A New Approach to Multi-Model Adaptive Control

by

Mohamad T. Shahab

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2020

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:	João P. Hespanha Professor Department of Electrical and Computer Engineering University of California, Santa Barbara
Supervisor:	Daniel E. Miller Professor Department of Electrical and Computer Engineering University of Waterloo
Internal Member:	Andrew J. Heunis Professor Department of Electrical and Computer Engineering University of Waterloo
Internal Member:	Christopher Nielsen Associate Professor Department of Electrical and Computer Engineering University of Waterloo
Internal-External Member:	Barış Fidan Associate Professor Department of Mechanical and Mechatronics Engineering University of Waterloo

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 with uncertain or time-varying parameters. A classical adaptive controller typically consists of a linear time-invariant (LTI) control law together with a tuning mechanism which adjusts its parameters. Usually, though not exclusively, discrete-time adaptive controllers provide only asymptotic stability and possibly bounded-noise bounded-state stability; neither exponential stability nor a bounded noise gain is typically proven. Recently it has been shown that if we employ a parameter estimator based on the original Projection Algorithm together with projecting the parameter estimates onto a given compact and convex set, then the adaptive controller guarantees linear-like closed-loop behavior: exponential stability, a bounded noise gain and a convolution bound on the exogenous inputs. In this thesis, the overarching objective is to show that we can prove these same desirable linear-like properties in a wide range of adaptive control problems without the convexity assumption: the main idea is to use multiple estimators and a switching algorithm. Indeed, we show that those properties arise in a surprisingly natural way.

We first prove a general result that exponential stability and a convolution bound on the closed-loop behavior can be leveraged to show tolerance to a degree of time-variations and unmodelled dynamics, i.e. such closed-loop properties guarantee robustness. After reviewing the original Projection Algorithm and introducing the reader to our slightly revised version, we turn our attention to controller design, with a focus on a non-convex set of plant uncertainty. As a starting point, we first consider first-order plants incorporating a simple switching algorithm. We then extend the approach to a class of nonlinear plants (which have stable zero dynamics); we consider both cases of convex and non-convex sets of parameter uncertainty. Afterwards, we turn to possibly non-minimum phase LTI plants; first we consider the stabilization problem for which we have two convex sets of uncertainty; then, we turn to the problem of tracking the sum of a finite number of sinusoids of known frequencies subject to an unknown plant order and a general compact set of uncertainty.

Acknowledgements

First, my sincerest gratitude is to my supervisor, Professor Daniel E. Miller; I thank him for giving me the opportunity to do interesting research work and for his excellent guidance and his patience during my doctoral study.

I would like also to thank Professor Andrew J. Heunis, Professor Christopher Nielsen and Professor Barış Fidan, for their support and for serving in my committee; I thank them for their feedback and comments.

I express my appreciation and thanks to Professor João P. Hespanha for serving as the thesis external examiner and for his valuable comments.

I would like to thank my friends in Waterloo for the companionship during our graduate studies. I am thankful to all of my friends from beyond Canada for their continuing friendship and for keeping in touch. I thank the many professors and teachers that I have learned from.

Finally, I am deeply grateful to my parents, Thohir Shahab and Shaikha Alaidros, for their continuing and unconditional support, sacrifice, love, and prayers; forever I am indebted to them for letting me be the person that I have become. I am grateful to my sisters, Shaima and Sukaina, for their love and for putting up with their big brother. I thank my family for bearing with me.

All praise be to Allah (God) Almighty; without His help and blessing, I would not have been able to keep on to seek knowledge.

Table of Contents

List of Figures

Chapter 1

Introduction

In this chapter, we first briefly talk about adaptive control. Next, we will motivate the results of the thesis through a detailed literature review and discussion. After that, objectives and contributions of the thesis are discussed. Then an outline of the rest of the thesis is provided, followed by introducing some notation that will be used throughout the thesis.

1.1 Adaptive Control

Adaptive control is an approach used to deal with systems with uncertain or time-varying parameters. Typically, an adaptive controller will consist of a linear time-invariant (LTI) control law together with a tuning algorithm which adjusts its parameters. In contrast, a robust controller is designed, according to some criteria, to be an LTI controller, dealing with uncertainty without the need of tuning or "adapting" its parameters. There are several reasons for using an adaptive controller, including

- if the goal is closed-loop stability, then adaptive controllers tend to tolerate more plant uncertainty and slowly changing operating conditions;
- if the goal is asymptotic tracking, then under suitable assumptions, an adaptive controller can tolerate a great deal of plant uncertainty, while a robust controller can tolerate none.

