1)| \Big) + \bar{\gamma}\, |r(t+1)|, \; t \geq t_0. \quad (3.1)$$

In Theorem 3.1 there is a uniform exponential decay bound on the effect of initial condition, and a convolution sum bound on the effect of exogenous signals (noise and reference). This result is analogous to the single uncertainty set case discussed in [19].

Furthermore, a similar convolution bound on the closed-loop system was presented in [34] for the memoryless switching rule (2.25).

*Proof.* Fix $\bar{\lambda} \in (0,1)$, $\lambda \in (0,1)$, $t_0 \in \mathbb{Z}$, $x_0 \in \mathbb{R}$, $\theta^* \in \mathcal{S}$, $\sigma_0 \in \{1,2\}$, $i \in \{1,2\}$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$, $n \in l_\infty$, and $r \in l_\infty$. Define the constants

$$\bar{a} := \max\left\{ |a| : \begin{bmatrix} a \\ b \end{bmatrix} \in \mathcal{S}_1 \cup \mathcal{S}_2 \right\}, \quad \bar{b} := \max\left\{ |b| : \begin{bmatrix} a \\ b \end{bmatrix} \in \mathcal{S}_1 \cup \mathcal{S}_2 \right\},$$

$$\bar{f} := \max\left\{ \left|\frac{a}{b}\right| : \begin{bmatrix} a \\ b \end{bmatrix} \in \mathcal{S}_1 \cup \mathcal{S}_2 \right\}, \quad \bar{g} := \max\left\{ \frac{1}{|b|} : \begin{bmatrix} a \\ b \end{bmatrix} \in \mathcal{S}_1 \cup \mathcal{S}_2 \right\}.$$

First we establish some general bounds to be used throughout the proof. Setting $c_1 := (1+\bar{f})$ and $c_2 := \bar{g}$, from the control law in (2.22) we obtain the general bound

$$||\phi(t)|| \le c_1|x(t)| + c_2|r(t+1)|\,; \tag{3.2}$$

if we define $c_3 := \max\{\bar{a} + \bar{b}\bar{f}, \bar{b}\bar{g}\}$, then from the plant equation (2.1) we obtain the crude bound

$$|x(t+1)| \le c_3|x(t)| + c_3|r(t+1)| + |n(t)|. \tag{3.3}$$

At this point we divide the proof into two cases; the easier case in which there is no noise followed by the harder case in which there is noise. As in [33], a separate analysis for the less general noise-free case is included to help the reader's understanding.

**Case 1:** $n(t) = 0$ for all $t \ge t_0$.

First, we will analyze the behaviour for two consecutive time instants and then we will consider the whole time horizon. From Proposition 2.3 we have that

$$||\tilde{\theta}_g(t)||^2 \le ||\tilde{\theta}_g(t_0)||^2 - \frac{1}{2}\sum_{j=t_0}^{t-1} \rho_g(j)\,\frac{|e_g(j+1)|^2}{||\phi(j)||^2}, \quad t > t_0.$$

Using the fact $||\tilde{\theta}_g(t_0)|| \le 2||\mathcal{S}_g||$, the above implies

$$\sum_{j=t_0}^{t-1} \rho_g(j)\,\frac{|e_g(j+1)|^2}{||\phi(j)||^2} \le 2||\tilde{\theta}_g(t_0)||^2 \le 8||\mathcal{S}_g||^2 \le 8\,\bar{\mathbf{s}}^2, \quad t > t_0. \tag{3.4}$$

The bound (3.4) shows that the input to (2.26) with $i = g$ is square summable. Applying Parseval's theorem for discrete-time systems to (2.26), noting that $\bar{\lambda} < 1$, we have

$$\sum_{t=t_0}^{\infty} J_g^2(t) \leq \frac{8\,\bar{\mathbf{s}}^2}{(1-\bar{\lambda})^2} =: \bar{c}_4. \tag{3.5}$$

Before proceeding further, define $c_4 \coloneqq \max\{1, \bar{c}_4\}$.

**Claim 3.1.** *For every $t > t_0$ and $i \in \{1, 2\}$, one of the following bounds hold*

(i) $|\varepsilon(t)| \leq ||\phi(t-1)||\,|J_i(t)|$, or

(ii) $|\varepsilon(t+1)| \leq ||\phi(t)||\,|J_i(t+1)|$.

*Proof of Claim 3.1.* Fix $t > t_0$ and $i \in \{1, 2\}$. Motivated by the result in Lemma 2.1 we split the proof of the claim into two cases.

**Case 1:** $J_{\sigma(t-1)}(t) \leq J_i(t)$.

In this case, from (2.26), we have

$$\bar{\lambda}\, J_{\sigma(t-1)}(t-1) + \rho_{\sigma(t-1)}(t-1)\,\frac{|e_{\sigma(t-1)}(t)|}{||\phi(t-1)||} \leq \bar{\lambda}\, J_i(t-1) + \rho_i(t-1)\,\frac{|e_i(t)|}{||\phi(t-1)||}. \tag{3.6}$$

On the one hand, if either $\rho_{\sigma(t-1)}(t-1) = 0$ or $\rho_i(t-1) = 0$, then from Proposition 2.2, we have $||\phi(t-1)|| = 0$, which implies that $x(t-1) = u(t-1) = 0$, so that $x(t) = 0$. From the formula for $u(t-1)$ in (2.22), we see that $r(t) = 0$ as well, so that $\varepsilon(t) = 0$ and (i) holds.

On the other hand, if $\rho_{\sigma(t-1)}(t-1) = \rho_i(t-1) = 1$, then using (3.6), we have

$$\begin{aligned} \frac{|e_{\sigma(t-1)}(t)|}{||\phi(t-1)||} &\leq \bar{\lambda}\, J_i(t-1) + \frac{|e_i(t)|}{||\phi(t-1)||} \\ \Rightarrow |e_{\sigma(t-1)}(t)| &\leq |e_i(t)| + \bar{\lambda}\, J_i(t-1)||\phi(t-1)||. \end{aligned}$$

Using the relation (2.24) between the tracking error and the estimation error, we can rewrite the above as

$$\begin{aligned}|\varepsilon(t)| &\leq |e_i(t)| + \bar{\lambda}\, J_i(t-1)||\phi(t-1)|| \\ &= ||\phi(t-1)|| \left[\bar{\lambda}\, J_i(t-1) + \frac{|e_i(t)|}{||\phi(t-1)||}\right] \\ &= ||\phi(t-1)||\ |J_i(t)|\end{aligned}$$

which shows that (i) holds.

**Case 2:** $J_{\sigma(t)}(t+1) \leq J_i(t+1)$.

In this case, from (2.26), we have

$$\bar{\lambda}\, J_{\sigma(t)}(t) + \rho_{\sigma(t)}(t)\,\frac{|e_{\sigma(t)}(t+1)|}{||\phi(t)||} \leq \bar{\lambda}\, J_i(t) + \rho_i(t)\,\frac{|e_i(t+1)|}{||\phi(t)||}. \tag{3.7}$$

On the one hand, if either $\rho_{\sigma(t)}(t) = 0$ or $\rho_i(t) = 0$, then from Proposition 2.2, we have that $||\phi(t)|| = 0$, which implies that $x(t) = u(t) = 0$, so that $x(t+1) = 0$. From the formula for $u(t)$ in (2.22), we see that $r(t+1) = 0$ as well, so that $\varepsilon(t+1) = 0$ and (ii) holds.

On the other hand, if $\rho_{\sigma(t)}(t) = \rho_i(t) = 1$, then using (3.7), we have

$$\begin{aligned}&\frac{|e_{\sigma(t)}(t+1)|}{||\phi(t)||} \leq \bar{\lambda}\, J_i(t) + \frac{|e_i(t+1)|}{||\phi(t)||} \\ \Rightarrow\ &|e_{\sigma(t)}(t+1)| \leq |e_i(t+1)| + \bar{\lambda}\, J_i(t)||\phi(t)||.\end{aligned}$$

Using (2.24), we can rewrite the above as

$$\begin{aligned}|\varepsilon(t+1)| &\leq |e_i(t+1)| + \bar{\lambda}\, J_i(t)||\phi(t)|| \\ &= ||\phi(t)|| \left[\bar{\lambda}\, J_i(t) + \frac{|e_i(t+1)|}{||\phi(t)||}\right] \\ &= ||\phi(t)||\ |J_i(t+1)|\end{aligned}$$

which shows that (ii) holds. ■

Applying Claim 3.1 with $i = g$ and using the general bound (3.2) on $\phi$, we have

$$|\varepsilon(t)| \leq \big(c_1|x(t-1)| + c_2|r(t)|\big)\ J_g(t)$$

or

$$|\varepsilon(t+1)| \leq \big(c_1|x(t)| + c_2|r(t+1)|\big)\ J_g(t+1), \quad t > t_0.$$

Since $t > t_0$ is arbitrary in the above bounds, it follows that for every $j \in \mathbb{Z}^+$ we have either

$$|\varepsilon(t_0+2j+1)| \leq \big(c_1|x(t_0+2j)| + c_2|r(t_0+2j+1)|\big)\ J_g(t_0+2j+1)$$

or

$$|\varepsilon(t_0+2j+2)| \leq \big(c_1|x(t_0+2j+1)| + c_2|r(t_0+2j+2)|\big)\ J_g(t_0+2j+2).$$

For $k \geq t_0$, we define $\alpha_k := J_g(k+1)$ (note that from (3.5), $\alpha_k \leq \sqrt{c_4}$) and we define $\bar{\alpha}_{t_0+2j} := \max\{\alpha_{t_0+2j}, \alpha_{t_0+2j+1}\}$. Using the definition (2.23) of the tracking error, we can conclude that either

$$|x(t_0+2j+1)| \leq c_1\,\bar{\alpha}_{t_0+2j}|x(t_0+2j)| + (c_2\sqrt{c_4}+1)\,|r(t_0+2j+1)|$$

or

$$|x(t_0+2j+2)| \leq c_1\,\bar{\alpha}_{t_0+2j}|x(t_0+2j+1)| + (c_2\sqrt{c_4}+1)\,|r(t_0+2j+2)|.$$

Combining this with the crude bound (3.3) and defining $c_5 := \max\{1, c_1c_3\}$ and $c_6 := 1 + 2c_3 + c_1\,c_3\,{c_4}^{1/2} + c_2\,{c_4}^{1/2} + c_2\,c_3\,{c_4}^{1/2}$, we see that in either case,

$$|x(t_0+2j+2)| \leq c_5\bar{\alpha}_{t_0+2j}|x(t_0+2j)| + c_6\big(|r(t_0+2j+1)| + |r(t_0+2j+2)|\big), \quad j \in \mathbb{Z}^+. \tag{3.8}$$

Before examining the behaviour across the whole time horizon, we prove a claim.

**Claim 3.2.** *There exists a constant $c_7 > 1$ so that the following bound holds:*

$$\prod_{j=q}^{p-1} c_5\bar{\alpha}_{t_0+2j} \leq c_7\lambda^{2(p-q)}, \qquad 0 \leq q < p. \tag{3.9}$$

*Proof of Claim 3.2.* Fix $p$, $q \in \mathbb{Z}^+$ such that $p > q \geq 0$. We begin by utilizing the inequality between arithmetic and geometric means [4]:

$$\prod_{j=q}^{p-1} \bar{\alpha}_{t_0+2j} \leq \left[\frac{1}{p-q}\sum_{j=p}^{p-1} \bar{\alpha}^2_{t_0+2j}\right]^{\frac{p-q}{2}}$$
$$\leq \left[\frac{c_4}{p-q}\right]^{\frac{p-q}{2}}. \tag{3.10}$$

Now define $\lambda_1 := \dfrac{\lambda^2}{c_5}$ and $\bar{k} := \left\lceil \dfrac{c_4}{{\lambda_1}^2} \right\rceil$ from which it follows that

$$\frac{c_4}{\bar{k}} \leq {\lambda_1}^2,$$

so it is easy to see that

$$\left[\left(\frac{c_4}{k}\right)^{\frac{1}{2}}\right]^k \leq {\lambda_1}^k, \qquad k \geq \bar{k}. \tag{3.11}$$

Since $\dfrac{c_4}{k}$ decreases as $k \geq 1$ increases, if we define

$$c_7 := \frac{{c_4}^{\frac{\bar{k}}{2}}}{{\lambda_1}^{\bar{k}}},$$

then

$$\left(\frac{c_4}{k}\right)^{\frac{k}{2}} \leq c_7{\lambda_1}^k, \quad k = 1, 2, \ldots, \bar{k},$$

as well. If we combine the above with (3.11), then from (3.10) we conclude that

$$\prod_{j=q}^{p-1} \bar{\alpha}_{t_0+2j} \leq c_7{\lambda_1}^{p-q}, \qquad 0 \leq q < p.$$

Then by the definition of $\lambda_1$,

$$\prod_{j=q}^{p-1} c_5 \bar{\alpha}_{t_0+2j} \leq c_7 {\lambda_1}^{p-q} {c_5}^{p-q}$$
$$= c_7 \lambda^{2(p-q)}, \qquad 0 \leq q < p.$$

■

Now we solve the difference inequality (3.8) recursively and apply the bound (3.9) to obtain

$$|x(t_0+2j)| \leq c_7 \lambda^{2j} |x_0| + \sum_{l=0}^{j-1} c_7\, c_6 \lambda^{2(j-l-1)} \big( |r(t_0+2l+1)| + |r(t_0+2l+2)| \big), \quad j \in \mathbb{Z}^+,$$

which simplifies to

$$|x(t_0+2j)| \leq c_7 \lambda^{2j} |x_0| + \sum_{l=0}^{2j-1} \frac{c_7 c_6}{\lambda} \lambda^{2j-l-1} |r(t_0+l+1)|, \quad j \in \mathbb{Z}^+. \tag{3.12}$$

We can use (3.3) to obtain a bound for the remaining time instants. So it follows that there exists a constant $\bar{\gamma}_1 := \frac{1}{\lambda^2} \max\{c_7, c_3, c_7 c_3, c_6 c_7 c_3, c_7 c_6\}$ so that

$$|x(t)| \leq \bar{\gamma}_1 \lambda^{t-t_0} |x_0| + \sum_{j=t_0}^{t-1} \bar{\gamma}_1 \lambda^{t-j-1} |r(j+1)|, \quad t \geq t_0. \tag{3.13}$$

**Case 2:** $n(t) \neq 0$ for some $t \geq t_0$.

To complete the proof we analyze the case when there is noise entering the system; this is more complicated since $||\tilde{\theta}_g(t)||^2$ is no longer monotonically decreasing. Following [19] and [21], we partition the timeline into two parts: one in which the noise $n$ is small compared to $\phi$ and the other where it is not. Before proceeding, we define

$$\nu := \left[ \frac{\lambda_1 (1-\bar{\lambda})}{\sqrt{32}} \right]^2$$

and, motivated by the noise-free case,

$$\lambda_1 := \frac{\lambda^2}{c_5}.$$

Let us now define two sets in relation to size of the noise $n$:

$$S_{\text{good}} = \left\{ t \geq t_0 : \phi(t) \neq 0 \text{ and } \frac{|n(t)|^2}{||\phi(t)||^2} < \nu \right\},$$
$$S_{\text{bad}} = \left\{ t \geq t_0 : \phi(t) = 0 \text{ or } \frac{|n(t)|^2}{||\phi(t)||^2} \geq \nu \right\};$$

the idea is that on $S_{\text{good}}$ the disturbance is small relative to $\phi$ so the closed-loop system acts like the noise-free case, at least if $\nu$ is small enough.

Now we partition the time index $\{t \in \mathbb{Z} : t \geq t_0\}$ into intervals which oscillate between $S_{\text{good}}$ and $S_{\text{bad}}$. We can clearly define a (possible infinite) sequence of intervals of the form $[k_l, k_{l+1})$, $l \in \mathbb{Z}^+$ which satisfy:

(i) without loss of generality, $k_0 = t_0$ serves as the initial instant of the first interval;

(ii) $[k_l, k_{l+1})$ either belongs to $S_{\text{good}}$ or $S_{\text{bad}}$; and

(iii) if $k_{l+1} \neq \infty$ and $[k_l, k_{l+1})$ belongs to $S_{\text{good}}$, then $[k_{l+1}, k_{l+2})$ belongs to $S_{\text{bad}}$ and *vice versa*.

We divide this part of the proof into two sub-cases.

**Sub-Case 2.1:** $[k_l, k_{l+1})$ is a subset of $S_{\text{bad}}$.

Let $j \in [k_l, k_{l+1})$ be arbitrary. So we have that $\phi(j) = 0$ or that $\frac{|n(j)|^2}{||\phi(j)||^2} \geq \nu$. In either case, when $\rho_g(j) = 1$, we have $||\phi(j)|| \leq \frac{1}{\sqrt{\nu}}|n(j)|$. But, when $\rho_g(j) = 0$, from Proposition 2.2 we have $||\phi(j)|| \leq \frac{1}{\delta}|n(j)|$. If we define

$$c_8 := \max\left\{ \frac{1}{\sqrt{\nu}}, \frac{1}{\delta} \right\}$$

and utilize the definition of $\phi(j)$ we have

$$|x(j)| \leq c_8 |n(j)|.$$

Also, from the plant's dynamics (2.1) we obtain

$$|x(t)| \le \Big((\bar a + \bar b)c_8 + 1\Big)\,|n(t-1)|, \qquad t \in [k_l+1, k_{l+1}].$$

Define $c_9 := \Big((\bar a + \bar b)c_8 + 1\Big)$ to conclude that

$$|x(t)| \le \begin{cases} c_8|n(t)|, & t = k_l, \\ c_9|n(t-1)|, & t \in [k_l+1, k_{l+1}]. \end{cases} \tag{3.14}$$

**Sub-Case 2.2:** $[k_l, k_{l+1})$ is a subset of $S_{\text{good}}$.

From Proposition 2.3, equation (2.17), we have

$$||\tilde\theta_g(\bar k)||^2 \le ||\tilde\theta_g(\underline{k})||^2 + \sum_{j=\underline{k}}^{\bar k-1} \rho_g(j)\left(-\frac{1}{2}\,\frac{|e_g(j+1)|^2}{||\phi(j)||^2} + 2\,\frac{|n(j)|^2}{||\phi(j)||^2}\right), \quad k_l \le \underline{k} < \bar k \le k_{l+1}, \tag{3.15}$$

which can be rearranged to obtain

$$\begin{aligned} \sum_{j=\underline{k}}^{\bar k-1} \rho_g(j)\,\frac{|e_g(j+1)|^2}{||\phi(j)||^2} &\le 2||\tilde\theta_g(\underline{k})||^2 + 4\sum_{j=\underline{k}}^{\bar k-1} \rho_g(j)\frac{|n(j)|^2}{||\phi(j)||^2} \\ &\le 8\,\bar{\mathbf{s}}^2 + 4\,\nu\,(\bar k - \underline{k}), \quad \bar k > \underline{k} \ge t_0. \end{aligned} \tag{3.16}$$

Using the bound (3.16) and applying Parseval's theorem for discrete-time systems to (2.26), noting that $\bar\lambda < 1$, we have

$$\sum_{j=\underline{k}}^{\bar k-1} J_g^2(j) \le \frac{2}{(1-\bar\lambda^2)} J_g^2(\underline{k}) + \frac{2}{(1-\bar\lambda)^2}[8\,\bar{\mathbf{s}}^2 + 4\,(\bar k - \underline{k})\,\nu], \quad \bar k > \underline{k} \ge t_0. \tag{3.17}$$

Since $\delta \in (0,\infty)$, the input to (2.26) is bounded and we get

$$J_g(j) \le \frac{1}{(1-\bar\lambda)}\,(2\,\bar{\mathbf{s}} + \delta), \quad j \ge t_0. \tag{3.18}$$

Combining (3.17) and (3.18), we conclude that

$$\sum_{j=\underline{k}}^{\bar k-1} J_g^2(j) \le \frac{2}{(1-\bar\lambda)^2}\left[\frac{(2\,\bar{\mathbf{s}}+\delta)^2}{(1-\bar\lambda^2)} + 8\,\bar{\mathbf{s}}^2 + 4\,(\bar k - \underline{k})\,\nu\right], \quad \bar k > \underline{k} \ge t_0. \tag{3.19}$$

Letting $\alpha_j := J_g(j)$, we can rewrite (3.19) as

$$\sum_{j=\underline{k}}^{\bar{k}-1} \alpha_j^2 \leq \frac{2}{(1-\bar{\lambda})^2}\left[\frac{(2\,\bar{\mathbf{s}}+\delta)^2}{(1-\bar{\lambda}^2)} + 8\,\bar{\mathbf{s}}^2 + 4\,(\bar{k}-\underline{k})\,\nu\right], \quad \bar{k} > \underline{k} \geq t_0.$$

If we define $\bar{\alpha}_{k_l+2j} := \max\{\alpha_{k_l+2j}, \alpha_{k_l+2j+1}\}$, then the above implies

$$\sum_{j=q}^{p-1} \bar{\alpha}_{k_l+2j}^2 \leq \frac{2}{(1-\bar{\lambda})^2}\left[\frac{(2\,\bar{\mathbf{s}}+\delta)^2}{(1-\bar{\lambda}^2)} + 8\,\bar{\mathbf{s}}^2 + 8\,(p-q)\,\nu\right], \quad k_l \leq k_l+2q < k_l+2p \leq k_{l+1}. \tag{3.20}$$

Now define the constant

$$c_{10} := \sqrt{\frac{2}{(1-\bar{\lambda})^2}\left[\frac{(2\,\bar{\mathbf{s}}+\delta)^2}{(1-\bar{\lambda}^2)} + 8\,\bar{\mathbf{s}}^2 + 8\,\nu\right]}.$$

If we now analyze the closed-loop system as in the noise-free case, we end up with a version of (3.8) with the noise now included. Let $c_{11} := 1+2c_3+c_2c_3c_{10}+c_1c_3c_{10}+c_1c_{10}+c_2c_{10}$ so that

$$\begin{aligned}|x(k_l+2j+2)| \leq c_5\bar{\alpha}_{k_l+2j}|x(k_l+2j)| + c_{11}\big(&|r(k_l+2j+1)| + |r(k_l+2j+2)| \\ &+ |n(k_l+2j)| + |n(k_l+2j+1)|\big), \quad j \in \mathbb{Z}^+ \text{ s.t. } k_l+2j+1 < k_{l+1}.\end{aligned} \tag{3.21}$$

**Claim 3.3.** *There exists a constant $c_{12} > 1$ so that the following bound holds:*

$$\prod_{j=q}^{p-1} c_5\bar{\alpha}_{k_l+2j} \leq c_{12}\lambda^{2(p-q)}, \qquad k_l \leq k_l+2q < k_l+2p \leq k_{l+1}. \tag{3.22}$$

*Proof of Claim 3.3.* Fix $p,\, q \in \mathbb{Z}^+$ such that $k_l \leq k_l+2q < k_l+2p \leq k_{l+1}$. Using a similar analysis to that in Claim 3.2, apply the inequality of arithmetic and geometric means [4]; from (3.20) and incorporating the definition of $\nu$, we get

$$\begin{aligned}\prod_{j=q}^{p-1} \bar{\alpha}_{k_l+2j} &\leq \left[\frac{1}{(p-q)}\sum_{j=q}^{p-1}\bar{\alpha}_{k_l+2j}^2\right]^{\frac{p-q}{2}} \\ &\leq \left[\frac{2\,(2\,\bar{\mathbf{s}}+\delta)^2}{(p-q)(1+\bar{\lambda})(1-\bar{\lambda})^3} + \frac{16\,\bar{\mathbf{s}}^2}{(p-q)(1-\bar{\lambda})^2} + \frac{16\,\nu}{(1-\bar{\lambda}^2)}\right]^{\frac{p-q}{2}} \\ &= \left[\frac{2\,(2\,\bar{\mathbf{s}}+\delta)^2}{(p-q)(1+\bar{\lambda})(1-\bar{\lambda})^3} + \frac{16\,\bar{\mathbf{s}}^2}{(p-q)(1-\bar{\lambda})^2} + \frac{{\lambda_1}^2}{2}\right]^{\frac{p-q}{2}}.\end{aligned} \tag{3.23}$$

Let us define

$$\bar{k} := \left\lceil \frac{16\bar{\mathbf{s}}^2}{\frac{{\lambda_1}^2}{2}(1-\bar{\lambda})^2} + \frac{2(2\,\bar{\mathbf{s}}+\delta)^2}{\frac{{\lambda_1}^2}{2}(1+\bar{\lambda})(1-\bar{\lambda})^3} \right\rceil .$$

Then we have

$$\bar{k} \geq \frac{16\bar{\mathbf{s}}^2}{\frac{{\lambda_1}^2}{2}(1-\bar{\lambda})^2} + \frac{2(2\,\bar{\mathbf{s}}+\delta)^2}{\frac{{\lambda_1}^2}{2}(1+\bar{\lambda})(1-\bar{\lambda})^3} \,,$$

which can be rearranged as

$$\frac{{\lambda_1}^2}{2} \geq \frac{16\bar{\mathbf{s}}^2}{\bar{k}(1-\bar{\lambda})^2} + \frac{2(2\,\bar{\mathbf{s}}+\delta)^2}{\bar{k}(1+\bar{\lambda})(1-\bar{\lambda})^3} \,,$$

which means that

$${\lambda_1}^k \geq \left( \frac{16\bar{\mathbf{s}}^2}{k(1-\bar{\lambda})^2} + \frac{2(2\,\bar{\mathbf{s}}+\delta)^2}{k(1+\bar{\lambda})(1-\bar{\lambda})^3} + \frac{{\lambda_1}^2}{2} \right)^{\frac{k}{2}} , \quad k \geq \bar{k}.$$

Then in a similar manner to that of Claim 3.2, if we define

$$c_{12} := \left( \frac{16\bar{\mathbf{s}}^2}{(1-\bar{\lambda})^2} + \frac{2(2\,\bar{\mathbf{s}}+\delta)^2}{(1+\bar{\lambda})(1-\bar{\lambda})^3} + 1 \right)^{\frac{\bar{k}}{2}} ;$$

it is easy to see that from (3.23), we get

$$\prod_{j=q}^{p-1} \bar{\alpha}_{k_l+2j} \leq c_{12}{\lambda_1}^{p-q}, \quad k_l \leq k_l + 2q < k_l + 2p \leq k_{l+1}.$$

Then by the definition of $\lambda_1$, we obtain

$$\prod_{j=q}^{p-1} c_5 \bar{\alpha}_{k_l+2j} \leq c_{12}{\lambda_1}^{p-q}{c_5}^{p-q}.$$
$$= c_{12}\lambda^{2(p-q)}, \quad k_l \leq k_l + 2q < k_l + 2p \leq k_{l+1}.$$

■

Before proceeding, observe from the definition of $\bar{\alpha}_{k_l+2j}$, that if $k_{l+1} - k_l$ is an odd number, then we would solve (3.21) and obtain a bound which is valid on $t = k_l, \dots, k_{l+1}-1$ and not on $t = k_{l+1}$; when $k_{l+1} - k_l$ is an even number, we would be able to obtain a bound on $t = k_l, \dots, k_{l+1}$. So in any case, we now proceed to solve (3.21) iteratively and apply the bound in (3.22). Using a similar analysis to that of Case 1 and defining

$$\bar{\gamma}_2 := \frac{1}{\lambda^2} \max\{c_{12}, c_3, c_{12}c_3, c_{11}c_{12}c_3, c_{12}c_{11}\},$$

we obtain

$$|x(t)| \le \bar{\gamma}_2 \lambda^{t-k_l} |x(k_l)| + \sum_{j=k_l}^{t-1} \bar{\gamma}_2 \lambda^{t-j-1} \big(|r(j+1)| + |n(j)|\big), \quad t \in [k_l, k_{l+1}). \tag{3.24}$$

Note that (3.24) does not apply for $t = k_{l+1}$; so to conclude Sub-Case 2.2, define

$$\bar{\gamma}_3 := c_3 \max\left\{1, \frac{\bar{\gamma}_2}{\lambda}\right\},$$

and utilizing (3.3) to obtain a bound accounting for the extra step yields

$$|x(t)| \le \bar{\gamma}_3 \lambda^{t-k_l} |x(k_l)| + \sum_{j=k_l}^{t-1} \bar{\gamma}_3 \lambda^{t-j-1} \big(|r(j+1)| + |n(j)|\big), \quad t \in [k_l, k_{l+1}]. \tag{3.25}$$

Finally, we will combine the results of Sub-Case 2.1 and Sub-Case 2.2 to find a general bound on $x$. Before proceeding, define

$$\bar{\gamma}_4 := \max\{\bar{\gamma}_3, c_9, \bar{\gamma}_3 c_9\}.$$

**Claim 3.4.** *The following bound holds:*

$$|x(t)| \le \bar{\gamma}_4 \lambda^{t-t_0} |x_0| + \sum_{j=t_0}^{t-1} \bar{\gamma}_4 \lambda^{t-j-1} \big(|r(j+1)| + |n(j)|\big), \quad t \ge t_0. \tag{3.26}$$

*Proof of Claim 3.4.* If $[k_0, k_l) = [t_0, k_l) \subset S_{\text{good}}$, then (3.26) is true for $t \in [k_0, k_l]$ by (3.25). If $[k_0, k_l) \subset S_{\text{bad}}$, then from (3.14) we have

$$|x(j)| \le \begin{cases} |x(k_0)| = |x_0|, & j = k_0, \\ c_9 |n(j-1)|, & j = k_0+1, k_0+2, \dots, k_l, \end{cases}$$

which means that (3.26) holds for $t \in [k_0, k_l]$ for this case as well.

We now proceed by induction: suppose that (3.26) is true for $t \in [k_0, k_l]$; we need to prove that it is true for $t \in (k_l, k_{l+1}]$. If $[k_l, k_{l+1}) \subset S_{\text{bad}}$, then from (3.14) we see that

$$|x(j)| \leq c_9 |n(j-1)|, \quad j = k_l + 1, k_l + 2, \ldots, k_{l+1},$$

so (3.26) clearly holds on $(k_l, k_{l+1}]$. On the other hand, if $[k_l, k_{l+1}) \subset S_{\text{good}}$, then $k_l - 1 \in S_{\text{bad}}$; from (3.14) we have

$$|x(k_l)| \leq c_9 |n(k_l - 1)|.$$

Using (3.25) to analyze the behaviour on $t \in [k_l, k_{l+1}]$, we have

$$\begin{aligned} |x(t)| &\leq \bar{\gamma}_3 \lambda^{t-k_l} |x(k_l)| + \sum_{j=k_l}^{t-1} \bar{\gamma}_3 \lambda^{t-j-1} \big( |r(j+1)| + |n(j)| \big) \\ &\leq c_9 \bar{\gamma}_3 \lambda^{t-k_l} |n(k_l - 1)| + \sum_{j=k_l}^{t-1} \bar{\gamma}_3 \lambda^{t-j-1} \big( |r(j+1)| + |n(j)| \big) \\ &\leq \sum_{j=k_l-1}^{t-1} \bar{\gamma}_4 \lambda^{t-j-1} \big( |r(j+1)| + |n(j)| \big), \end{aligned}$$

which implies that (3.26) holds. ■

At this point we have bounds on $x$ for both cases with noise and without. To combine the bounds (3.13) and (3.26), define

$$\bar{\gamma}_5 := \max\{\bar{\gamma}_1, \bar{\gamma}_4\}.$$

Then the overall bound is given by

$$|x(t)| \leq \bar{\gamma}_5 \lambda^{t-t_0} |x_0| + \sum_{j=t_0}^{t-1} \bar{\gamma}_5 \lambda^{t-j-1} \big( |r(j+1)| + |n(j)| \big), \quad t \geq t_0. \tag{3.27}$$

To conclude the proof of Theorem 3.1, we need a bound on $u$: using (2.22) we obtain

$$|u(t)| \leq \bar{f} |x(t)| + \bar{g} |r(t+1)|, \quad t \geq t_0;$$

so by substituting (3.27) into the above, we get

$$|u(t)| \leq \bar{f}\bar{\gamma}_5\lambda^{t-t_0}|x_0| + \sum_{j=t_0}^{t-1} \bar{f}\bar{\gamma}_5\lambda^{t-j-1}|n(j)| + \sum_{j=t_0}^{t} \left(\frac{\bar{f}\bar{\gamma}_5}{\lambda} + \bar{g}\right)\lambda^{t-j}|r(j+1)|, \quad t \geq t_0. \tag{3.28}$$

By combining (3.27) and (3.28), and defining

$$\bar{\gamma} := \max\left\{\bar{\gamma}_5, \bar{f}\bar{\gamma}_5, \frac{\bar{f}\bar{\gamma}_5}{\lambda} + \bar{g}\right\},$$

we conclude the proof. ■

Next, building on the observation in Remark 2.2, we show that the closed-loop system enjoys the homogeneity property [21]. Towards this end, with $\gamma \in \mathbb{R}$, $\gamma \neq 0$, let the scaled system state $x^\gamma$ evolve according to

$$x^\gamma(t+1) = a\,x^\gamma(t) + b\,u^\gamma(t) + \gamma\,n(t), \qquad x^\gamma(t_0) = \gamma\,x_0. \tag{3.29}$$

From (2.12), (2.13), with $\hat{\theta}_i^\gamma(t_0) \in \mathcal{S}_i$, $i \in \{1, 2\}$, we have

$$\check{\theta}_i^\gamma(t+1) := \hat{\theta}_i^\gamma(t) + \rho_i(t)\frac{\phi^\gamma(t)}{||\phi^\gamma(t)||^2}\left(\phi^\gamma(t)^\top(\theta^* - \hat{\theta}_i^\gamma(t)) + \gamma\,n(t)\right), \tag{3.30}$$

$$\hat{\theta}_i^\gamma(t+1) := \text{Proj}_{\mathcal{S}_i}\left\{\check{\theta}_i^\gamma(t+1)\right\}, \tag{3.31}$$

where the scaled quantity $\phi^\gamma$ is defined as

$$\phi^\gamma(t) := \begin{bmatrix} x^\gamma(t) \\ u^\gamma(t) \end{bmatrix},$$

with

$$\hat{\theta}_{\sigma(t)}^\gamma(t) =: \begin{bmatrix} \hat{a}_{\sigma(t)}^\gamma(t) \\ \hat{b}_{\sigma(t)}^\gamma(t) \end{bmatrix}$$

and control input $u^\gamma$ given by

$$u^\gamma(t) = \frac{1}{\hat{b}_{\sigma(t)}^\gamma(t)}\left(\gamma\,r(t+1) - \hat{a}_{\sigma(t)}^\gamma(t)\,x^\gamma(t)\right). \tag{3.32}$$

**Proposition 3.1.** *Suppose* $t_0 \in \mathbb{Z}$, $x_0 \in \mathbb{R}$, $\theta^* \in \mathcal{S}$, $\sigma_0 \in \{1,2\}$, $\hat{\theta}_i^{\gamma}(t_0) = \hat{\theta}_i(t_0) \in \mathcal{S}_i$, $i \in \{1,2\}$, $r$, $n \in l_\infty$, *and* $\gamma \in \mathbb{R}$, $\gamma \neq 0$. *If the combination of initial condition, reference, and noise* $(\phi_0, r, n)$ *yields system response* $\phi(t)$ *for (2.1),(2.11), (2.12), (2.13), (2.22), (2.26), and (2.27), and the combination* $(\gamma\,\phi_0, \gamma\, r, \gamma\, n)$ *yields system response* $\phi^\gamma$ *for (2.26), (2.27), (2.11), (3.29), (3.30), (3.31) and (3.32), then for every* $t \geq t_0$

$$\phi^\gamma(t) = \gamma\,\phi(t); \tag{3.33}$$

*and*

$$\hat{\theta}_i^{\gamma}(t) = \hat{\theta}_i(t). \tag{3.34}$$

*Proof.* Let $t_0 \in \mathbb{Z}$, $x_0 \in \mathbb{R}$, $\theta^* \in \mathcal{S}$, $r$, $n \in l_\infty$, $\gamma \in \mathbb{R}$, $\gamma \neq 0$ be arbitrary and fix $\sigma_0 \in \{1,2\}$, $\hat{\theta}_i^{\gamma}(t_0) = \hat{\theta}_i(t_0) \in \mathcal{S}_i$, $i \in \{1,2\}$ . We proceed by induction.