The main goals in adaptive control are stability and tracking. One of the popular approaches addressing the tracking goal is the One-Step-Ahead Adaptive Control problem;

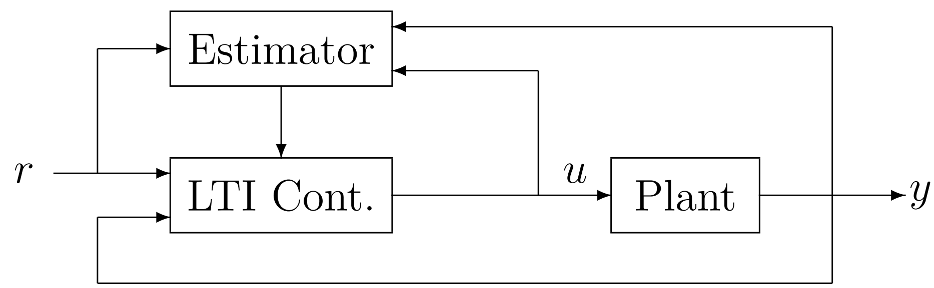


Figure 1.1: An Adaptive Control System.

the goal is to have the output of the plant track a reference signal asymptotically. One typical approach to achieving the stabilization goal is the Pole-Placement Adaptive Control problem; the goal here is to asymptotically have the closed-loop behavior according to prespecified closed-loop poles. These are among the most basic problems considered in adaptive control—see [21], [31], [30], [37], [53].

1.2 Background and Motivation

1.2.1 Early Results

The first general results of adaptive control came about around 1980, e.g. see [16], [20], [49], [59], [60]. However, these early results are not without issues. First of all, these controllers typically do not tolerate unmodeled dynamics, time-variations, and/or noise/disturbance very well—see the paper by Rohrs *et al.* [65]. Second of all, they put stringent assumptions on *a priori* information about the structure of the plant, e.g. order, delay/relative degree, etc.

Over the following two decades, an attempt was made to handle unmodeled dynamics, slow time-variations and noise/disturbance. A common approach was to make small controller design changes, such as the use of signal normalization, deadzones and σ-modifications, e.g. see [29], [34], [40], [41], [75]. An alternative approach imposes a convexity assumption on the set of parameter uncertainty, which is utilized in the estimation process to restrain the estimates of the plant parameters to the convex set, e.g. see [38], [55], [77], [78], [80], [81]. Although these controllers typically provide some tolerance to unmodelled dynamics and/or slow time-variations, in general, they only provide

- asymptotic stability and not exponential stability,
- and a bounded-noise bounded-state property, but not a bounded gain on the noise.

The goal of this thesis is to design adaptive controllers for which the closed-loop system exhibits highly desirable linear-like system properties, such as exponential stability, a bounded gain on the noise, and ideally a convolution bound on the input-output behavior. As far as the author is aware, in the classical approach to adaptive control a bounded gain on the noise is proven only in [81]; however, neither a crisp exponential bound on the effect of the initial condition nor a convolution bound on the closed-loop behavior is proven.

Remark 1.1. *While we can prove a form of exponential stability if a persistent excitation condition is satisfied, e.g. see [3], this places a stringent requirement on exogenous inputs, which we would like to avoid.*

1.2.2 Multiple Estimators and Multiple Controllers in Adaptive Control

A great deal of work focused on removing (or reducing) the assumptions on structural information of the plant, which typically included the following:

- the plant is minimum phase (this is waived if the goal is stability only);
- the plant order is known (or an upper bound on the plant order);
- the sign of the high-frequency gain is known.

There are various non-classical approaches to adaptive control that address these issues and provide some linear-like system properties, like those that include multiple estimators, multiple controller designs and/or a switching mechanism.

Pre-routed Switching

Some of the early works that tried to tackle each of the aforementioned restrictions are [17], [39], [45], [42], [54], [62], [79]; here, a *pre-routed* switching mechanism is employed to switch between a list of pre-designed candidate controllers. In Nussbaum [62], it is proven that knowledge of the sign of high-frequency gain is not needed in the first order case. In Mårtensson [39], it is proven that the knowledge of the order of an LTI stabilizing compensator is sufficient for adaptive stabilization. In Fu and Barmish [17], with plant parameters assumed to lie in a compact set, exponential stability is proven; this is extended in Miller and Davison [45]. In all of these results discussed above, a bounded gain on the noise is not proven. Also a side effect of the switching mechanism is that they typically yield poor transient behavior, poor noise tolerance, or a large control input.