For the base case, at time step $t_0$, from (2.8) and (2.12)

$$\check{\theta}_i(t_0+1) = \hat{\theta}_i(t_0) + \rho_i(t_0)\frac{\phi(t_0)}{||\phi(t_0)||^2}\big(\phi(t_0)^\top(\theta^* - \hat{\theta}_i(t_0)) + n(t_0)\big), \tag{3.35}$$

and from (3.30), we get

$$\check{\theta}_i^{\gamma}(t_0+1) = \hat{\theta}_i^{\gamma}(t_0) + \rho_i(t_0)\frac{\phi(t_0)}{||\phi(t_0)||^2}\big(\phi(t_0)^\top(\theta^* - \hat{\theta}_i^{\gamma}(t_0)) + n(t_0)\big).$$

Using the fact that $\hat{\theta}_i^{\gamma}(t_0) = \hat{\theta}_i(t_0)$, the above implies

$$\check{\theta}_i^{\gamma}(t_0+1) = \check{\theta}_i(t_0+1),$$

which implies that

$$\hat{\theta}_i^{\gamma}(t_0+1) = \hat{\theta}_i(t_0+1),$$

as required.

Now, from (3.32), we have

$$u^\gamma(t_0) = \frac{\gamma}{\hat{b}^\gamma_{\sigma(t_0)}(t_0)} \big(r(t_0+1) - \hat{a}^\gamma_{\sigma(t_0)}(t_0)\, x(t_0)\big),$$

and so using the fact that $\hat{\theta}^\gamma_i(t_0) = \hat{\theta}_i(t_0)$, we get $u^\gamma(t_0) = \gamma\, u(t_0)$. From (3.29) we have

$$x^\gamma(t_0+1) = a\,\gamma\, x(t_0) + b\, u^\gamma(t_0) + \gamma\, n(t_0),$$

which gives $x^\gamma(t_0+1) = \gamma\, x(t_0+1)$. As a result, again using (3.32) and the previously established fact that $\hat{\theta}^\gamma_i(t_0+1) = \hat{\theta}_i(t_0+1)$, we conclude that $u^\gamma(t_0+1) = \gamma\, u(t_0+1)$ which shows that

$$\phi^\gamma(t_0+1) = \gamma\, \phi(t_0+1),$$

as required.

For the induction step, assume that for some $t > t_0$, $t \in \mathbb{Z}$, it is true that

$$\hat{\theta}^\gamma_i(t) = \hat{\theta}_i(t), \tag{3.36}$$
$$\phi^\gamma(t) = \gamma\, \phi(t). \tag{3.37}$$

From (3.30), we obtain

$$\check{\theta}^\gamma_i(t+1) = \hat{\theta}^\gamma_i(t) + \rho_i(t) \frac{\phi(t)}{||\phi(t)||^2} \big(\phi(t)^\top (\theta^* - \hat{\theta}^\gamma_i(t)) + n(t)\big).$$

Using the induction hypotheses (3.36) and (3.37), we get

$$\check{\theta}^\gamma_i(t+1) = \check{\theta}_i(t+1),$$

which, implies that

$$\hat{\theta}^\gamma_i(t+1) = \hat{\theta}_i(t+1),$$

as required.

Now, from (3.32), we get

$$u^\gamma(t) = \frac{\gamma}{\hat{b}^\gamma_{\sigma(t)}(t)}\big(r(t+1) - \hat{a}^\gamma_{\sigma(t)}(t)\,x(t)\big),$$

which, using the induction hypotheses (3.36) and (3.37), gives

$$u^\gamma(t) = \gamma\,u(t).$$

From (3.29)

$$x^\gamma(t+1) = a\,\gamma\,x(t) + b\,u^\gamma(t) + \gamma\,n(t),$$

which yields

$$x^\gamma(t+1) = \gamma\,x(t+1).$$

Finally, again using (3.32) and the previously established fact that $\hat{\theta}^\gamma_i(t+1) = \hat{\theta}_i(t+1)$, we conclude that $u^\gamma(t+1) = \gamma\,u(t+1)$ so that

$$\phi^\gamma(t+1) = \gamma\,\phi(t+1).$$

as required. ∎

## 3.2 Tolerance to Time-Variations and Unmodelled Dynamics

In this section, we will show that the convolution bound and exponential stability proven in Theorem 3.1 guarantees robustness to a degree of time-variations and unmodelled dynamics. The proof of this claim uses a general result from [36] applied to our specific setup. In order to summarize the results from [36], we introduce a multi-input multi-output plant with finite memory and some noise (disturbance) $n(t) \in \mathbb{R}^r$. To this end, with an input $u(t) \in \mathbb{R}^m$, an output $x(t) \in \mathbb{R}^r$, a modeling parameter of

$$\theta^* \in \tilde{\mathcal{S}} \subset \mathbb{R}^{p\times r},$$

and a vector of input-output data of the form

$$\phi(t) = \begin{bmatrix} x(t) \\ x(t-1) \\ \vdots \\ x(t-n_y+1) \\ u(t) \\ u(t-1) \\ \vdots \\ u(t-n_u+1) \end{bmatrix} \in \mathbb{R}^{n_y \cdot r + n_u \cdot m},$$

the nominal plant is given by

$$x(t+1) = \theta^{*\top} f\big(\phi(t)\big) + n(t), \quad t \geq t_0, \quad \phi(t_0) = \phi_0\,; \tag{3.38}$$

where $f : \mathbb{R}^{n_y \cdot r + n_u \cdot m} \to \mathbb{R}^p$ has a bounded gain and $\tilde{\mathcal{S}}$ is a bounded set. We represent this system by the pair $(f, \tilde{\mathcal{S}})$.

The general results in [36] considered a general dynamic controller with its state partitioned into two parts:

- $z(t) \in \mathbb{R}^{l_1}$ and
- $\hat{\theta}(t) \in \mathbb{R}^{l_2}$,

an exogenous signal $w(t) \in \mathbb{R}^r$, and with equations of the form

$$z(t+1) = g_1\big(z(t), \hat{\theta}(t), \phi(t), w(t), t, t_0\big), \quad z(t_0) = z_0 \tag{3.39}$$

$$\hat{\theta}(t+1) = g_2\big(z(t), \hat{\theta}(t), \phi(t), w(t), t, t_0\big), \quad \hat{\theta}(t_0) = \hat{\theta}_0 \tag{3.40}$$

$$u(t) = h\big(z(t), \hat{\theta}(t), \phi(t-1), x(t), w(t), t, t_0\big). \tag{3.41}$$

With $\Omega \subset \mathbb{R}^{l_2}$ a bounded set, in [36] it is assumed that

$$g_2 : \mathbb{R}^{l_1} \times \Omega \times \mathbb{R}^{n_y \cdot r + n_u \cdot m} \times \mathbb{R}^r \times \mathbb{Z} \times \mathbb{Z} \longrightarrow \Omega\,,$$

i.e., if $\hat{\theta}(t) \in \Omega$, then $\hat{\theta}(t+1) \in \Omega$.

We now define the exponential stability and convolution bound property.

**Definition 3.1.** *We say that the controller (3.39)-(3.41) provides exponential stability and a convolution bound for* $(f, \tilde{\mathcal{S}})$ *with gain* $c \geq 1$ *and decay rate* $\lambda \in (0,1)$ *if, for every* $\theta^* \in \tilde{\mathcal{S}}$, $t_0 \in \mathbb{Z}$, $\phi_0 \in \mathbb{R}^{n_y \cdot r + n_u \cdot m}$, $z_0 \in \mathbb{R}^{l_1}$, $\hat{\theta}_0 \in \Omega \subset \mathbb{R}^{l_2}$, *and* $n$, $w \in \mathbb{S}(\mathbb{R}^r)$, *when the controller (3.39)-(3.41) is applied to (3.38), the following holds:*

$$\left\| \begin{bmatrix} \phi(t) \\ z(t) \end{bmatrix} \right\| \leq c\lambda^{t-\tau} \left\| \begin{bmatrix} \phi(\tau) \\ z(\tau) \end{bmatrix} \right\| + \sum_{j=\tau}^{t-1} c\lambda^{t-j-1} \big( \|w(j)\| + \|n(j)\| \big) + c\,\|w(t)\|, \quad t \geq \tau \geq t_0. \tag{3.42}$$

Next, consider a plant with possibly time-varying parameter $\theta^*(t) \in \tilde{\mathcal{S}} \subset \mathbb{R}^{p \times r}$:

$$x(t+1) = \theta^*(t)^\top f\big(\phi(t)\big) + n(t), \quad \phi(t_0) = \phi_0. \tag{3.43}$$

Using the standard time-variation model commonly adopted in adaptive control setting – see [12] – with $c_0 \geq 0$ and $\epsilon > 0$, let $s(\tilde{\mathcal{S}}, c_0, \epsilon)$ denote the subset of $l_\infty(\mathbb{R}^{p \times r})$ whose elements $\theta^*$ satisfy the following

- $\theta^*(t) \in \tilde{\mathcal{S}}$ for every $t \in \mathbb{Z}$,
- and

$$\sum_{t=t_1}^{t_2-1} ||\theta^*(t+1) - \theta^*(t)|| \leq c_0 + \epsilon\,(t_2 - t_1), \quad t_2 > t_1,\ t_1 \in \mathbb{Z}.$$

The above time-variations cover both slow time variations and/or occasional jumps. We represent the time-varying system by the pair $(f, s(\tilde{\mathcal{S}}, c_0, \epsilon))$. Definition 3.1 extends in a natural way to handle time-variations.

**Definition 3.2.** *We say that the controller (3.39)-(3.41) provides exponential stability and a convolution bound for* $\big(f, s(\tilde{\mathcal{S}}, c_0, \epsilon)\big)$ *with gain* $c \geq 1$ *and decay rate* $\lambda \in (0,1)$ *if, for every* $\theta^* \in s(\tilde{\mathcal{S}}, c_0, \epsilon)$, $t_0 \in \mathbb{Z}$, $\phi_0 \in \mathbb{R}^{n_y \cdot r + n_u \cdot m}$, $z_0 \in \mathbb{R}^{l_1}$, $\hat{\theta}_0 \in \Omega \subset \mathbb{R}^{l_2}$, *and* $n$, $w \in \mathbb{S}(\mathbb{R}^r)$, *when the controller (3.39)-(3.41) is applied to (3.38), the following holds:*

$$\left\| \begin{bmatrix} \phi(t) \\ z(t) \end{bmatrix} \right\| \leq c\lambda^{t-\tau} \left\| \begin{bmatrix} \phi(\tau) \\ z(\tau) \end{bmatrix} \right\| + \sum_{j=\tau}^{t-1} c\lambda^{t-j-1} \big( \|w(j)\| + \|n(j)\| \big) + c\,\|w(t)\|, \quad t \geq \tau \geq t_0. \tag{3.44}$$

The following theorem shows that if the controller (3.39)-(3.41) provides exponential stability and a convolution bound for the plant (3.38), then the same will be true for the time-varying plant (3.43), as long as $\epsilon$ is small enough.

**Theorem 3.2** ([36, Theorem 1])**.** *Suppose that the controller (3.39)-(3.41) provides exponential stability and a convolution bound for* $(f, \tilde{\mathcal{S}})$ *with gain* $c \geq 1$ *and decay rate* $\lambda \in (0,1)$*. Then for every* $\lambda_1 \in (\lambda, 1)$ *and* $c_0 > 0$*, there exist* $c_1 \geq c$ *and* $\epsilon > 0$ *so that the controller (3.39)-(3.41) provides exponential stability and a convolution bound for* $\big(f, s(\tilde{\mathcal{S}}, c_0, \epsilon)\big)$ *with gain* $c_1$ *and decay rate* $\lambda_1$*.*

We now consider the time-varying plant (3.43) with unmodelled dynamics $d_\Delta(t) \in \mathbb{R}^r$:

$$x(t+1) = \theta^*(t)^\top f\big(\phi(t)\big) + n(t) + d_\Delta(t), \quad \phi(t_0) = \phi_0. \tag{3.45}$$

We also adopt a common model of unmodelled dynamics; with $g : \mathbb{R}^{n_y \cdot r + n_u \cdot m} \to \mathbb{R}$ a map with a bounded gain, $\beta \in (0,1)$ and $\mu > 0$, we consider

$$m(t+1) = \beta\, m(t) + \beta\, \big|g\big(\phi(t)\big)\big|, \quad m(t_0) = m_0 \tag{3.46}$$

$$\big\|d_\Delta(t)\big\| \leq \mu\, m(t) + \mu\, \big|g\big(\phi(t)\big)\big|, \quad t \geq t_0. \tag{3.47}$$

It turns out that the model (3.46), (3.47) encapsulates classical additive, multiplicative and uncertainty in a coprime factorization – see [21] for detailed discussion.

The following result proves that if the controller (3.39)-(3.41) provides exponential stability and a convolution bound for the plant (3.43) with $s(\tilde{\mathcal{S}}, c_0, \epsilon)$, then the closed-loop system enjoys a degree of tolerance to unmodelled dynamics.

**Theorem 3.3** ([36, Theorem 2]). *Suppose that the controller (3.39)-(3.41) provides exponential stability and a convolution bound for* $\big(f, s(\tilde{\mathcal{S}}, c_0, \epsilon)\big)$ *with a gain* $c_1$ *and decay rate* $\lambda_1 \in (0,1)$. *Then for every* $\beta \in (0,1)$ *and* $\lambda_2 \in \big(\max\{\lambda_1, \beta\}, 1\big)$, *there exist* $\bar{\mu} > 0$ *and* $c_2 > 0$ *so that for every* $\theta^* \in s(\tilde{\mathcal{S}}, c_0, \epsilon)$, $\mu \in (0, \bar{\mu})$, $t_0 \in \mathbb{Z}$, $\phi_0 \in \mathbb{R}^{n_y \cdot r + n_u \cdot m}$, $z_0 \in \mathbb{R}^{l_1}$, $\hat{\theta}_0 \in \Omega \subset \mathbb{R}^{l_2}$ *and* $n$, $w \in \mathbb{S}(\mathbb{R}^r)$, *when the controller (3.39)-(3.41) is applied to the plant (3.45) with* $d_\Delta$ *satisfying (3.46)-(3.47), the following holds:*

$$\left\| \begin{bmatrix} \phi(t) \\ z(t) \\ m(t) \end{bmatrix} \right\| \leq c_2 \, {\lambda_2}^{t-t_0} \left\| \begin{bmatrix} \phi_0 \\ z_0 \\ m_0 \end{bmatrix} \right\| + \sum_{j=t_0}^{t-1} c_2 \, {\lambda_2}^{t-j-1} \big( \|w(j)\| + \|n(j)\| \big) + c_2 \, ||w(t)||, \quad t \geq t_0.$$

Finally, we now show that the adaptive controller (2.22),(2.26) and (2.27), presented in Chapter 2, fits into the paradigm of the controller (3.39)-(3.41) and that our closed-loop system enjoys the aforementioned tolerance to time-variations and unmodelled dynamics.

**Lemma 3.1.** *The adaptive controller consisting of the parameter estimator (2.11), (2.12), (2.13), the control law (2.22), and the switching rule (2.26), (2.27) provides exponential stability and a convolution bound for* $(I_2, \mathcal{S})$ *with gain* $\gamma \geq 1$ *and decay rate* $\lambda \in (0,1)$.

*Proof.* We start by showing that our adaptive controller is an instance of the general control law (3.39)-(3.41). To do so we make the following identifications:

$$\begin{aligned}
\Omega &:= \mathcal{S}_1 \times \mathcal{S}_2 \times \{1,2\} \times \left[0, \frac{2\,\bar{\mathbf{s}}+\delta}{1-\bar{\lambda}}\right] \times \left[0, \frac{2\,\bar{\mathbf{s}}+\delta}{1-\bar{\lambda}}\right], \\
z(t) &:= \varnothing\,, \\
\hat{\theta}(t) &:= \begin{bmatrix} \hat{\theta}_1(t) \\ \hat{\theta}_2(t) \\ \sigma(t) \\ J_1(t) \\ J_2(t) \end{bmatrix}, \\
w(t) &:= r(t+1)\,,
\end{aligned}$$

and also identify the control law (3.41) with (2.22). With these identifications, we must show that if $\hat{\theta}(t) \in \Omega$, then $\hat{\theta}(t+1) \in \Omega$. It suffices to prove that if $J_i(t) \in \left[0, \frac{2\,\bar{\mathbf{s}}+\delta}{1-\bar{\lambda}}\right]$, then $J_i(t+1) \in \left[0, \frac{2\,\bar{\mathbf{s}}+\delta}{1-\bar{\lambda}}\right]$.

Suppose that for some $i \in \{1,2\}$ and $t \in \mathbb{Z}$, $J_i(t) \in \left[0, \frac{2\,\bar{\mathbf{s}}+\delta}{1-\bar{\lambda}}\right]$. Then from (2.26) we have that

$$\begin{aligned}
J_i(t+1) &= \bar{\lambda}\, J_i(t) + \rho_i(t)\, \frac{|e_i(t+1)|}{||\phi(t)||} \\
&\leq \bar{\lambda}\, \frac{2\,\bar{\mathbf{s}}+\delta}{1-\bar{\lambda}} + 2\,\bar{\mathbf{s}} + \delta \quad \text{(by definition (2.11) of } \rho_i) \\
&= \bar{\lambda}\, \frac{2\,\bar{\mathbf{s}}+\delta}{1-\bar{\lambda}} + \frac{(2\,\bar{\mathbf{s}}+\delta)\,(1-\bar{\lambda})}{1-\bar{\lambda}} \\
&= \frac{2\,\bar{\mathbf{s}}+\delta}{1-\bar{\lambda}}.
\end{aligned}$$

Therefore $J_i(t+1) \in \left[0, \frac{2\,\bar{\mathbf{s}}+\delta}{1-\bar{\lambda}}\right]$ as required. Since $i$ was arbitrary, this shows that $\hat{\theta}(t+1) \in \Omega$ as claimed. The result now follows from Theorem 3.1. ■

We now consider the plant (2.1) subject to parameter variations $\theta^*(t) \in \mathcal{S}$ and with unmodelled dynamics $d_\Delta(t) \in \mathbb{R}$:

$$x(t+1) = \phi(t)^\top\, \theta^*(t) + n(t) + d_\Delta(t), \quad \phi(t_0) = \phi_0. \tag{3.48}$$

with $d_\Delta$ satisfying (3.46)-(3.47). By Lemma 3.1, Theorem 3.2 and Theorem 3.3, we conclude that the adaptive controller (3.39)-(3.41) provides exponential stability and a convolution bound for (3.48), i.e., our approach is robust.

# Chapter 4

# Tracking Problem

We now move from the stability problem to the much harder tracking problem. We remind the reader that with $\theta^* \in \mathcal{S}$, we let $g(\theta^*)$ denote the index $i \in \{1, 2\}$ for which $\theta^* \in \mathcal{S}_i$, see (2.2) for a precise definition. To minimize notation, when there is no risk of confusion, we will drop the argument and simply denote this index by $g$. Throughout this chapter we consider tracking when the noise is zero.

## 4.1 Preliminary Results

In this section we present three results which provide a building block for two key theorems. We start with a result on the performance signal and the prediction error for the correct estimator.

**Proposition 4.1.** *Suppose that the adaptive controller consisting of the parameter estimator (2.11), (2.12), (2.13), the control law (2.22), and the switching rule (2.26), (2.27) is applied to the plant (2.1). For every* $\bar{\lambda} \in (0, 1)$, $t_0 \in \mathbb{Z}$, $x_0 \in \mathbb{R}$, $\theta^* \in \mathcal{S}$, $\sigma_0 \in \{1, 2\}$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$, $i \in \{1, 2\}$, *and* $r \in l_\infty$, *if* $n = 0$, *then the following limits hold:*

$$\lim_{t \to \infty} J_g(t + 1) = 0, \tag{4.1}$$

$$\lim_{t \to \infty} e_g(t + 1) = 0. \tag{4.2}$$

*Proof.* Fix $\bar{\lambda} \in (0,1)$, $t_0 \in \mathbb{Z}$, $x_0 \in \mathbb{R}$, $\theta^* \in \mathcal{S}$, $\sigma_0 \in \{1,2\}$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$, $i \in \{1,2\}$, and $r \in l_\infty$, and set $n = 0$. By Proposition 2.3 and equation (2.17) we have that

$$||\tilde{\theta}_g(t)||^2 \leq ||\tilde{\theta}_g(t_0)||^2 - \frac{1}{2}\sum_{j=t_0}^{t-1} \rho_g(j)\, \frac{|e_g(j+1)|^2}{||\phi(j)||^2}.$$

Using the fact $||\tilde{\theta}_g(t_0)|| \leq 2||\mathcal{S}_g||$, the above implies that

$$\sum_{j=t_0}^{t-1} \rho_g(j)\, \frac{|e_g(j+1)|^2}{||\phi(j)||^2} \leq 2||\tilde{\theta}_g(t_0)||^2 \leq 8||\mathcal{S}_g||^2 \leq 8\,\bar{\mathbf{s}}^2, \quad t > t_0. \tag{4.3}$$

Using the bound (4.3) and applying Parseval's theorem for discrete-time systems to the difference equation for $J_g(t)$ given in (2.26), we have

$$\sum_{j=t_0}^{\infty} J_g^2(j+1) \leq \frac{8\,\bar{\mathbf{s}}^2}{(1-\bar{\lambda})^2},$$

from which we conclude that

$$\lim_{j\to\infty} J_g(j+1) = 0.$$

For integers $j \geq t_0$, define

$$a_j := \begin{cases} \frac{|e_g(j+1)|}{||\phi(j)||}, & \text{if } \rho_g(j) = 1 \\ 0, & \text{otherwise.} \end{cases} \tag{4.4}$$

From (4.3), it follows that $a_j \to 0$ as $j \to \infty$. When $\rho_g(j) = 1$, from (4.4) we have

$$|e_g(j+1)| = a_j\,||\phi(j)||. \tag{4.5}$$

On the other hand, when $\rho_g(j) = 0$, by Proposition 2.2 we see that $||\phi(j)|| = 0$, which is equivalent to $x(j) = u(j) = 0$; from (2.8), we see that $e_g(j+1) = 0$, which means that (4.5) holds in this case as well. From Theorem 3.1 we have that $\phi$ is bounded, from which we deduce that

$$\lim_{j\to\infty} e_g(j+1) = 0.$$

■

Before proceeding further, we define the index $b$ by

$$b := \{1,2\} \setminus \{g\}.$$

We now present a useful proposition for the case when $\lim_{t\to\infty} J_b(t) = 0$.

**Proposition 4.2.** *Suppose that the adaptive controller consisting of the parameter estimator (2.11), (2.12), (2.13), the control law (2.22), and the switching rule (2.26), (2.27) is applied to the plant (2.1). For every* $\bar{\lambda} \in (0,1)$, $t_0 \in \mathbb{Z}$, $x_0 \in \mathbb{R}, \theta^* \in \mathcal{S}$, $\sigma_0 \in \{1,2\}$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$, $i \in \{1,2\}$, *and* $r \in l_\infty$, *if* $n = 0$ *and* $\lim_{t\to\infty} J_b(t) = 0$, *then*

$$\lim_{t\to\infty} \varepsilon(t) = 0. \tag{4.6}$$

*Proof.* Fix $\bar{\lambda} \in (0,1)$, $t_0 \in \mathbb{Z}$, $x_0 \in \mathbb{R}, \theta^* \in \mathcal{S}$, $\sigma_0 \in \{1,2\}$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$, $i \in \{1,2\}$, and $r \in l_\infty$. Suppose that $n = 0$ and that $\lim_{t\to\infty} J_b(t) = 0$. For $t \geq t_0$, define

$$a_t := \begin{cases} \frac{|e_b(t+1)|}{||\phi(t)||}, & \text{if } \rho_b(t) = 1 \\ 0, & \text{otherwise.} \end{cases} \tag{4.7}$$

From (2.26), we have

$$J_b(t+1) = \bar{\lambda}\, J_b(t) + a_t.$$

Since $\lim_{t\to\infty} J_b(t) = 0$, it follows that $a_t \to 0$ as $t \to \infty$. When $\rho_b(t) = 1$, we have

$$|e_b(t+1)| = a_t\, ||\phi(t)||. \tag{4.8}$$

On the other hand, when $\rho_b(t) = 0$, by Proposition 2.2 we see that $||\phi(t)|| = 0$, which is equivalent to $x(t) = u(t) = 0$; from (2.8), we see that $e_b(t+1) = 0$ as well, which means that (4.8) holds in this case as well. From Theorem 3.1 we have that $\phi$ is bounded, from which we deduce that

$$\lim_{t\to\infty} e_b(t+1) = 0.$$

From (2.24), for every $t \geq t_0$, either

$$\varepsilon(t+1) = e_g(t+1)$$

or

$$\varepsilon(t+1) = e_b(t+1).$$

Since from Proposition 4.1, we already know that $e_g(t+1) \to 0$ as $t \to \infty$, we conclude that

$$\lim_{t\to\infty} \varepsilon(t) = 0.$$

■

We now prove that the tracking error $\varepsilon$ goes to zero if switching stops.

**Proposition 4.3.** *Suppose that the adaptive controller consisting of the parameter estimator (2.11), (2.12), (2.13), the control law (2.22), and the switching rule (2.26), (2.27) is applied to the plant (2.1). For every $\bar{\lambda} \in (0,1)$, $t_0 \in \mathbb{Z}$, $x_0 \in \mathbb{R}, \theta^* \in \mathcal{S}$, $\sigma_0 \in \{1,2\}$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$, $i \in \{1,2\}$, and $r \in l_\infty$, if $n = 0$ and the switching signal $\sigma$ stops switching, then*

$$\lim_{t\to\infty} \varepsilon(t) = 0. \tag{4.9}$$

*Proof.* Fix $\bar{\lambda} \in (0,1)$, $t_0 \in \mathbb{Z}$, $x_0 \in \mathbb{R}, \theta^* \in \mathcal{S}$, $\sigma_0 \in \{1,2\}$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$, $i \in \{1,2\}$, and $r \in l_\infty$.

**Case 1:** $\sigma(t)$ stops at $g$.

In this case, there exists a $\bar{t} \geq t_0$ so that

$$\sigma(t) = g, \qquad t \geq \bar{t}.$$

Hence, (2.24) implies that

$$\varepsilon(t+1) = e_g(t+1), \qquad t \geq \bar{t}. \tag{4.10}$$

From Proposition 4.1, we conclude that $\varepsilon(t) \to 0$ as $t \to \infty$.

**Case 2:** $\sigma(t)$ stops at $b$.

In this case, there exists a $\bar{t} \geq t_0$ so that

$$\sigma(t) = b, \qquad t \geq \bar{t}.$$

This, in turn, implies that

$$J_b(t) \leq J_g(t), \qquad t \geq \bar{t}. \tag{4.11}$$

In this case, (2.24) implies that

$$\varepsilon(t+1) = e_b(t+1), \qquad t \geq \bar{t}. \tag{4.12}$$

Now define

$$a_t := \begin{cases} \frac{|e_b(t+1)|}{||\phi(t)||}, & \text{if } \rho_b(t) = 1 \\ 0, & \text{otherwise.} \end{cases} \tag{4.13}$$

From Proposition 4.1, we know that $J_g(t+1) \to 0$ as $t \to \infty$. Combining (4.11) and (4.13), we have that

$$J_b(t+1) = \bar{\lambda}\, J_b(t) + a_t \leq J_g(t+1), \qquad t \geq \bar{t},$$

which can only be true if $a_t \to 0$ as $t \to \infty$. Now we need to relate $a_t$ to $\varepsilon(t+1)$. If $\rho_b(t) = 1$, then using (4.12), we have

$$|\varepsilon(t+1)| = a_t\, ||\phi(t)||. \tag{4.14}$$

On the other hand, if $\rho_b(t) = 0$, then from Proposition 2.2, we see that $||\phi(t)|| = 0$, which is equivalent to $x(t) = u(t) = 0$, so that $x(t+1) = 0$; from the formula for $u(t)$ in (2.22), we see that $r(t+1) = 0$ as well, so that $\varepsilon(t+1) = 0$. This shows that regardless of the value of $\rho_b(t)$, the expression (4.14) holds. From Theorem 3.1 we have that $\phi$ is bounded, so we conclude that $\varepsilon(t) \to 0$ as $t \to \infty$. ■

While all of our simulations indicate that, in the absence of noise, switching stops and asymptotic tracking ensues, we have been unable to prove it. In the following two sections we will prove that this is the case in two circumstances.

## 4.2 Tracking With a Persistence of Excitation Assumption

To prove asymptotic convergence of tracking error, we will impose an assumption on the nature of the reference signal. A common concept used in the adaptive control literature is that of a persistently exciting signal, e.g., see [1] and [8]. In that work this concept is used to guarantee that a single estimator converges to the correct parameter. Here we will use this in a different way: to prove that if $\mathcal{S}_1$ and $\mathcal{S}_2$ are disjoint, then switching will stop and tracking will occur, we start with a classical definition, e.g., see [8].

**Definition 4.1.** *$r \in l_\infty$ is strongly persistently exciting (SPE) of order $p$ if there exists an $n \in \mathbb{N}$, $c_1 > 0$, $c_2 > 0$ such that*

$$c_1 I \le \sum_{t=k}^{k+n} \begin{bmatrix} r(t) \\ \vdots \\ r(t+p-1) \end{bmatrix} \begin{bmatrix} r(t) & \cdots & r(t+p-1) \end{bmatrix} \le c_2 I, \quad k \in \mathbb{Z}.$$

**Remark 4.1.** *If $p = 2$, then from Definition 4.1, we have*

$$\sum_{t=k}^{k+n} \begin{bmatrix} r(t) \\ r(t+1) \end{bmatrix} \begin{bmatrix} r(t) & r(t+1) \end{bmatrix} = \begin{bmatrix} r(k) & \cdots & r(k+n) \\ r(k+1) & \cdots & r(k+n+1) \end{bmatrix} \begin{bmatrix} r(k) & r(k+1) \\ \vdots & \vdots \\ r(k+n) & r(k+n+1) \end{bmatrix}.$$

In our situation we will need $r$ to be sufficiently exciting in the limit. This brings us to a non-standard definition:

**Definition 4.2.** *$r \in l_\infty$ is asymptotically strongly persistently exciting (ASPE) of order $p$ if there exists an $n \in \mathbb{N}$, $c_1 > 0$, $c_2 > 0$ and $T > t_0$ such that*

$$c_1 I \le \sum_{t=k}^{k+n} \begin{bmatrix} r(t) \\ \vdots \\ r(t+p-1) \end{bmatrix} \begin{bmatrix} r(t) & \cdots & r(t+p-1) \end{bmatrix} \le c_2 I, \quad k \ge T\,.$$

Now we prove asymptotic trajectory tracking for the case when the reference signal $r \in l_\infty$ is asymptotically strongly persistently exciting (ASPE) of order 2.

**Theorem 4.1.** *Assume that $\mathcal{S}_1 \cap \mathcal{S}_2 = \varnothing$ and suppose that the adaptive controller consisting of the parameter estimator (2.11), (2.12), (2.13), the control law (2.22), and the switching rule (2.26), (2.27) is applied to the plant (2.1). For every $\bar{\lambda} \in (0,1)$, $t_0 \in \mathbb{Z}$, $x_0 \in \mathbb{R}$, $\theta^* \in \mathcal{S}$, $\sigma_0 \in \{1,2\}$, and $\hat{\theta}_i(t_0) \in \mathcal{S}_i$, $i \in \{1,2\}$, if $n = 0$ and $r \in l_\infty$ is ASPE of order 2, then*

$$\lim_{t\to\infty} \varepsilon(t) = 0.$$

*Proof.* Fix $t_0 \in \mathbb{Z}$, $x_0 \in \mathbb{R}$, $\theta^* \in \mathcal{S}$, $\sigma_0 \in \{1,2\}$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$, $i \in \{1,2\}$, $r \in l_\infty$ and $\bar{\lambda} \in (0,1)$. Assume that $n = 0$ and that $r \in l_\infty$ is ASPE of order 2. Before proceeding, we define two sets of times according to the estimator chosen by the switching signal $\sigma$:

$$S_{\text{good}} = \{t \geq t_0 \,:\, \sigma(t) = g\}$$
$$S_{\text{bad}} = \{t \geq t_0 \,:\, \sigma(t) = b\}.$$

Clearly $\{t \in \mathbb{Z} : t \geq t_0\} = S_{\text{good}} \cup S_{\text{bad}}$.

Now we partition the time index $\{j \in \mathbb{Z} : j \geq t_0\}$ into intervals which oscillate between $S_{\text{good}}$ and $S_{\text{bad}}$. We can clearly define a (possibly infinite) sequence of intervals of the form $[k_l, k_{l+1})$, $l \in \mathbb{Z}^+$, which satisfy:

(i) without loss of generality, $k_0 = t_0$ serves as the initial instant of the first interval;

(ii) $[k_l, k_{l+1})$ either belongs to $S_{\text{good}}$ or $S_{\text{bad}}$; and

(iii) if $k_{l+1} \neq \infty$ and $[k_l, k_{l+1})$ belongs to $S_{\text{good}}$, then $[k_{l+1}, k_{l+2})$ belongs to $S_{\text{bad}}$ and *vice versa*.

Since $n = 0$ and $r$ is asymptotically strongly persistently exciting of order 2, for the case when $J_b(t) \to 0$ as $t \to \infty$, by Proposition 4.2 we have that $\lim_{t\to\infty} \varepsilon(t) = 0$. So for the rest of the proof we suppose that this is not the case, i.e., we have $\limsup_{t\to\infty} |J_b(t+1)| > 0$. We will show that eventually we will end up in $S_{\text{good}}$ and switching stops, so we can conclude from Proposition 4.3 that $\lim_{t\to\infty} \varepsilon(t) = 0$.