Supervisory Control

Later, motivated by works discussed above and by robust control, more sophisticated logic-based switching approaches to adaptive control emerged. While an initial idea appeared in Middleton *et al.* [41], the powerful approach of *Supervisory Control* is introduced in Morse [50] and Morse [51]; this was then extended in Hespanha *et al.* [27], Hespanha *et al.* [26] and Hespanha *et al.* [25], as well as in [11] for the discrete-time setting. Here, a bank of candidate controllers is built; each is designed to stabilize a different "nominal" system model. A high-level *supervisor* is responsible for switching between the candidate controllers. The supervisor uses a *performance signal* or *monitoring signal*, typically a function of the estimation error, to assess the potential performance of each candidate controller. From time to time, the supervisor selects the controller corresponding to the "best" performance signal. Analysis of the switching behavior utilize various concepts, namely, *Hysteresis Switching*, see Morse *et al.* [52], and Hespanha and Morse [28], and *Dwell-time Switching*, see Morse [50] and Hespanha and Morse [24]. Stability results are proven and in certain circumstances a bounded gain on the noise can be proven, e.g. see Morse [51]. In Vu and Liberzon [76], supervisory control is applied to handle slowly time-varying plants; while stability results are proven handling time-variation, crisp bounds on closed-loop behavior are not provided, and convolution bounds are not proven.

In the above works, convexity of the uncertainty set is generally not required. However, a problem with these approaches is that, in many cases, there is a need to design a large number of candidate controllers. To this end, the complexity of the approach grows with the size of the plant uncertainty. This *covering* problem is still an ongoing research problem; a discussion about this issue can be found in [1], [2], [63], [15], [33]; this issue arises in other approaches discussed below.

Multiple Parameter Estimators

In Narendra and Balakrishnan [56] and Narendra and Balakrishnan [57], a combination of both switching and parameter estimation is applied. Multiple estimators provide parallel parameter estimates, and from time to time a switching algorithm selects which corresponding controller to be used. All estimators are essentially the same; they differ only in the initial values. Transient behavior is argued to be improved due to the possibility of closeness of one of the estimates to the correct value; however, exponential stability is not explicitly proven. In Narendra and Xiang [61], similar results are discussed when dealing with noise.

Mixing-based Approach

In a closely-related approach to the previous one, an approach utilizing multi-estimators along with multi-controllers is discussed in Kuipers and Ioannou [36], Baldi *et al.* [4], Baldi *et al.* [7], and Baldi and Ioannou [6]; this approach is called *Adaptive Mixing Control*. Here, multiple parameter estimates are calculated and a list of corresponding candidate controllers is constructed; the control law is a weighted combination of the candidate controllers, where a "mixing" signal determines the participating level of each; in [4] and [6], this mixing is determined by some switching logic.

A similar approach is found in Han and Narendra [23] and Narendra and Han [58]. It, too, utilizes a mixing approach, but is different from the above; the control law is, instead, designed using a weighted combination of the parameter estimates directly. The weights for this combination is also adaptively set, hence the approach is sometimes labeled *Second-level Adaptation*.

For some of the approaches discussed above, stability and tracking results are presented with some tolerance to noise and some unmodelled dynamics; while improvement in closed-loop behavior is shown, assumptions of convexity and knowledge of plant order are enforced. It is not clear how these approaches deal with plant parameter time-variation.

Falsification-based Approach

Another approach utilizes the multi-controller idea without any estimation; it is based on the concept of *Unfalsified Control*—see Safonov and Tsao [67]. It addresses the problem of detecting instability, or in other words, avoiding destabilizing controllers (see Stefanovic and Safonov [74]). Here also, a list of candidate controllers is available; the approach aims to switch any destabilizing controller out of the loop completely. Unlike other switching methods discussed earlier, the switching decision here is based solely on the system's input-output information—see Battistelli *et al.* [8] and Baldi *et al.* [5]. So this approach can be considered a data-driven "model-free" approach, sidestepping the issue of dealing with structural assumptions on the plant. The switching algorithms utilized here borrows similar ideas from ones in Supervisory Control. In Battistelli *et al.* [9] and Battistelli *et al.* [10], an interesting switching algorithm is designed to handle time-varying plants, as well. Stability results are proven, but there is no guarantee of good transient behavior, let alone a bounded noise gain.

Another approach employing a falsification idea, although inspired differently, can be found in Zhivoglyadov *et al.* [87] and Zhivoglyadov *et al.* [88]. It uses a localization-based switching to prove exponential stability with some tolerance to noise. One drawback of the approach is that the number of candidate controllers may be extremely large if the set of plant uncertainty is large; a bounded gain on the noise is not proven.