To proceed, suppose that we do not stop switching. Next, define

$$\bar{\delta} := \min \left\{ \frac{1}{2}, \frac{1}{2} \limsup_{t \to \infty} \rho_b(t) \frac{|e_b(t+1)|}{||\phi(t)||} \right\};$$

since we have assumed that $\limsup_{t\to\infty} |J_b(t+1)| > 0$, it must be that $\bar{\delta} > 0$ as well. Since $r$ is assumed to be ASPE of order 2, we know that there exist $m \in \mathbb{N}$, $\bar{c}_1 > 0$, $\bar{c}_2 > 0$, $T_0 > t_0$, such that

$$\bar{c}_1 I_2 \le \sum_{t=\underline{t}}^{\underline{t}+m} \begin{bmatrix} r(t) \\ r(t+1) \end{bmatrix} \begin{bmatrix} r(t) & r(t+1) \end{bmatrix} \le \bar{c}_2 I_2, \quad \underline{t} \ge T_0, \tag{4.15}$$

so fix such quantities. Now, from Proposition 4.1, for each $n \ge 2$, we can define $T_n \ge T_0$, such that

$$J_g(t+1) \le \bar{\delta}^{2n}, \quad t \ge T_n \tag{4.16}$$
$$|e_g(t+1)| \le \bar{\delta}^{2n}, \quad t \ge T_n\,; \tag{4.17}$$

without loss of generality $T_{n+1} > T_n$ for $n \ge 2$. From the definition of $\bar{\delta}$, we know that

$$\frac{1}{2} \limsup_{t\to\infty} \rho_b(t) \frac{|e_b(t+1)|}{||\phi(t)||} \ge \bar{\delta},$$

which means that there exists a $t_n \ge T_n$ so that

$$\rho_b(t_n) \frac{|e_b(t_n+1)|}{||\phi(t_n)||} \ge \bar{\delta}; \tag{4.18}$$

this, in turn, implies that

$$J_b(t_n+1) \ge \bar{\delta}$$

as well. Since $T_n$ is defined so that both (4.16) and (4.17) holds, this means that

$$J_g(t_n+1) \le \bar{\delta}^{2n} < \bar{\delta},$$

so $\sigma(t_n+1) = g$. Since we have assumed that we switch forever, there exists a $\bar{t}_n > t_n + 1$ so that

$$\sigma(t) = g, \quad t \in [t_n+1, \bar{t}_n),$$

but

$$\sigma(\bar{t}_n) = b,$$

i.e., $[t_n+1,\bar{t}_n) \subset S_{\text{good}}$. But in order to switch from $S_{\text{good}}$ to $S_{\text{bad}}$ at $t=\bar{t}_n$, we need

$$J_b(\bar{t}_n) \le J_g(\bar{t}_n) \le \bar{\delta}^{2n}\,;$$

this means that at some time $t \in (t_n+1,\bar{t}_n]$, we have

$$\rho_b(t-1)\,\frac{|e_b(t)|}{||\phi(t-1)||} \le \bar{\delta}^{2n}.$$

Indeed, if we combine this with (4.18) we see that there must exist a $k \in [t_n+1,\bar{t}_n)$ so that

$$\rho_b(t-1)\,\frac{|e_b(t)|}{||\phi(t-1)||} \le \bar{\delta}^{n}, \quad t \in (k,\bar{t}_n)\,; \tag{4.19}$$

let $\underline{t}_n$ denote the smallest such $k$. This implies, in particular, that

$$J_b(\underline{t}_n) \ge \rho_b(\underline{t}_n-1)\,\frac{|e_b(\underline{t}_n)|}{||\phi(\underline{t}_n-1)||} \ge \bar{\delta}^{n}. \tag{4.20}$$

**Claim 4.1.** *$\bar{t}_n - \underline{t}_n$ is **large** if $n$ is large, in the sense that*

$$\bar{t}_n - \underline{t}_n \ge n\,\frac{\ln\bar{\delta}}{\ln\bar{\lambda}}. \tag{4.21}$$

*Proof of Claim 4.1.* With $[t_n+1,\bar{t}_n) \subset S_{\text{good}}$, to switch from $S_{\text{good}}$ to $S_{\text{bad}}$ at $t=\bar{t}_n$, we need

$$J_b(\bar{t}_n) \le J_g(\bar{t}_n) \le \bar{\delta}^{2n}.$$

But from (4.20)

$$J_b(\underline{t}_n) \ge \bar{\delta}^{n}\,,$$

so by solving (2.26) recursively, we have

$$J_b(\bar{t}_n) \ge \bar{\lambda}^{(\bar{t}_n-\underline{t}_n)}\,\bar{\delta}^{n}.$$

So we need to have

$$\bar{\lambda}^{(\bar{t}_n-\underline{t}_n)}\bar{\delta}^{n} \le \bar{\delta}^{2n}$$
$$\Leftrightarrow \bar{\lambda}^{(\bar{t}_n-\underline{t}_n)} \le \bar{\delta}^{n}$$
$$\Leftrightarrow (\bar{t}_n-\underline{t}_n)\,\ln\bar{\lambda} \le n\,\ln\bar{\delta}$$

Since $\bar{\lambda} \in (0,1)$, we have

$$\bar{t}_n - \underline{t}_n \ge n\,\frac{\ln\bar{\delta}}{\ln\bar{\lambda}} > 0.$$

■

Now choose $N \geq 2$ satisfying

$$N \frac{\ln \bar{\delta}}{\ln \bar{\lambda}} > m + 1,$$

and restrict $n \geq N$. To proceed, we now analyze $e_b$; from (2.8), we have

$$e_b(t+1) = \phi(t)^\top \left(\theta^* - \hat{\theta}_b(t)\right).$$

If we define $\tilde{\theta}_b(t) := \hat{\theta}_b(t) - \theta^*$, then the above can be rewritten as

$$e_b(t+1) = -\phi(t)^\top \tilde{\theta}_b(t). \tag{4.22}$$

We will now analyze the above equation to prove that we will obtain a contradiction to the claim that the scaled version of $e_b(t)$ will be small on the interval $[\underline{t}_n + 1, \underline{t}_n + m + 1]$, at least if $n$ is large enough. The key to proving this are two insights:

(i) $\tilde{\theta}_b(t)$ will be moving very slowly on this interval, and

(ii) $\phi(t)$ will be approximately equal to a scaled version of $\begin{bmatrix} r(t) \\ r(t+1) \end{bmatrix}$ on this interval.

To proceed, we first examine $\phi(t)$ in detail. Using (2.1) and (2.23), we have

$$\phi(t) = \begin{bmatrix} r(t) \\ \frac{1}{b}\left(r(t+1) - a\, r(t)\right) \end{bmatrix} + \begin{bmatrix} \varepsilon(t) \\ \frac{1}{b}\left(\varepsilon(t+1) - a\, \varepsilon(t)\right) \end{bmatrix};$$

define

$$\bar{r}(t) := \begin{bmatrix} r(t) \\ r(t+1) \end{bmatrix},$$
$$\bar{\varepsilon}(t) := \begin{bmatrix} \varepsilon(t) \\ \varepsilon(t+1) \end{bmatrix},$$
$$\bar{A} := \begin{bmatrix} 1 & 0 \\ -\frac{a}{b} & \frac{1}{b} \end{bmatrix},$$

so the above can be rewritten as

$$\phi(t) = \bar{A}\, \bar{r}(t) + \bar{A}\, \bar{\varepsilon}(t). \tag{4.23}$$

Substituting (4.23) into (4.22), we obtain

$$\begin{aligned} e_b(t+1) &= -\bar{r}(t)^\top \bar{A}^\top \tilde{\theta}_b(t) - \bar{\varepsilon}(t)^\top \bar{A}^\top \tilde{\theta}_b(t) \\ \Rightarrow e_b(t+1) + \bar{\varepsilon}(t)^\top \bar{A}^\top \tilde{\theta}_b(t) &= -\bar{r}(t)^\top \bar{A}^\top \tilde{\theta}_b(t). \end{aligned} \tag{4.24}$$

From Proposition 2.3 and equation (4.19) with $k = \underline{t}_n$, we have that

$$||\hat{\theta}_b(t) - \hat{\theta}_b(t-1)|| \leq \bar{\delta}^n, \quad t \in [\underline{t}_n + 1, \bar{t}_n).$$

So

$$||\tilde{\theta}_b(t) - \tilde{\theta}_b(\underline{t}_n + 1)|| \leq m\,\bar{\delta}^n, \quad t \in [\underline{t}_n + 1, \underline{t}_n + m + 1]. \tag{4.25}$$

Now rewrite equation (4.24) as

$$\underbrace{-\Big(e_b(t+1) + \bar{\varepsilon}(t)^\top \bar{A}^\top \tilde{\theta}_b(t) + \bar{r}(t)^\top \bar{A}^\top \big(\tilde{\theta}_b(t) - \tilde{\theta}_b(\underline{t}_n + 1)\big)\Big)}_{=:\,\nu(t)} = \bar{r}(t)^\top \bar{A}^\top \tilde{\theta}_b(\underline{t}_n + 1). \tag{4.26}$$

Now we will obtain a bound on $\nu(t)$. First of all from (4.19) with $k = \underline{t}_n$, observe that

$$\rho_b(t)\,\frac{|e_b(t+1)|}{||\phi(t)||} \leq \bar{\delta}^n, \quad t \in [\underline{t}_n + 1, \underline{t}_n + m + 1]\,;$$

on this interval, if $\phi(t) \neq 0$, then

$$|e_b(t+1)| \leq \bar{\delta}^n\,||\phi(t)||, \tag{4.27}$$

but if $\phi(t) = 0$ then $e_b(t+1) = 0$ as well, so (4.27) is also true. Hence,

$$|e_b(t+1)| \leq \bar{\delta}^n \sup_{j \geq t_0} ||\phi(j)||, \quad t \in [\underline{t}_n + 1, \underline{t}_n + m + 1].$$

Second of all, observe that

$$\begin{aligned} ||\bar{\varepsilon}(t)|| &= \sqrt{|\varepsilon(t)|^2 + |\varepsilon(t+1)|^2} \\ &= \sqrt{|e_g(t)|^2 + |e_g(t+1)|^2} \\ &= \sqrt{\bar{\delta}^{4n} + \bar{\delta}^{4n}} \\ &\leq 2\,\bar{\delta}^{2n}, \quad t \in [\underline{t}_n + 1, \underline{t}_n + m + 1]. \end{aligned}$$

Since $\hat{\theta}_b(\cdot) \in \mathcal{S}_b$, we see that

$$|\bar{\varepsilon}(t)^\top \bar{A}^\top \tilde{\theta}_b(t)| \leq 4\,||\bar{A}^\top||\,\bar{\mathbf{s}}\,\bar{\delta}^n, \quad t \in [\underline{t}_n + 1, \underline{t}_n + m + 1].$$

Last of all, using (4.25) we see that

$$\left|\bar{r}(t)^\top \bar{A}^\top \left(\tilde{\theta}_b(t) - \tilde{\theta}_b(\underline{t}_n + 1)\right)\right| \leq 2\,||r||_\infty\,||\bar{A}^\top||\,m\,\bar{\delta}^n, \quad t \in [\underline{t}_n + 1, \underline{t}_n + m + 1].$$

We conclude that

$$|\nu(t)| \leq \underbrace{\left(\sup_{j \geq t_0} ||\phi(j)|| + 4\,||\bar{A}^\top||\,\bar{\mathbf{s}} + 2\,m\,||r||_\infty\,||\bar{A}^\top||\right)}_{=:d} \bar{\delta}^n, \quad t \in [\underline{t}_n + 1, \underline{t}_n + m + 1]. \quad (4.28)$$

Now let us examine equation (4.26) in detail on the interval $[\underline{t}_n + 1, \underline{t}_n + m + 1]$. To proceed, we define

$$\bar{\nu}(\underline{t}_n + 1) := \begin{bmatrix} \nu(\underline{t}_n + 1) \\ \nu(\underline{t}_n + 2) \\ \vdots \\ \nu(\underline{t}_n + m + 1) \end{bmatrix}$$

and

$$\Psi(\underline{t}_n) := \begin{bmatrix} r(\underline{t}_n + 1) & r(\underline{t}_n + 2) \\ r(\underline{t}_n + 2) & r(\underline{t}_n + 3) \\ \vdots & \vdots \\ r(\underline{t}_n + m + 1) & r(\underline{t}_n + m + 2) \end{bmatrix};$$

from (4.28) it immediately follows that

$$||\bar{\nu}(\underline{t}_n + 1)|| \leq (m + 1)\,d\,\bar{\delta}^n, \quad n \geq N. \quad (4.29)$$

If we incorporate these variables into (4.26), we obtain

$$\bar{\nu}(\underline{t}_n + 1) = \Psi(\underline{t}_n)\,\bar{A}^\top\,\tilde{\theta}_b(\bar{t}_n).$$

Because of the condition on $r$ given in (4.15), it is easy to see that $\Psi(\underline{t}_n)^\top\,\Psi(\underline{t}_n)$ is invertible for $\underline{t}_n \geq T_0$, so this equation has a unique solution for $\tilde{\theta}_b(\bar{t}_n)$:

$$\begin{aligned} \Psi(\underline{t}_n)^\top\,\bar{\nu}(\underline{t}_n + 1) &= \left(\Psi(\underline{t}_n)^\top \Psi(\underline{t}_n)\right)\bar{A}^\top\,\tilde{\theta}_b(\bar{t}_n) \\ \Rightarrow \tilde{\theta}_b(\bar{t}_n) &= \left(\bar{A}^\top\right)^{-1}\left(\Psi(\underline{t}_n)^\top \Psi(\underline{t}_n)\right)^{-1}\Psi(\underline{t}_n)^\top\,\bar{\nu}(\underline{t}_n + 1). \end{aligned} \quad (4.30)$$

To proceed, we need a bound on the size of $\left(\Psi(\underline{t}_n)^\top \Psi(\underline{t}_n)\right)^{-1}$.

**Claim 4.2.**

$$\left\| \left( \Psi(\underline{t}_n)^\top \Psi(\underline{t}_n) \right)^{-1} \right\| \leq \frac{1}{\bar{c}_1}, \quad n \geq N.$$

*Proof of Claim 4.2.* From (4.15), it follows that

$$\text{smallest eigenvalue of } \left( \Psi(\underline{t}_n)^\top \Psi(\underline{t}_n) \right) \geq \bar{c}_1. \tag{4.31}$$

Hence, the largest eigenvalue of $\left( \Psi(\underline{t}_n)^\top \Psi(\underline{t}_n) \right)^{-1}$ is at most $\frac{1}{\bar{c}_1}$. Hence,

$$\left\| \left( \Psi(\underline{t}_n)^\top \Psi(\underline{t}_n) \right)^{-1} \right\| \leq \frac{1}{\bar{c}_1}.$$

■

If we now take the norm of both sides of (4.30) and use the bound on $\bar{\nu}(\underline{t}_n + 1)$ given in (4.29) together with Claim 4.2, and use the fact that $\left\| \Psi(\underline{t}_n) \right\| = \left\| \Psi(\underline{t}_n)^\top \Psi(\underline{t}_n) \right\|^{\frac{1}{2}} \leq \bar{c}_2^{\frac{1}{2}}$, we see that

$$||\tilde{\theta}_b(\bar{t}_n)|| \leq \left\| \left( \bar{A}^\top \right)^{-1} \right\| \times \frac{1}{\bar{c}_1} \times \bar{c}_2^{\frac{1}{2}} \times (m+1)\, d\, \bar{\delta}^n, \quad n \geq N. \tag{4.32}$$

But $\tilde{\theta}_b(\bar{t}_n) = \hat{\theta}_b(\bar{t}_n) - \theta^*$, with $\hat{\theta}_b(\bar{t}_n) \in \mathcal{S}_b$ and $\theta^* \in \mathcal{S}_g$. Since $\mathcal{S}_g \cap \mathcal{S}_b = \varnothing$ and both $\mathcal{S}_g$ and $\mathcal{S}_b$ are compact, there is a gap between them, i.e.,

$$\inf\{\|x - y\| : x \in \mathcal{S}_b \text{ and } y \in \mathcal{S}_g\} > 0.$$

This means that for $n \geq N$ sufficiently large, inequality (4.32) can not be true which is a contradiction to the hypothesis that switching does not stop. We conclude that switching does stop; so by Proposition 4.3 we have $\lim_{t\to\infty} \varepsilon(t) = 0$. ■

## 4.3 Tracking for a Fairly General Reference Signal

To prove tracking of a fairly general reference signal, we make several assumptions on the plant model. To proceed, we first define

$$\underline{a}_1 := \min\{a \in \mathbb{R} : \begin{bmatrix} a \\ b \end{bmatrix} \in \mathcal{S}_1 \text{ for some } b \in \mathbb{R}\},$$
$$\underline{a}_2 := \min\{a \in \mathbb{R} : \begin{bmatrix} a \\ b \end{bmatrix} \in \mathcal{S}_2 \text{ for some } b \in \mathbb{R}\},$$
$$\underline{b}_1 := \min\{b \in \mathbb{R} : \begin{bmatrix} a \\ b \end{bmatrix} \in \mathcal{S}_1 \text{ for some } a \in \mathbb{R}\},$$
$$\bar{b}_1 := \max\{b \in \mathbb{R} : \begin{bmatrix} a \\ b \end{bmatrix} \in \mathcal{S}_1 \text{ for some } a \in \mathbb{R}\},$$
$$\underline{b}_2 := \min\{b \in \mathbb{R} : \begin{bmatrix} a \\ b \end{bmatrix} \in \mathcal{S}_2 \text{ for some } a \in \mathbb{R}\},$$
$$\bar{b}_2 := \max\{b \in \mathbb{R} : \begin{bmatrix} a \\ b \end{bmatrix} \in \mathcal{S}_2 \text{ for some } a \in \mathbb{R}\}.$$

Here we will assume that all admissible models are unstable with a positive pole location, and that the sign of the $b$ term for the sets are opposite. That is, we would like to impose

**Assumption 4.1.**

*(i)* $\underline{a}_1 > 1$

*(ii)* $\underline{a}_2 > 1$

*(iii) Either*

$$\underline{b}_1 > 0 \qquad \textit{and} \qquad \bar{b}_2 < 0$$

*or*

$$\bar{b}_1 < 0 \qquad \textit{and} \qquad \bar{b}_2 > 0.$$

If Assumption 4.1 holds, we define

$$\underline{a} \coloneqq \min\left\{\underline{a}_1, \underline{a}_2\right\}$$

and

$$\underline{b} \coloneqq \begin{cases} \min\left\{\underline{b}_1, -\bar{b}_2\right\}, & \text{if } \underline{b}_1 > 0, \\ \min\left\{-\bar{b}_1, \underline{b}_2\right\}, & \text{if } \underline{b}_1 < 0. \end{cases}$$

Hence, $\underline{a}$ provides a lower bound on the $a$ variable, and $\underline{b}$ provides a lower bound on the magnitude of the $b$ variable.

Now we present the main result of this section.

**Theorem 4.2.** *Suppose that Assumption 4.1 holds and the adaptive controller consisting of the parameter estimator (2.11), (2.12), (2.13), the control law (2.22), and the switching rule (2.26), (2.27) is applied to the plant (2.1). For every $\bar{\lambda} \in (0,1)$, $t_0 \in \mathbb{Z}$, $x_0 \in \mathbb{R}$, $\theta^* \in \mathcal{S}$, $\sigma_0 \in \{1,2\}$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$, $i \in \{1,2\}$, and $r \in l_\infty$, if $n = 0$ and $\liminf_{t\to\infty} |r(t)| > 0$, then switching stops and*

$$\lim_{t\to\infty} \varepsilon(t) = 0.$$

*Proof.* Fix $\bar{\lambda} \in (0,1)$, $t_0 \in \mathbb{Z}$, $x_0 \in \mathbb{R}$, $\theta^* \in \mathcal{S}$, $\sigma_0 \in \{1,2\}$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$, $i \in \{1,2\}$, $n = 0$, and $r \in l_\infty$ so that $\liminf_{t\to\infty} |r(t)| > 0$. Before proceeding, we define two sets of times according to the estimator chosen by the switching signal $\sigma$:

$$\begin{aligned} S_{\text{good}} &= \{t \geq t_0 \,:\, \sigma(t) = g\} \\ S_{\text{bad}} &= \{t \geq t_0 \,:\, \sigma(t) = b\}. \end{aligned}$$

Clearly $\{t \in \mathbb{Z} : t \geq t_0\} = S_{\text{good}} \cup S_{\text{bad}}$.

Now we partition the time index $\{j \in \mathbb{Z} : j \geq t_0\}$ into intervals which oscillate between $S_{\text{good}}$ and $S_{\text{bad}}$. We can clearly define a (possible infinite) sequence of intervals of the form $[k_l, k_{l+1})$, $l \in \mathbb{Z}^+$ which satisfy:

(i) without loss of generality, $k_0 = t_0$ serves as the initial instant of the first interval;

(ii) $[k_l, k_{l+1})$ either belongs to $S_{\text{good}}$ or $S_{\text{bad}}$; and

(iii) if $k_{l+1} \neq \infty$ and $[k_l, k_{l+1})$ belongs to $S_{\text{good}}$, then $[k_{l+1}, k_{l+2})$ belongs to $S_{\text{bad}}$ and *vice versa*.

Since $n = 0$ and $r \in l_\infty$, for the case when $J_b(t) \to 0$ as $t \to \infty$, by Proposition 4.2 we have that $\lim_{t\to\infty} \varepsilon(t) = 0$. So for the rest of the proof we suppose that this is not the case, i.e., we have $\limsup_{t\to\infty} |J_b(t+1)| > 0$. We will prove that switching stops, so we can conclude from Proposition 4.3 that $\lim_{t\to\infty} \varepsilon(t) = 0$.

To proceed, suppose that we do not stop switching. The goal is to obtain a difference equation for $r(t)$ which is driven by the prediction error $e_b(t)$ and the tracking error $\varepsilon(t)$. We first analyze the prediction error $e_b(t)$. For each $t \geq t_0$ from (2.3) we have

$$e_b(t+1) = x(t+1) - \phi(t)^\top \hat{\theta}_b(t),$$

which can be rewritten as

$$\begin{aligned} x(t+1) &= e_b(t+1) + \phi(t)^\top \hat{\theta}_b(t) \\ &= \hat{a}_b(t)x(t) + \hat{b}_b(t)u(t) + e_b(t+1). \end{aligned}$$

So

$$\varepsilon(t+1) + r(t+1) = \hat{a}_b(t)x(t) + \hat{b}_b(t)u(t) + e_b(t+1).$$

From (2.23), we have

$$\begin{aligned} \varepsilon(t+1) + r(t+1) &= \hat{a}_b(t)\Big(x(t) - r(t) + r(t)\Big) + \hat{b}_b(t)u(t) + e_b(t+1) \\ \Rightarrow r(t+1) &= \hat{a}_b(t)r(t) + \hat{b}_b(t)u(t) + \underbrace{\hat{a}_b(t)\varepsilon(t) + e_b(t+1) - \varepsilon(t+1)}_{=:\nu_1(t)}. \end{aligned} \tag{4.33}$$

We would like to get rid of the $u(t)$ term; to proceed, we rewrite the plant model (2.1) as

$$\varepsilon(t+1) + r(t+1) = a\Big(\varepsilon(t) + r(t)\Big) + b\,u(t),$$

or equivalently

$$r(t+1) = a\,r(t) + b\,u(t) + \underbrace{a\,\varepsilon(t) - \varepsilon(t+1)}_{=:\nu_2(t)};$$

if we solve for $u(t)$ we obtain

$$u(t) = \frac{1}{b}\Big(r(t+1) - a\,r(t) - \nu_2(t)\Big). \tag{4.34}$$

Now we substitute the formula for $u(t)$ given in (4.34) into the update equation for $r(t)$ given in (4.33):

$$\begin{aligned} r(t+1) &= \hat{a}_b(t)r(t) + \frac{\hat{b}_b(t)}{b}\Big(r(t+1) - a\,r(t) - \nu_2(t)\Big) + \nu_1(t) \\ &= \Big(\hat{a}_b(t) - \frac{a}{b}\,\hat{b}_b(t)\Big)r(t) + \frac{\hat{b}_b(t)}{b}\,r(t+1) + \nu_1(t) - \frac{\hat{b}_b(t)}{b}\,\nu_2(t)\,; \end{aligned}$$

if we define

$$\nu_3(t) \coloneqq \nu_1(t) - \frac{\hat{b}_b(t)}{b}\,\nu_2(t)\,,$$

then the above becomes

$$\Big(1 - \underbrace{\frac{\hat{b}_b(t)}{b}}_{<0}\Big)\,r(t+1) = \Big(\hat{a}_b(t) - \frac{a}{b}\,\hat{b}_b(t)\Big)r(t) + \nu_3(t).$$

Since $1 - \dfrac{\hat{b}_b(t)}{b} > 1$, it is invertible. Therefore, the above can be rewritten as

$$r(t+1) = \frac{b}{b - \hat{b}_b(t)}\Big(\hat{a}_b(t) - \frac{a}{b}\,\hat{b}_b(t)\Big)r(t) + \frac{b}{b - \hat{b}_b(t)}\,\nu_3(t)\,; \tag{4.35}$$

define

$$\nu_4(t) \coloneqq \frac{b}{b - \hat{b}_b(t)}\,\nu_3(t)$$

and

$$a_{cl}(t) \coloneqq \frac{b\,\hat{a}_b(t) - a\,\hat{b}_b(t)}{b - \hat{b}_b(t)}.$$

It follows that

$$a_{cl}(t) = \underbrace{\frac{b}{b - \hat{b}_b(t)}}_{=:\,\alpha(t)}\,\hat{a}_b(t) + a\,\Bigg(\underbrace{\frac{-\hat{b}_b(t)}{b - \hat{b}_b(t)}}_{=:\,\beta(t)}\Bigg) \tag{4.36}$$

**Claim 4.3.** $\alpha(t) \in (0,1)$ *for* $t \geq t_0$.

*Proof of Claim 4.3* . If $b < 0$, then $\hat{b}_g(t) < 0$ and $\hat{b}_b(t) > 0$, so it is clear that

$$\alpha(t) = \frac{b}{b - \hat{b}_b(t)} \in (0,1).$$

On the other hand, if $b > 0$, then $\hat{b}_g(t) > 0$ and $\hat{b}_b(t) < 0$, so it is clear that

$$\alpha(t) = \frac{b}{b - \hat{b}_b(t)} \in (0,1).$$

∎

Now observe that

$$\alpha(t) + \beta(t) = \frac{b - \hat{b}_b(t)}{b - \hat{b}_b(t)} = 1,$$

which means that

$$\beta(t) = 1 - \alpha(t).$$

From Claim 4.3, this means that $\beta(t) \in (0,1)$ as well. Hence, $a_{cl}(t)$ is a convex combination of $\hat{a}_b(t)$ and $a$. Since

$$\hat{a}_b(t) \geq \underline{a}$$

and

$$a \geq \underline{a},$$

it follows that

$$a_{cl}(t) \geq \underline{a} > 1 \tag{4.37}$$

as well. Hence, the update equation (4.35) for $r(t+1)$ is unstable and driven by a weighted sum of $e_b(t)$ and $\varepsilon(t)$. More specifically, we can conclude that

$$r(t+1) \geq \underline{a}\, r(t) + \nu_4(t), \quad t \geq t_0. \tag{4.38}$$

To understand the behaviour of the above inequality, we need to examine $\nu_4(t)$ in more detail:

$$\begin{aligned}
\nu_4(t) &= \frac{b}{b-\hat{b}_b(t)}\Big(\hat{a}_b(t)\varepsilon(t) + e_b(t+1) - \varepsilon(t+1) - \frac{\hat{b}_b(t)}{b}\Big(a\,\varepsilon(t) - \varepsilon(t+1)\Big)\Big)\\
&= \frac{b}{b-\hat{b}_b(t)}\Big(\hat{a}_b(t)\varepsilon(t) + e_b(t+1) - \varepsilon(t+1)\Big) + \frac{\hat{b}_b(t)}{b-\hat{b}_b(t)}\Big(\varepsilon(t+1) - a\,\varepsilon(t)\Big)\\
&= -\varepsilon(t+1) + \left(\frac{\hat{a}_b(t)\,b - a\,\hat{b}_b(t)}{b-\hat{b}_b(t)}\right)\varepsilon(t) + \left(\frac{b}{b-\hat{b}_b(t)}\right)e_b(t+1).
\end{aligned}$$

Given that

$$|b - \hat{b}_b(t)| \geq 2\,\underline{b}$$

and $\begin{bmatrix} a \\ b \end{bmatrix}$, $\begin{bmatrix} \hat{a}_b(t) \\ \hat{b}_b(t) \end{bmatrix}$, and $\begin{bmatrix} \hat{a}_g(t) \\ \hat{b}_g(t) \end{bmatrix}$ lie in the compact set $\mathcal{S}_1 \cup \mathcal{S}_2$, we conclude that there exists a constant $\bar{c}$ so that regardless of the values of $\hat{a}_b(t)$, $\hat{b}_b(t)$, $a$, and $b$, we have

$$|\nu_4(t)| \leq \bar{c}\Big(|\varepsilon(t)| + |\varepsilon(t+1)| + |e_b(t+1)|\Big), \quad t \geq t_0. \tag{4.39}$$

Now we turn to analyzing the performance signals. Since $\limsup_{t\to\infty} J_b(t+1) > 0$ by hypothesis, from (2.26) it follows that

$$\limsup_{t\to\infty} \rho_b(t)\,\frac{|e_b(t+1)|}{||\phi(t)||} > 0.$$

Now define

$$\bar{\delta} := \min\left\{\frac{1}{2},\ \frac{1}{2}\limsup_{t\to\infty} \rho_b(t)\,\frac{|e_b(t+1)|}{||\phi(t)||},\ \frac{1}{2}\liminf_{t\to\infty} |r(t)|\right\} > 0$$

and

$$T_1 := \min\{t \geq t_0 : |r(j)| \geq \bar{\delta} \text{ for } j \geq t\}.$$

We would like to examine (4.38) when $e_g(t+1)$ and $J_g(t+1)$ are small. Since both tend to zero by Proposition 4.1, for each $n \geq 2$ we can define $T_n \geq T_1$ such that

$$J_g(t+1) < \bar{\delta}^{2n}, \quad t \geq T_n \tag{4.40}$$

$$|e_g(t+1)| < \bar{\delta}^{2n}, \quad t \geq T_n\,; \tag{4.41}$$

without loss of generality $T_{n+1} > T_n$ for $n \geq 2$. From the definition of $\bar{\delta}$, we know that

$$\frac{1}{2} \limsup_{t \to \infty} \rho_b(t) \frac{|e_b(t+1)|}{||\phi(t)||} \geq \bar{\delta};$$

this means that there exists a $t_n \geq T_n$ so that

$$\rho_b(t_n) \frac{|e_b(t_n+1)|}{||\phi(t_n)||} \geq \bar{\delta}, \tag{4.42}$$

which implies that

$$J_b(t_n+1) \geq \bar{\delta}$$

as well. Since $T_n$ is defined so that both (4.40) and (4.41) holds, this means that

$$J_g(t_n+1) < \bar{\delta}^{2n} < \bar{\delta} \leq J_b(t_n+1),$$

so $\sigma(t_n+1) = g$. Since we have assumed that we switch forever, there exists a $\bar{t}_n > t_n + 2$ so that

$$\sigma(t) = g, \quad t \in [t_n+1, \bar{t}_n),$$

and

$$\sigma(\bar{t}_n) = b\,;$$

this means that, $[t_n+1, \bar{t}_n) \subset S_{\text{good}}$. But in order to switch from $S_{\text{good}}$ to $S_{\text{bad}}$, we need

$$J_b(\bar{t}_n) \leq J_g(\bar{t}_n) < \bar{\delta}^{2n}\,;$$

this means that at some time $t \in (t_n+1, \bar{t}_n]$, we have

$$\rho_b(t-1) \frac{|e_b(t)|}{||\phi(t-1)||} \leq \bar{\delta}^n.$$

From (4.42), it follows that there must be a **minimum time** $\tilde{t}_n \in [t_n+1, \bar{t}_n)$, such that

$$\rho_b(t-1) \frac{|e_b(t)|}{||\phi(t-1)||} < \bar{\delta}^n, \quad t \in (\tilde{t}_n, \bar{t}_n)\,; \tag{4.43}$$

this means, in particular, that

$$J_b(\tilde{t}_n) \geq \rho_b(\tilde{t}_n - 1) \frac{|e_b(\tilde{t}_n)|}{||\phi(\tilde{t}_n - 1)||} \geq \bar{\delta}^n. \tag{4.44}$$

**Claim 4.4.** *$\bar{t}_n - \tilde{t}_n$ satisfies*

$$\bar{t}_n - \tilde{t}_n \geq n \, \frac{\ln \bar{\delta}}{\ln \bar{\lambda}}. \tag{4.45}$$

*Proof of Claim 4.4.* To switch from $S_{\text{good}}$ to $S_{\text{bad}}$ at $t = \bar{t}_n$, we need

$$J_b(\bar{t}_n) \leq J_g(\bar{t}_n) \leq \bar{\delta}^{2n}.$$

If we solve the (2.26) with an initial condition of (4.44), we see that

$$J_b(\bar{t}_n) \geq \bar{\lambda}^{(\bar{t}_n - \tilde{t}_n)} \, \bar{\delta}^n.$$

So we need to have

$$\begin{aligned} \bar{\lambda}^{(\bar{t}_n - \tilde{t}_n)} \bar{\delta}^n &\leq \bar{\delta}^{2n} \\ \Leftrightarrow \bar{\lambda}^{(\bar{t}_n - \tilde{t}_n)} &\leq \bar{\delta}^n \\ \Leftrightarrow (\bar{t}_n - \tilde{t}_n) \ln \bar{\lambda} &\leq n \ln \bar{\delta} \end{aligned}$$

Since $\bar{\lambda} \in (0, 1)$, we have

$$\bar{t}_n - \tilde{t}_n \geq n \, \frac{\ln \bar{\delta}}{\ln \bar{\lambda}} > 0.$$

■

**Claim 4.5.** *There exists $n_1 \geq 2$ such that*

$$|r(\bar{t}_n - 1)| > \left(\frac{\underline{a} + 1}{2}\right)^{\bar{t}_n - \tilde{t}_n - 2} \bar{\delta}, \quad n \geq n_1.$$

*Proof of Claim 4.5.* From (4.41) and (4.43), we have

$$\begin{aligned} |e_g(t)| = |\varepsilon(t)| &< \bar{\delta}^{2n}, \quad t \in (\tilde{t}_n, \bar{t}_n) \\ \rho_b(t-1) \, \frac{|e_b(t)|}{||\phi(t-1)||} &< \bar{\delta}^n, \quad t \in (\tilde{t}_n, \bar{t}_n); \end{aligned} \tag{4.46}$$

Since $\rho_b(t-1) = 0$ implies that $e_b(t) = 0$ and $\phi(t-1) = 0$, this means that

$$|e_b(t)| \leq \bar{\delta}^n \, ||\phi(t-1)||, \quad t \in (\tilde{t}_n, \bar{t}_n).$$

From Theorem 3.1, we know that $\phi$ is bounded, so

$$|e_b(t)| \le \bar{\delta}^n \sup_{j\ge t_0} ||\phi(j)||, \quad t \in (\tilde{t}_n, \bar{t}_n). \tag{4.47}$$

Combining (4.38), (4.39), (4.46) and (4.47), the above implies that

$$\begin{aligned} |r(t+1)| &\ge \underline{a}\,|r(t)| - |\nu_4(t)| \\ &> \underline{a}\,|r(t)| - \bar{c}(2\,\bar{\delta}^{2n} + \bar{\delta}^n \sup_{j\ge t_0} ||\phi(j)||), \quad t \in (\tilde{t}_n, \bar{t}_n - 1). \end{aligned}$$

Now choose $n_1 \ge 2$ such that

$$\bar{c}(2\,\bar{\delta}^{2n} + \bar{\delta}^n \sup_{j\ge t_0} ||\phi(j)||) \le \frac{1}{2}\,(\underline{a}-1)\bar{\delta}, \quad n \ge n_1 \ge 2.$$

So for $n \ge n_1$

$$\begin{aligned} |r(t+1)| &\ge \underline{a}\,|r(t)| - \frac{1}{2}\,(\underline{a}-1)\bar{\delta} \\ &= \left(\frac{\underline{a}+1}{2}\right)|r(t)| + \left(\frac{\underline{a}-1}{2}\right)|r(t)| - \left(\frac{\underline{a}-1}{2}\right)\bar{\delta}, \quad t \in (\tilde{t}_n, \bar{t}_n - 1)\,. \end{aligned}$$

Using the definition of $T_1$, we see that

$$\left(\frac{\underline{a}-1}{2}\right)|r(t)| - \left(\frac{\underline{a}-1}{2}\right)\bar{\delta} \ge 0, \quad t \ge T_1,$$

so we have

$$|r(t+1)| \ge \left(\frac{\underline{a}+1}{2}\right)|r(t)|, \quad t \in (\tilde{t}_n, \bar{t}_n - 1).$$

Since $|r(\tilde{t}_n + 1)| \ge \bar{\delta}$, solving this recursively yields

$$|r(\bar{t}_n - 1)| \ge \left(\frac{\underline{a}+1}{2}\right)^{\bar{t}_n - \tilde{t}_n - 2}\bar{\delta}, \quad n \ge n_1 \ge 2\,.$$

■

If we combine Claim 4.4 and Claim 4.5, we see that

$$|r(\bar{t}_n - 1)| > \left(\frac{\underline{a}+1}{2}\right)^{n_1 \frac{\ln \bar{\delta}}{\ln \lambda} - 2} \bar{\delta}, \quad n \geq n_1.$$

This means that

$$\limsup_{t \to \infty} |r(t)| = \infty,$$

which contradicts the fact that $r$ is bounded. Hence, we deduce that switching signal $\sigma$ does indeed stop switching, so from Proposition 4.3 we have $\lim_{t\to\infty} \varepsilon(t) = 0$. ■

# Chapter 5

# Numerical Simulations

In this chapter we present a couple of simulation scenarios to illustrate the efficacy of the proposed controller with parameter estimator (2.11), (2.12), (2.13), and performance signal based switching rule (2.26), (2.27) combined with adaptive control law (2.22).