1.2.3 Linear-like Behavior in Adaptive Control

In all of the approaches discussed earlier, a linear-like convolution bound on the closed-loop behavior is not proven. More recently, an approach is provided, in both the one-step-ahead control setting by Miller [43], Miller and Shahab [47] and Miller and Shahab [48][1], and the pole-placement control setting by Miller [44] and Miller and Shahab [46], which guarantees a *linear-like* closed-loop behavior:

- exponential stability,
- a bounded gain on the noise,
- and a convolution bound on the exogenous signals,

which are clearly very desirable properties. As far as the author knows, such linear-like convolution bounds have never before been proven in the adaptive setting. The key idea is to employ, as part of the discrete-time adaptive controller, a **single parameter estimator** based on the original Projection Algorithm together with projection of the parameter estimates onto a given **compact and convex** set. In [43] and [46], the resulting convolution bound is leveraged to prove tolerance to a degree of time-variation and unmodelled dynamics.

[1]The author of this thesis contributed to the tracking results of [47] and [48]; these papers deal with the case of a convex set of uncertainty, which is outside the scope of the thesis, but a version of the tracking results appears in Chapter 5.

1.3 Objective and Contributions

The goal of this thesis is to extend the approach discussed in Section 1.2.3 to the non-convex setting. As discussed earlier, the requirement of convexity on the set of uncertainty is shown to play a crucial role in obtaining desirable closed-loop properties. Since convexity is a very restrictive requirement, the main objective of the thesis is to prove the same desirable linear-like closed-loop behavior in the case of a **non-convex** uncertainty set. To this end, we replace the assumption of a convex and compact set of parameter uncertainty with only the assumption of a compact set. The key idea of the proposed discrete-time adaptive control approach include:

- first we "cover" the compact set of admissible parameters by a finite number of convex sets;
- then we design a parameter estimator based on the original Projection Algorithm for each convex set,
- and finally we use a switching algorithm to switch between the corresponding control laws.

In Miller [43] and Miller and Shahab [46], it was shown that for the case of using a single estimator, if a convolution bound is proven, then the resulting adaptive controller is robust with respect to slow time-varying parameters and a degree of unmodelled dynamics. In this thesis, the first main result is to show this is true in a more general formulation, which allows for nonlinear plants and multi-estimator settings. This allows us to prove robustness in a modular fashion: we can focus on analyzing the ideal plant, knowing that robustness comes for free.

At this point, we look at the case of non-convex uncertainty sets:

- In the first-order setting, we consider the one-step-ahead adaptive control problem subject to a compact uncertainty set; we provide a switching mechanism and prove that the desired linear-like closed-loop properties are achieved.
- While we attempted to extend the results on the high order one-step-head adaptive control of Miller and Shahab [48] to the case of a non-convex uncertainty set, we did not succeed. However, we succeeded in extending the first-order approach to a special class of high order nonlinear systems with stable zero dynamics.

- At this point, we turn our attention to possibly non-minimum phase LTI plants. We first look at the simplest case: the stabilization problem subject to known plant order and two convex sets of uncertainty. We develop a switching mechanism and prove that the desired linear-like closed-loop properties are achieved.

- Last of all, in the context of possibly non-minimum phase LTI plants, we tackle the more difficult problem of tracking the sum of a finite number of sinusoids of known frequencies subject to an unknown plant order and a general compact set of uncertainty. Once again, we are able to develop a switching mechanism and prove that the desired linear-like closed-loop properties are achieved.

We now provide an outline for the rest of the thesis. In **Chapter 2**, we define what we mean by a convolution bound and then we prove robustness given a convolution bound on the closed-loop behavior as discussed above. In **Chapter 3**, we provide the details of the parameter estimation used throughout the thesis, based on the original Projection Algorithm. Then, as a showcase for the contribution of the whole thesis, we consider the case of one-step-ahead adaptive control of first-order plants in **Chapter 4**, which includes a simple switching algorithm. Then in **Chapter 5**, we provide the results of adaptive control of a special class of nonlinear plants; this chapter includes two parts: one where the uncertainty set is convex and one where it is not. In **Chapter 6** we consider the stabilization problem of possibly non-minimum phase plants given two convex sets. The penultimate chapter, **Chapter 7**, includes the main Multi-Model Adaptive Control and Tracking approach dealing with possibly non-minimum phase plants subject to an unknown plant order and a general compact uncertainty set. Finally, a summary of results and a discussion about the limitations of our approach and future directions are provided in **Chapter 8**.