## 5.1 Simulation Parameters

Throughout this chapter we consider the time-varying plant

$$x(t+1) = a(t)x(t) + b(t)u(t) + n(t) \tag{5.1}$$

with $\theta^*(t) = \begin{bmatrix} a(t) & b(t) \end{bmatrix}^\top$ in the uncertainty set

$$\mathcal{S} = \left\{ \begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{R}^2 : a \in [1,5],\ b \in [1,5] \cup [-5,-1] \right\}.$$

The sets $\mathcal{S}_1$ and $\mathcal{S}_2$ of Proposition 2.1 are taken to be

$$\mathcal{S}_1 = \left\{ \begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{R}^2 : a \in [1,5],\ b \in [1,5] \right\}, \quad \mathcal{S}_2 = \left\{ \begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{R}^2 : a \in [1,5],\ b \in [-5,-1] \right\}.$$

We present two simulation scenarios and in each scenario we consider two cases. In the first case, referred to as the "continuous parameter" case, the plant parameters are

$$\begin{aligned} a(t) &= 3 + 2\,\sin(0.03\,t) \\ b(t) &= -3 - 2\,\cos(0.05\,t). \end{aligned}$$

In the continuous parameter case $\theta^*(t)$ lies in $\mathcal{S}_2$ and sweeps the entire parameter range.

The second case, referred to as the "discontinuous parameter" case, the plant parameters are

$$a(t) = 3 + 2\,\sin(0.03\,t)$$
$$b(t) = \big(-3 - 2\,\cos(0.05\,t)\big)\cdot \operatorname{sign}\left(\sin\left(\frac{t\,\pi}{1000}\right)\right).$$

In the discontinuous parameter case $\theta^*(t)$ jumps between sets $\mathcal{S}_1$ and $\mathcal{S}_2$ every 1000 times steps.

For every simulation, we take $\delta = 0.1$ in (2.11) and the plant and the controller are initialized as follows: $t_0 = 0$, $\phi(0) = \begin{bmatrix}1 & 1\end{bmatrix}^\top$, $\hat{\theta}_1(0) = \begin{bmatrix}1 & 1\end{bmatrix}^\top$, $\hat{\theta}_2(0) = \begin{bmatrix}1 & -1\end{bmatrix}^\top$, $\rho_1(0) = \rho_2(0) = 0$, $\sigma(0) = 1$, $J_1(0) = J_2(0) = 0$. We simulate all scenarios for 10000 time steps. With a slight abuse of notation, the 2-norm of a discrete-time signal $x$ obtained through simulation is denoted

$$\|x\|_2 \coloneqq \sqrt{\sum_{t=0}^{10000} |x(t)|^2}.$$

## 5.2 Scenario 1: Tracking

In this scenario, we set the noise $n$ to zero and take the reference signal to be

$$r(t) = \sin(\omega\,t), \quad \omega \in \{1, 2, \ldots, 100\}. \tag{5.2}$$

For each frequency $\omega$, we compute the tracking gain $W_1(\omega)$ defined as

$$W_1(\omega) \coloneqq \frac{\|\varepsilon\|_2}{\|r\|_2}.$$

Table 5.1 illustrates the average and range of the tracking gain $W_1$ for different values of $\bar{\lambda}$. On one hand, in the continuous parameter case, we observe that both the mean and the maximum value of $W_1(\omega)$ increases as $\bar{\lambda}$ decreases. On the other hand, in discontinuous parameter case, the mean of $W_1(\omega)$ significantly decreases as $\bar{\lambda}$ decreases.

Table 5.2 shows the mean and the standard deviation of the number of switches over the range of $\omega$ for different values of $\bar{\lambda}$. From this table we notice that both the mean and the standard deviation increase as the value of $\bar{\lambda}$ decreases in both the continuous and discontinuous parameter cases. We also deduce that the proposed performance signal based adaptive controller outperforms the one presented in [34], i.e., when $\bar{\lambda} = 0$.

Table 5.1: Mean and range of tracking gain $W_1(\omega)$ when $n = 0$ and $r$ is time-varying

| $\bar{\lambda}$ | continuous parameter | | discontinuous parameter | |
|---|---|---|---|---|
| | Mean | Range | Mean | Range |
| 0.9 | 0.49 | [0.24, 1.33] | $1.13 \times 10^5$ | [305, $1.53 \times 10^6$] |
| 0.8 | 0.48 | [0.24, 1.22] | 343.6 | [7.12, 2039] |
| 0.7 | 0.5 | [0.24, 1.53] | 52 | [3.57, 155] |
| 0.6 | 0.49 | [0.24, 1.45] | 31.09 | [1.13, 145] |
| 0.5 | 0.49 | [0.25, 1.47] | 16.89 | [1.11, 138] |
| 0.4 | 0.50 | [0.25, 1.29] | 5.27 | [0.53, 15.19] |
| 0.3 | 0.64 | [0.25, 4.80] | 4.70 | [0.49, 10.84] |
| 0.2 | 0.78 | [0.27, 4.81] | 4.54 | [0.65, 10.94] |
| 0.1 | 1.07 | [0.31, 5.29] | 4.56 | [0.74, 10.25] |
| 0 | 1.55 | [0.38, 5.33] | 4.50 | [0.69, 16.26] |

## 5.3 Scenario 2: Noise Gain

In this scenario, we set the reference signal to zero and take the noise as

$$n(t) = \frac{1}{50} \sin(\omega\, t), \quad \omega \in \{1, 2, \dots, 100\}. \tag{5.3}$$

For each frequency $\omega$ we compute the noise gain $W_2(\omega)$ defined as

$$W_2(\omega) := \frac{\|x\|_2}{\|n\|_2}\,.$$

Table 5.3 illustrates the mean and the range of the noise gain $W_2$ over $\omega$ for different values of $\bar{\lambda}$. We observe that the mean of the gain $W_2$ gets considerably smaller as $\bar{\lambda}$ decreases in the continuous parameter case. For the discontinuous parameter case, we again observe that the mean of the noise gain $W_2$ gets significantly smaller (by more than a factor of $10^3$) for smaller values of $\bar{\lambda}$.

Table 5.2: Number of switches when $n = 0$ and $r$ is time-varying

| $\bar{\lambda}$ | # switches with continuous parameter | | # switches with discontinuous parameter | |
|---|---|---|---|---|
| | Mean | Standard deviation | Mean | Standard deviation |
| 0.9 | 1.18 | 1.45 | 10.56 | 1.77 |
| 0.8 | 2.66 | 11.70 | 12.12 | 11.43 |
| 0.7 | 3.68 | 18.15 | 13.76 | 18.78 |
| 0.6 | 4.90 | 23.85 | 14.68 | 24.20 |
| 0.5 | 7.06 | 28.13 | 17.04 | 28.11 |
| 0.4 | 11.88 | 33.67 | 22.58 | 34.26 |
| 0.3 | 21.19 | 42.79 | 31.95 | 42.94 |
| 0.2 | 38.37 | 52.67 | 49.07 | 52.44 |
| 0.1 | 87.89 | 64.99 | 98.53 | 64.37 |
| 0 | 312 | 145 | 322 | 144 |

To draw comparisons between the mean and the standard deviation of the number of switches over the range of $\omega$, and different values of $\bar{\lambda}$, we move to Table 5.4, which shows that both the mean and the standard deviation of the number of switches increases drastically as $\bar{\lambda}$ decreases, in both the continuous and discontinuous parameter cases.

## 5.4 Representative Examples

On the one hand, from Table 5.1 we observed that in the continuous parameter case the mean of $W_1$ gradually increases with a decrease in the value of $\bar{\lambda}$. In the discontinuous parameter case, however, the mean of $W_1$ sharply decreases as $\bar{\lambda}$ decreases. On the other hand, from Table 5.3, we clearly see that in both the continuous and discontinuous parameter cases, the mean of $W_2$ considerably decreases as $\bar{\lambda}$ decreases. Therefore, we argue that $\bar{\lambda} = 0.3$ provides nice trade-off between small tracking gain ($W_1$) and small noise gain

Table 5.3: Mean and range of noise gain $W_2(\omega)$ when $n$ is time-varying and $r = 0$

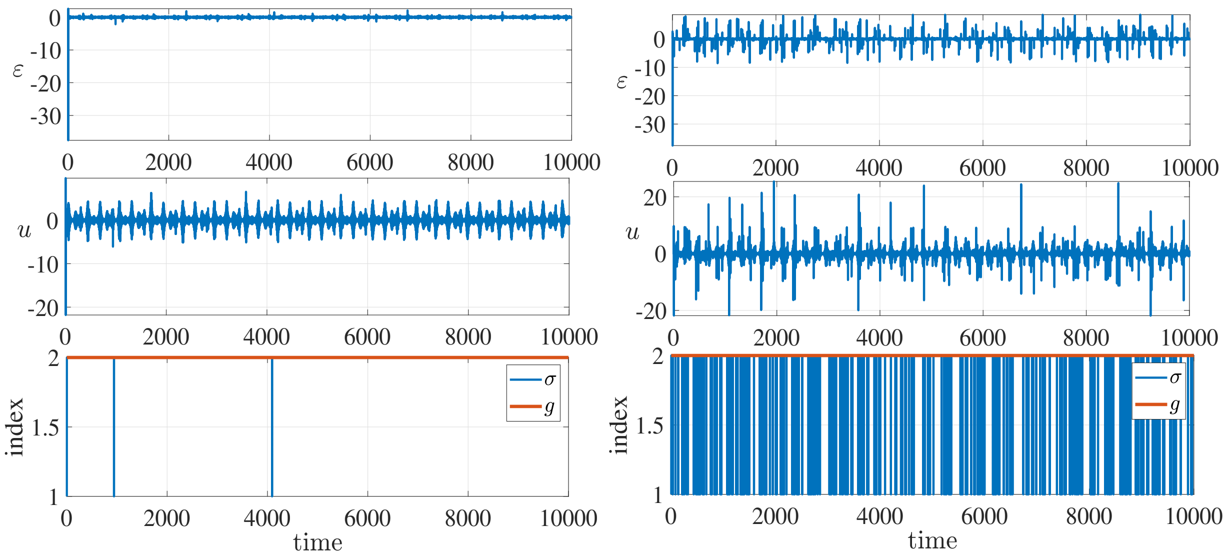
| $\bar{\lambda}$ | continuous parameter | | discontinuous parameter | |
|---|---|---|---|---|
| | Mean | Range | Mean | Range |
| 0.9 | 829 | [12.37, $1.64 \times 10^4$] | $5.95 \times 10^4$ | [426, $3.30 \times 10^6$] |
| 0.8 | 1090 | [12.37, 9229] | 1445 | [43.02, $1.05 \times 10^4$] |
| 0.7 | 765 | [12.37, 9229] | 970 | [21.98, 9628] |
| 0.6 | 232 | [13.41, 1480] | 291 | [18.43, 1910] |
| 0.5 | 183 | [15.2, 1274] | 232 | [17.37, 2079] |
| 0.4 | 101 | [19.84, 606] | 118.3 | [21.82, 622] |
| 0.3 | 46.45 | [18.78, 337] | 61.44 | [19.17, 217] |
| 0.2 | 34.03 | [18.21, 131] | 41.77 | [17.67, 131] |
| 0.1 | 28.92 | [17.10, 71.65] | 33.90 | [17.53, 74] |
| 0 | 26.49 | [17.44, 59.81] | 31.64 | [16.67, 118] |

($W_2$) in both the continuous and discontinuous parameter cases.

To illustrate this observation, we simulate with $\bar{\lambda} = 0.3$, $\omega = 1$ for $r$ as given in (5.2) and $\omega = 10$ for $n$ as given in (5.3). Representative figures for Scenario 1 are shown in Figures 5.1 and 5.2, which show the performance of our proposed adaptive controller for $\bar{\lambda} = 0.3$ against that of presented in [34], i.e., $\bar{\lambda} = 0$, for both the continuous and discontinuous case, respectively. It is clear that our performance signal based adaptive controller has smaller spikes and significantly fewer number of switches than [34] for both the continuous and discontinuous parameter cases. Figure 5.3 shows the actual and estimated parameters and the performance signals $J_1$ and $J_2$ for the continuous case and discontinuous parameter cases. For Scenario 2, the proposed performance signal based switching algorithm with $\bar{\lambda} = 0.3$ produces bigger spikes than simple switching rule in [34] for both the continuous (Figure 5.4) and discontinuous (Figure 5.5) parameter cases. Figure 5.6 compares the actual and estimated parameters along with performance signals $J_1$ and $J_2$ for the continuous and discontinuous parameter cases.

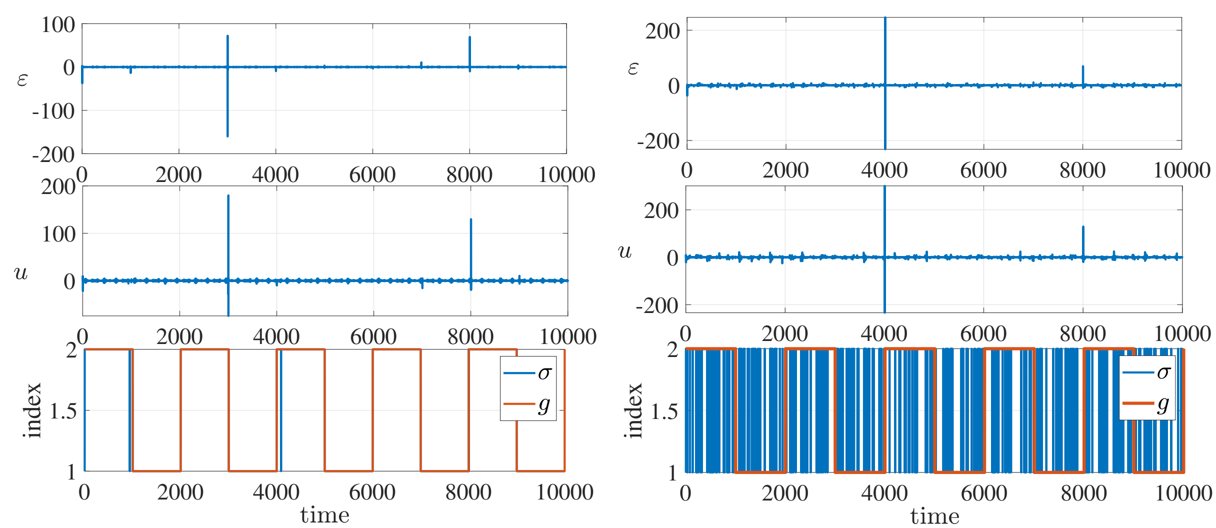
Table 5.4: Number of switches when $n$ is time-varying and $r = 0$

| $\bar{\lambda}$ | # switches with continuous parameter | | # switches with discontinuous parameter | |
|---|---|---|---|---|
| | Mean | Standard deviation | Mean | Standard deviation |
| 0.9 | 20.26 | 7.84 | 30.3 | 7.66 |
| 0.8 | 32.88 | 15.40 | 43.28 | 15.27 |
| 0.7 | 179 | 99.51 | 187 | 96.94 |
| 0.6 | 543 | 166 | 549 | 164 |
| 0.5 | 1401 | 257 | 1401 | 253 |
| 0.4 | 2069 | 388 | 2067 | 385 |
| 0.3 | 2320 | 407 | 2323 | 398 |
| 0.2 | 2497 | 420 | 2496 | 416 |
| 0.1 | 2639 | 425 | 2643 | 417 |
| 0 | 2904 | 389 | 2842 | 400 |



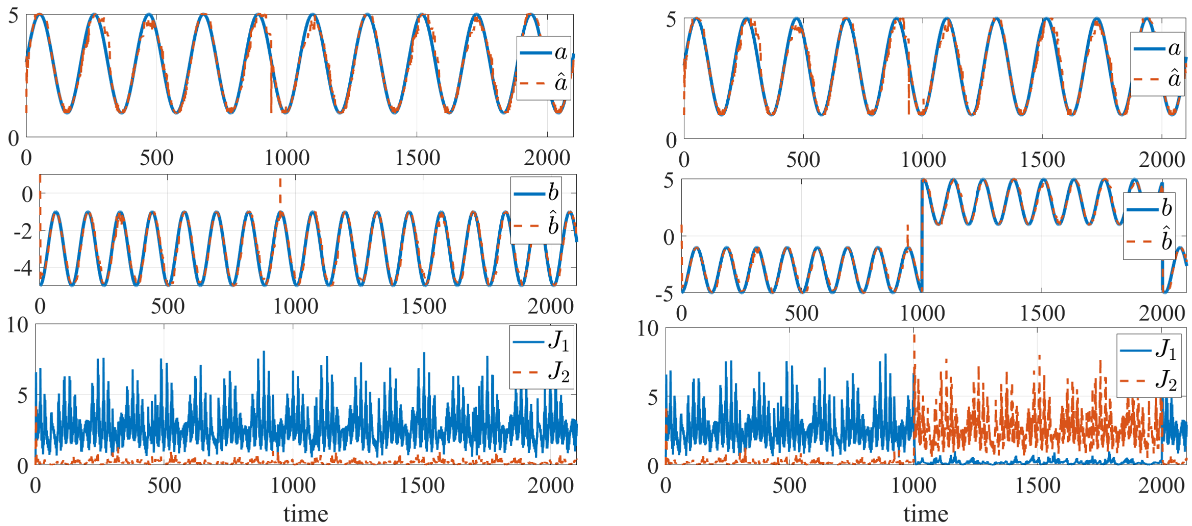

(a) $\bar{\lambda} = 0.3$ (b) $\bar{\lambda} = 0$

Figure 5.1: Plots for continuous parameter, time-varying reference and no noise. Column (a) shows our proposed algorithm whereas column (b) shows the one presented in [34].



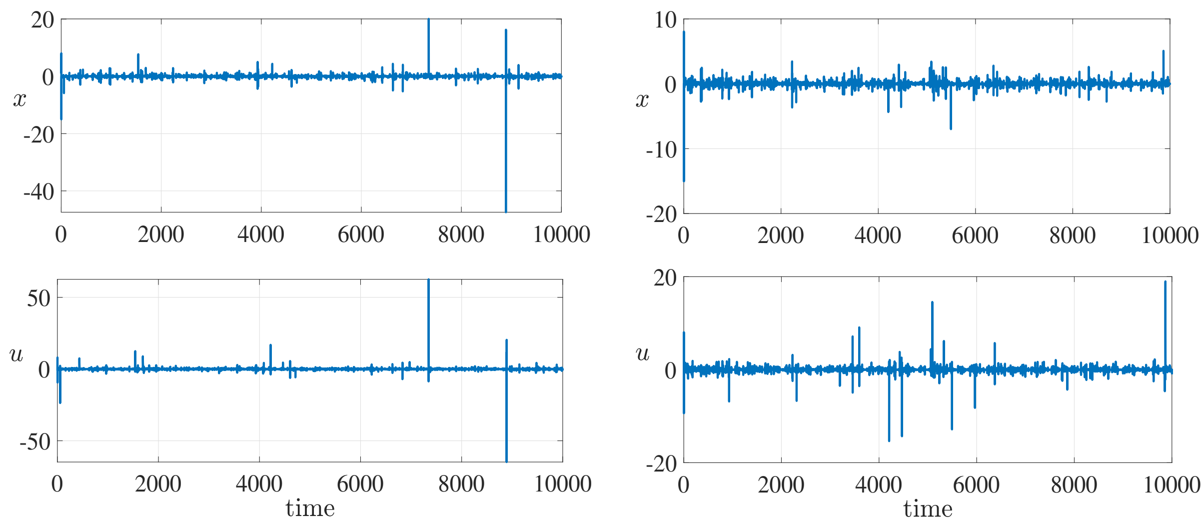

(a) $\bar{\lambda} = 0.3$ (b) $\bar{\lambda} = 0$

Figure 5.2: Plots for discontinuous parameter, time-varying reference and no noise. Column (a) shows our proposed algorithm whereas column (b) shows the one presented in [34].




(a) continuous parameters (b) discontinuous parameters

Figure 5.3: Plots for time-varying reference and no noise with $\bar{\lambda} = 0.3$. Column (a) shows the continuous parameter case whereas column (b) shows the discontinuous parameter case.




(a) $\bar{\lambda} = 0.3$ (b) $\bar{\lambda} = 0$

Figure 5.4: Plot for continuous parameter, time-varying noise and no reference. Column (a) shows our proposed algorithm whereas column (b) shows the one presented in [34].

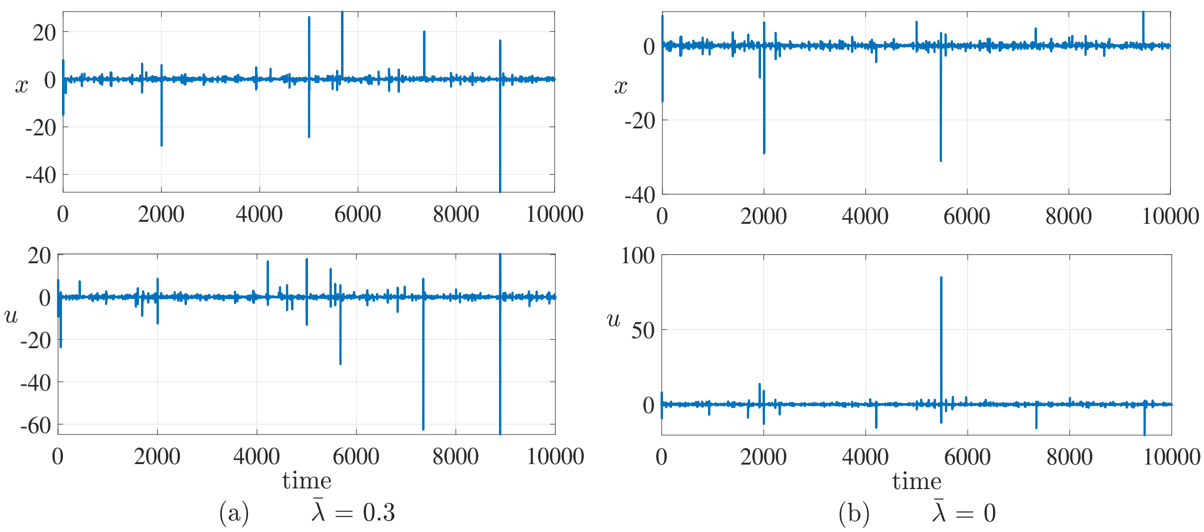



Figure 5.5: Plots for discontinuous parameter, time-varying noise and no reference. Column (a) shows our proposed algorithm whereas column (b) shows the one presented in [34].

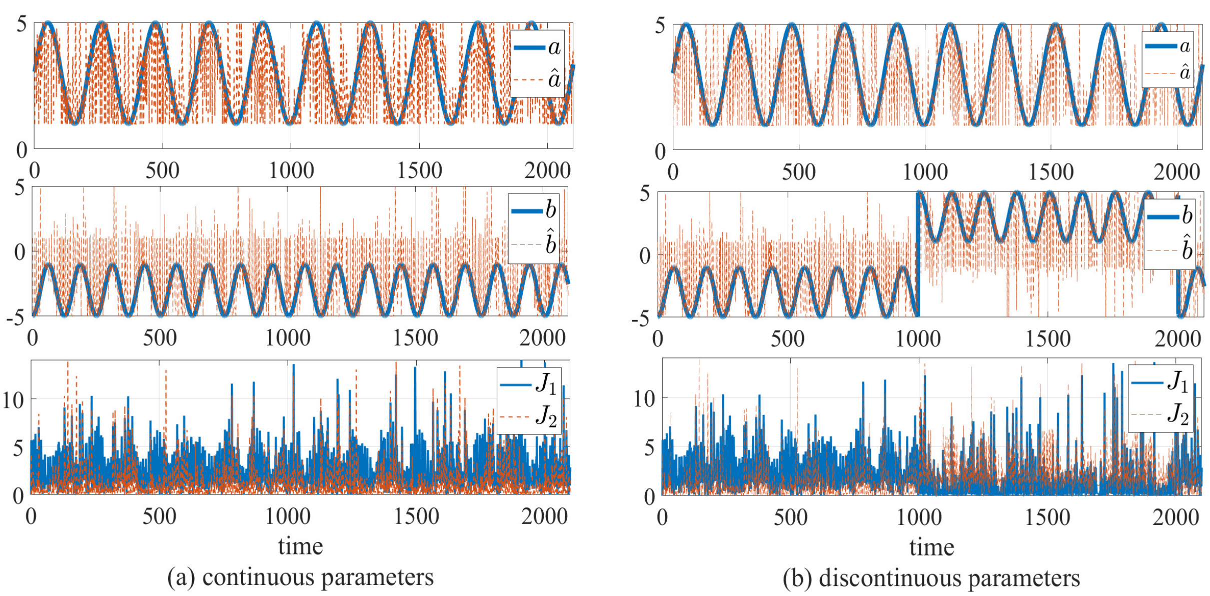



Figure 5.6: Plots for time-varying noise and no reference with $\bar{\lambda} = 0.3$. Column (a) shows the continuous parameter case whereas column (b) shows the discontinuous parameter case.

# Chapter 6

# Summary and Future Work

In this chapter we summarize the approach presented here and the results it yielded. Furthermore, we provide some future directions of research inside the domain of our thesis.

## 6.1 Summary of the Results

In this thesis, we have extended the work presented in [34] to show that the asymptotic trajectory tracking is achievable. The system in consideration is the first-order discrete-time plant, with an unknown sign of the input gain $b$. The proposed approach can be explained in the following way. Firstly, we cover the compact set of admissible plant parameters with a pair of convex sets. Next, we design two estimators (one for each convex set) using the modified version of the original projection algorithm; here, each estimator has the corresponding one-step-ahead controller. Finally, we choose the best suitable controller using a performance signal based switching rule at every time step.

Our proposed scheme has shown to have very desirable linear-like properties like exponential stability and convolution bounds on both the noise and the reference signal. Furthermore, in the presence of slow time-variations and unmodelled dynamics, the closed-loop system retains this desired behaviour. We also have examined the tracking ability of our proposed algorithm in the absence of noise and have proved asymptotic trajectory tracking for asymptotically strongly persistently exciting of order 2 reference signals, as well as for a fairly general reference signal with an unstable plant.

## 6.2 Future Work

We list some future avenues of exploration with regard to the work presented in this thesis.

- A natural extension of this work is to extend the proposed approach to a higher-order system.
- We would also like to prove asymptotic trajectory tracking for a general reference signal for both stable and unstable plants.
- In this research we have covered the set of plant parameters with the union of two convex sets, and therefore have two estimators. One possible extension is to cover the uncertainty set with more than two convex sets while retaining the linear-like behaviour and tracking ability of the closed-loop system.

# References


[1] B. D. O. Anderson. Exponential convergence and persistent excitation. In *1982 21st IEEE Conference on Decision and Control*, pages 12–17. IEEE, 1982. 45

[2] S. Baldi. Multiple estimation architecture in discrete-time adaptive mixing control. *Machines*, 1(1):33–49, 2013. 2

[3] S. Baldi, P. A. Ioannou, and E. B. Kosmatopoulos. Adaptive mixing control with multiple estimators. *International Journal of Adaptive Control and Signal Processing*, 26(8):800–820, 2012. 2

[4] E. F. Beckenbach and R. Bellman. *Inequalities*, volume 30. Springer Science & Business Media, 2012. 22, 26

[5] A. Feuer and A. S. Morse. Adaptive control of single-input, single-output linear systems. *IEEE Transactions on Automatic Control*, 23(4):557–569, 1978. 1

[6] M. Fu and B. Barmish. Adaptive stabilization of linear systems via switching control. *IEEE Transactions on Automatic Control*, 31(12):1097–1103, 1986. 2

[7] G. C. Goodwin, P. Ramadge, and P. Caines. Discrete-time multivariable adaptive control. *IEEE Transactions on Automatic Control*, 25(3):449–456, 1980. 1, 7

[8] G. C. Goodwin and K. S. Sin. Adaptive filtering prediction and control,(book) prentice-hall. *Englewood Cliffs*, 1984. 6, 7, 45

[9] J. P. Hespanha, D. Liberzon, and A. S. Morse. Hysteresis-based switching algorithms for supervisory control of uncertain systems. *Automatica*, 39(2):263–272, 2003. 2

[10] J. P. Hespanha, D. Liberzon, and A. S. Morse. Overcoming the limitations of adaptive control by means of logic-based switching. *Systems & Control Letters*, 49(1):49–65, 2003. 2

[11] P. A. Ioannou and K. S. Tsakalis. A robust direct adaptive controller. *IEEE Transactions on Automatic Control*, 31(11):1033–1043, 1986. 1

[12] G. Kreisselmeier. Adaptive control of a class of slowly time-varying plants. *Systems & Control Letters*, 8(2):97–103, 1986. 35

[13] G. Kreisselmeier and B. D. O. Anderson. Robust model reference adaptive control. *IEEE Transactions on Automatic Control*, 31(2):127–133, 1986. 1

[14] M. Kuipers and P. A. Ioannou. Multiple model adaptive control with mixing. *IEEE Transactions on Automatic Control*, 55(8):1822–1836, 2010. 2

[15] Y. Li and H-F Chen. Robust adaptive pole placement for linear time-varying systems. *IEEE Transactions on Automatic Control*, 41(5):714–719, 1996. 1

[16] R. H. Middleton and G. C. Goodwin. Adaptive control of time-varying linear systems. *IEEE Transactions on Automatic Control*, 33(2):150–155, 1988. 1

[17] R. H. Middleton, G. C. Goodwin, D. J. Hill, and D. Q. Mayne. Design issues in adaptive control. *IEEE Transactions on Automatic Control*, 33(1):50–58, 1988. 1

[18] D. E. Miller. Classical discrete-time adaptive control revisited: Exponential stabilization. In *2017 IEEE Conference on Control Technology and Applications (CCTA)*, pages 1975–1980. IEEE, 2017. 7

[19] D. E. Miller. A parameter adaptive controller which provides exponential stability: The first order case. *Systems & Control Letters*, 103:23–31, 2017. 2, 7, 17, 23

[20] D. E. Miller and E. J. Davison. An adaptive controller which provides lyapunov stability. *IEEE Transactions on Automatic Control*, 34(6):599–609, 1989. 2

[21] D. E. Miller and M. T. Shahab. Classical pole placement adaptive control revisited: linear-like convolution bounds and exponential stability. *Mathematics of Control, Signals, and Systems*, 30(4):19, 2018. 2, 6, 8, 13, 14, 23, 30, 36

[22] D. E. Miller and M. T. Shahab. Adaptive tracking with exponential stability and convolution bounds using vigilant estimation. *Mathematics of Control, Signals, and Systems*, 32(3):241–291, 2020. 2

[23] A. S. Morse. Global stability of parameter-adaptive control systems. *IEEE Transactions on Automatic Control*, 25(3):433–439, 1980. 1

[24] A. S. Morse. Supervisory control of families of linear set-point controllers-part i. exact matching. *IEEE Transactions on Automatic Control*, 41(10):1413–1431, 1996. 2

[25] A. S. Morse. Supervisory control of families of linear set-point controllers. 2. robustness. *IEEE Transactions on Automatic Control*, 42(11):1500–1515, 1997. 2

[26] A. S. Morse. A bound for the disturbance—to—tracking—error gain of a supervised set-point control system. In *Perspectives in Control*, pages 23–41. Springer, 1998. 2

[27] S. M. Naik, P. R. Kumar, and B. E. Ydstie. Robust continuous-time adaptive control by parameter projection. *IEEE Transactions on Automatic Control*, 37(2):182–197, 1992. 1

[28] K. S. Narendra and J. Balakrishnan. Improving transient response of adaptive control systems using multiple models and switching. *IEEE Transactions on Automatic Control*, 39(9):1861–1866, 1994. 2

[29] K. S. Narendra and J. Balakrishnan. Adaptive control using multiple models. *IEEE Transactions on Automatic Control*, 42(2):171–187, 1997. 2

[30] K. S. Narendra and Y.-H. Lin. Stable discrete adaptive control. *IEEE Transactions on Automatic Control*, 25(3):456–461, 1980. 1

[31] K. S. Narendra, Y.-H. Lin, and L. S. Valavani. Stable adaptive controller design, part ii: Proof of stability. *IEEE Transactions on Automatic Control*, 25(3):440–448, 1980. 1

[32] C. Rohrs, L. Valavani, M. Athans, and G. Stein. Robustness of continuous-time adaptive control algorithms in the presence of unmodeled dynamics. *IEEE Transactions on Automatic Control*, 30(9):881–889, 1985. 1

[33] M. T. Shahab. A new approach to multi-model adaptive control. 2020. 18

[34] M. T. Shahab and D. E. Miller. Multi-estimator based adaptive control which provides exponential stability: the first-order case. In *2018 IEEE Conference on Decision and Control (CDC)*, pages 2223–2228. IEEE, 2018. viii, 2, 3, 6, 13, 14, 18, 64, 67, 69, 70, 71, 72

[35] M. T. Shahab and D. E. Miller. Adaptive set-point regulation using multiple estimators. In *2019 IEEE 58th Conference on Decision and Control (CDC)*, pages 84–89. IEEE, 2019. 2, 6, 10

[36] M. T. Shahab and D. E. Miller. The inherent robustness of a new approach to adaptive control. In *2020 IEEE Conference on Control Technology and Applications (CCTA)*, pages 510–515. IEEE, 2020. 33, 34, 36, 37

[37] M. T. Shahab and D. E. Miller. Asymptotic tracking and linear-like behavior using multi-model adaptive control. *IEEE Transactions on Automatic Control*, 2021 (Accepted). 2

[38] K. S. Tsakalis and P. A. Ioannou. Adaptive control of linear time-varying plants: A new model reference controller structure. *IEEE Transactions on Automatic Control*, 34(10):1038–1046, 1989. 1

[39] C. Wen. A robust adaptive controller with minimal modifications for discrete time-varying systems. *IEEE Transactions on Automatic Control*, 39(5):987–991, 1994. 1

[40] C. Wen and D. J. Hill. Global boundedness of discrete-time adaptive control just using estimator projection. *Automatica*, 28(6):1143–1157, 1992. 1

[41] B. E. Ydstie. Stability of discrete model reference adaptive control—revisited. *Systems & Control Letters*, 13(5):429–438, 1989. 1

[42] B. E. Ydstie. Transient performance and robustness of direct adaptive control. *IEEE Transactions on Automatic Control*, 37(8):1091–1105, 1992. 1