We would like to point out to the reader that, as of the time of writing this thesis, multiple papers have been published that include versions of the results of this thesis:

- in the conference paper Shahab and Miller [68], we have proven the desired results in the context of one-step-ahead adaptive control of first-order plants subject only to a compact uncertainty set;

- in the journal paper Miller and Shahab [46], we have proven the desired stability results for a possibly non-minimum phase plant subject to a known order and two compact and convex uncertainty sets;

- in the conference paper Shahab and Miller [70], we have proven the desired results in the context of one-step-ahead adaptive control of first-order nonlinear plants;

- in the conference paper Shahab and Miller [69], we have proven desired stability results as well as step tracking, for a possibly non-minimum phase plant subject to a known order and compact uncertainty set;
- in the journal paper Shahab and Miller [71] (under review), we have proven stability results as well as tracking of the sum of a finite number of sinusoids of known frequencies, for a possibly non-minimum phase plant subject to an unknown plant order and a general compact set of uncertainty;
- in the conference submission Shahab and Miller [72], we have proven a condensed version of Chapter 2.

1.4 Notation and Mathematical Preliminaries

We use standard notation throughout the thesis:

- $\mathbb{R}$ denotes the set of real numbers.
- $\mathbb{R}^+$ denotes the set of nonnegative real numbers.
- $\mathbb{Z}$ denotes the set of integers.
- $\mathbb{Z}^+$ denotes the set of nonnegative integers.
- $\mathbb{N}$ denotes the set of natural numbers.
- $\lceil\cdot\rceil$ denotes the ceiling function: for any $x \in \mathbb{R}$, $\lceil x \rceil := \min\{z \in \mathbb{Z} : z \geq x\}$.
- $\|\cdot\|$, the subscript-less default norm notation, denotes the Euclidean-norm of a vector and the induced norm of a matrix.
- For a square matrix A, let $\det\big(A\big)$ denotes the determinant of A.
- $\mathbb{S}(\mathbb{R}^{p\times q})$ denotes the set of $\mathbb{R}^{p\times q}$-valued sequences.
- $\boldsymbol{\ell}_\infty(\mathbb{R}^{p\times q})$ denotes the set of $\mathbb{R}^{p\times q}$-valued *bounded* sequences.
- $\boldsymbol{\ell}_\infty$ denotes the special case of $\boldsymbol{\ell}_\infty(\mathbb{R})$.

- For a signal $s \in \boldsymbol{\ell}_\infty$, define the ∞-norm by

$$\|s\|_\infty := \sup_{t \in \mathbb{Z}} |s(t)|.$$

- If $\Omega \subset \mathbb{R}^p$ is a bounded set, we define the notation

$$\|\Omega\| := \sup_{x \in \Omega} \|x\|.$$

- $\mathbf{e}_j \in \mathbb{R}^p$ denotes the normal vector of appropriate length p defined as

$$\mathbf{e}_j := \begin{bmatrix} \overbrace{0 \quad \cdots \quad 0}^{j-1 \text{ elements}} & 1 & 0 & \cdots & 0 \end{bmatrix}^\top.$$

- $\mathbf{0}_{p \times q}$ denotes the $p \times q$ matrix whose entries are all zeros.
- $\mathbf{0}_p$ denotes the $p \times 1$ vector whose entries are all zeros.
- I_p denotes the identity matrix of size p.
- In general, signals are represented in the time-domain by small letters, e.g. $s(t)$, and in the z-transform domain by capital letters, e.g. $S(z)$.
- We say that a function $\Gamma : \mathbb{R}^p \to \mathbb{R}^q$ has a *bounded gain* if there exists a $\nu > 0$ such that for all $x \in \mathbb{R}^p$, we have $\|\Gamma(x)\| \leq \nu \|x\|$; the smallest such ν is the gain, and is denoted by $\|\Gamma\|$.
- For an index set $\mathcal{I} = \{1, 2, \ldots, m\}$ and $J_i \geq 0$ ($i \in \mathcal{I}$), in computing

$$\operatorname*{argmin}_{i \in \mathcal{I}} J_i,$$

  it could very well be that there is more than one value $i \in \mathcal{I}$ which achieves the minimum. **In such a case, we (somewhat arbitrarily) choose the smallest such index.**

- Throughout the thesis, we will utilize the "inequality of arithmetic and geometric means": for nonnegative numbers $h_1, h_2, h_3, \ldots, h_M$, we have that

$$\left[\prod_{j=1}^{M} h_j\right]^{\frac{1}{M}} \leq \frac{1}{M} \sum_{j=1}^{M} h_j \quad \Leftrightarrow \quad \prod_{j=1}^{M} h_j \leq \left[\frac{1}{M} \sum_{j=1}^{M} h_j\right]^M.$$

Chapter 2

Robustness and Convolution Bounds

2.1 Introduction

In control system design, a common requirement is that the closed-loop system not only be stable, but also be robust, in the sense that the desired closed-loop properties are maintained, at the very least, in the presence of small time-variations in the plant parameters and a small amount of unmodelled dynamics. Of course, if the plant and controller are both linear and time-invariant, and the desired objective is closed-loop stability, then such robustness follows from the Small Gain Theorem [82] and the study of time-varying linear systems [14]. On the other hand, if either the plant or controller is nonlinear, this is often not the case and/or it is not easy to prove.

Indeed, one special class of nonlinear controllers is that of adaptive controllers. In general, as mentioned earlier, adaptive controllers provide asymptotic stability but not exponential stability, with no bounded gain on the noise, let alone a convolution bound. However, it has been proven (see [43], [46]) that if discrete-time adaptive control is carried out using the original projection algorithm, then exponential stability and a convolution bound on the closed-loop behavior can be proven; hence, the closed-loop system acts 'linear-like'. It is shown in the first-order one-step-ahead case [43] and the pole-placement case with a single estimator [46], that this approach is *robust*; this is **proven in a modular fashion**—we leverage the exponential stability and the convolution bounds proven for the nominal plant model without reopening its proof. This differs markedly from the approach that most work on robust adaptive control adopt: there one proves robustness by taking the proof for the ideal case and creating a more complicated version with a time-varying plant with some unmodelled dynamics added. The goal of this chapter is to prove that this *modularity* property holds in a very general adaptive control setting; modularity is a highly desirable property, since it allows us to focus on analyzing the ideal plant model, knowing that robustness will come for free.

To this end, here we consider a class of finite-dimensional, nonlinear plant and adaptive controller combinations; if exponential stability and a convolution bound holds, then we prove that tolerance to small time-variations in the plant parameters and a small amount of unmodelled dynamics follows[1]. An immediate application of this result is to prove robustness of the proposed adaptive controllers presented in the following chapters of this thesis; this includes approaches which use multi-estimators, and it allows us to focus on the ideal plant model in the analysis.

2.2 The Setup

Here the nominal plant is multi-input multi-output[2] with finite memory and an additive disturbance, such that the uncertain plant parameter enters linearly. To this end, with an output $y(t) \in \mathbb{R}^r$, an input $u(t) \in \mathbb{R}^m$, a disturbance $w(t) \in \mathbb{R}^r$, a modeling parameter of

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

and a vector of input-output data of the form

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

we consider the plant

$$y(t+1) \quad = \quad \theta^{*\top} f\Big(\vartheta(t)\Big) + w(t), \quad \vartheta(t_0) = \vartheta_0; \tag{2.1}$$

[1]A version of this chapter was submitted as a conference paper [72].

[2]This model is more general than we need throughout this thesis, but the cost of this is minimal.

we assume that $f : \mathbb{R}^{n_y \cdot r + n_u \cdot m} \to \mathbb{R}^p$ has a bounded gain and that $\mathcal{S}$ is a bounded set; both requirements are reasonable given that we will require uniform bounds in our analysis. We represent this system by the pair $\left(f, \mathcal{S}\right)$.

Here we consider a large class of controllers which subsumes LTI ones as well as a large class of adaptive ones. To this end, we consider a 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 $r(t) \in \mathbb{R}^r$ (typically a reference signal), together with equations of the form

$$z(t+1) = g_1\left(z(t), \hat{\theta}(t), \vartheta(t), r(t)\right), \quad z(t_0) = z_0 \tag{2.2a}$$

$$\hat{\theta}(t+1) = g_2\left(z(t), \hat{\theta}(t), \vartheta(t), r(t)\right), \quad \hat{\theta}(t_0) = \theta_0 \tag{2.2b}$$

$$u(t) = h\left(z(t), \hat{\theta}(t), \vartheta(t-1), y(t), r(t)\right). \tag{2.2c}$$

With $\Omega \subset \mathbb{R}^{l_2}$ a bounded set, we assume that

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

i.e. if $\hat{\theta}$ is initialized in Ω, then it remains in Ω throughout.

Remark 2.1. *This class subsumes finite-dimensional LTI controllers: simply set $l_2 = 0$ so that the sub-state $\hat{\theta}(t)$ disappears, and make the functions g_1 and h be linear.*

Remark 2.2. *This class subsumes many adaptive controllers: simply set $l_1 = 0$ and let $\hat{\theta}(t)$ be the state of a parameter estimator constrained to the set Ω. If we are using multiple estimators, then the size of $\hat{\theta}(t)$ is typically larger than that of θ^*.*

We now provide a definition of the desired **linear-like** closed-loop property:

Definition 2.1. We say that (2.2) **provides exponential stability and a convolution bound for** $\big(f, \mathcal{S}\big)$ with gain $c \geq 1$ and decay rate $\lambda \in (0,1)$ if, for every $\theta^* \in \mathcal{S}$, $t_0 \in \mathbb{Z}$, $\vartheta_0 \in \mathbb{R}^{n_y \cdot r + n_u \cdot m}$, $z_0 \in \mathbb{R}^{l_1}$, $\theta_0 \in \Omega \subset \mathbb{R}^{l_2}$, $w \in \mathbb{S}(\mathbb{R}^r)$ and $r \in \mathbb{S}(\mathbb{R}^r)$, when (2.2) is applied to (2.1), the following holds:

$$\left\| \begin{bmatrix} \vartheta(t) \\ z(t) \end{bmatrix} \right\| \leq c\lambda^{t-\tau} \left\| \begin{bmatrix} \vartheta(\tau) \\ z(\tau) \end{bmatrix} \right\| + \sum_{j=\tau}^{t-1} c\lambda^{t-j-1}(\|r(j)\| + \|w(j)\|) + c\|r(t)\|, \qquad t \geq \tau \geq t_0. \tag{2.3}$$

Remark 2.3. *In this definition there is no mention of* $\hat{\theta}(t)$*, since it is of secondary interest and constrained to lie in the bounded set* Ω*.*

2.3 Tolerance to Time-Variation

We now consider plants with a possibly time-varying parameter vector $\theta^*(t)$ instead of a static θ^*:

$$y(t+1) = \theta^*(t)^\top f\Big(\vartheta(t)\Big) + w(t), \quad \vartheta(t_0) = \vartheta_0. \tag{2.4}$$

Definition 2.2. With $c_0 \geq 0$ and $\epsilon > 0$ let $\mathbb{S}(\mathcal{S}, c_0, \epsilon)$ denote the subset of $\boldsymbol{\ell}_\infty(\mathbb{R}^{p \times r})$ whose elements θ^* satisfy:

- $\theta^*(t) \in \mathcal{S}$ 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-variation model encompasses both slow variations ($c_0 = 0$) and/or occasional jumps ($c_0 \neq 0$); this class is well-known in the adaptive control literature, e.g. see [35].

We can extend Definition 2.1 in a natural way to handle time-variations.

Definition 2.3. We say that (2.2) **provides exponential stability and a convolution bound for** $\big(f, \mathbb{S}(\mathcal{S}, c_0, \epsilon)\big)$ with gain $c \geq 1$ and decay rate $\lambda \in (0,1)$ if, for every $\theta^* \in \mathbb{S}(\mathcal{S}, c_0, \epsilon)$, $t_0 \in \mathbb{Z}$, $\vartheta_0 \in \mathbb{R}^{n_y \cdot r + n_u \cdot m}$, $z_0 \in \mathbb{R}^{l_1}$, $\theta_0 \in \Omega \subset \mathbb{R}^{l_2}$, $w \in \mathbb{S}(\mathbb{R}^r)$ and $r \in \mathbb{S}(\mathbb{R}^r)$, when (2.2) is applied to (2.4), the following holds:

$$\left\| \begin{bmatrix} \vartheta(t) \\ z(t) \end{bmatrix} \right\| \leq c\lambda^{t-\tau} \left\| \begin{bmatrix} \vartheta(\tau) \\ z_\tau) \end{bmatrix} \right\| + \sum_{j=\tau}^{t-1} c\lambda^{t-j-1}(\|r(j)\| + \|w(j)\|) + c\|r(t)\|, \qquad t \geq \tau \geq t_0. \tag{2.5}$$

We now will show that if a controller (2.2) provides exponential stability and a convolution bound for the plant (2.1), then the same will be true for the time-varying plant (2.4), as long as ϵ is small enough. We consider two cases: one where there is a desired decay rate, and one where there is not.

Theorem 2.1. *Suppose that the controller (2.2) provides exponential stability and a convolution bound for $\big(f, \mathcal{S}\big)$ 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 a $c_1 \geq c$ and $\epsilon > 0$ so that (2.2) provides exponential stability and a convolution bound for $\big(f, \mathbb{S}(\mathcal{S}, c_0, \epsilon)\big)$ with gain c_1 and decay rate λ_1.*

Remark 2.4. *The following proof is based, in part, on the proof of Theorem 2 of [46], which deals with a much simpler setup.*

Proof of Theorem 2.1. Suppose the controller (2.2) provides exponential stability and a convolution bound for (2.1) with gain $c \geq 1$ and a decay rate of λ. Fix $\lambda_1 \in (\lambda, 1)$ and $c_0 > 0$; let $t_0 \in \mathbb{Z}$, $\vartheta_0 \in \mathbb{R}^{n_y \cdot r + n_u \cdot m}$, $z_0 \in \mathbb{R}^{l_1}$, $\theta_0 \in \Omega$, $w \in \mathbb{S}(\mathbb{R}^r)$ and $r \in \mathbb{S}(\mathbb{R}^r)$ be arbitrary.

Now fix $m \in \mathbb{N}$ to be any number satisfying

$$m \geq \frac{\ln(c) + \frac{4c_0 c\|f\|}{\lambda_1 - \lambda}\left[\ln\left(1 + 2c\|f\|\|\mathcal{S}\|\right) + \ln(2) - \ln(\lambda + \lambda_1)\right]}{\ln(2\lambda_1) - \ln(\lambda + \lambda_1)},$$

(the rationale for this choice will be more clear shortly), and set

$$\epsilon = \frac{c_0}{m^2};$$

let $\theta^* \in \mathbb{S}(\mathcal{S}, c_0, \epsilon)$ be arbitrary and apply the controller (2.2) to the time-varying plant (2.4). To proceed, we analyze the closed-loop system behavior on intervals of length m, which we further analyze in groups of m^2.

To proceed, let $\bar{t} \geq t_0$ be arbitrary. Define a sequence $\{\bar{t}_i\}$ by

$$\bar{t}_i = \bar{t} + im, \qquad i \in \mathbb{Z}^+.$$

We can rewrite the time-varying plant as

$$y(t+1) = \theta^*(\bar{t}_i)^\top f(\vartheta(t)) + w(t) + \underbrace{\Big[\theta^*(t) - \theta^*(\bar{t}_i)\Big]^\top f(\vartheta(t))}_{=:\tilde{n}_i(t)}, \qquad t \in [\bar{t}_i, \bar{t}_{i+1}).$$

On the interval $[\bar{t}_i, \bar{t}_{i+1}]$, we can regard the plant as time-invariant, but with an extra disturbance; so by hypothesis,

$$\left\| \begin{bmatrix} \vartheta(t) \\ z(t) \end{bmatrix} \right\| \leq c\lambda^{t-\bar{t}_i} \left\| \begin{bmatrix} \vartheta(\bar{t}_i) \\ z(\bar{t}_i) \end{bmatrix} \right\| + \sum_{j=\bar{t}_i}^{t-1} c\lambda^{t-j-1}(\|r(j)\| + \|w(j)\| + \|\tilde{n}_i(j)\|) + c\|r(t)\|,$$
$$t \in [\bar{t}_i, \bar{t}_{i+1}], \; i \in \mathbb{Z}^+. \tag{2.6}$$

To analyze this difference inequality, we first construct an associated difference equation:

$$\psi(t+1) = \lambda\psi(t) + \|r(t)\| + \|w(t)\| + \|\tilde{n}_i(t)\|, \qquad t \in [\bar{t}_i, \bar{t}_{i+1}),$$

with an initial condition of

$$\psi(\bar{t}_i) = \left\| \begin{bmatrix} \vartheta(\bar{t}_i) \\ z(\bar{t}_i) \end{bmatrix} \right\|.$$

Using the fact that $c \geq 1$, it is straightforward to prove that

$$\left\| \begin{bmatrix} \vartheta(t) \\ z(t) \end{bmatrix} \right\| \leq c\psi(t) + c\|r(t)\|, \qquad t \in [\bar{t}_i, \bar{t}_{i+1}]. \tag{2.7}$$

Now we analyze this equation for $i = 0, 1, \ldots, m-1$.

Case 1: $\|\tilde{n}_i(t)\| \leq \frac{\lambda_1 - \lambda}{2c}\|\vartheta(t)\|$ for all $t \in [\bar{t}_i, \bar{t}_{i+1})$.

Using the above bound (2.7) and the fact that $\lambda_1 - \lambda \in (0, 1)$, we obtain

$$\begin{aligned} \psi(t+1) &= \lambda\psi(t) + \|r(t)\| + \|w(t)\| + \|\tilde{n}_i(t)\| \\ &\leq \lambda\psi(t) + \|r(t)\| + \|w(t)\| + \tfrac{\lambda_1-\lambda}{2c}\|\vartheta(t)\| \\ &\leq \lambda\psi(t) + \|r(t)\| + \|w(t)\| + \tfrac{\lambda_1-\lambda}{2}\left(\psi(t) + \|r(t)\|\right) \\ &\leq \tfrac{\lambda_1+\lambda}{2}\psi(t) + 2\|r(t)\| + \|w(t)\|, \qquad t \in [\bar{t}_i, \bar{t}_{i+1}), \end{aligned} \tag{2.8}$$

which means that

$$|\psi(t)| \leq \left(\tfrac{\lambda_1+\lambda}{2}\right)^{t-\bar{t}_i} |\psi(\bar{t}_i)| + \sum_{j=\bar{t}_i}^{t-1} \left(\tfrac{\lambda_1+\lambda}{2}\right)^{t-j-1} \left(2\|r(j)\| + \|w(j)\|\right), \qquad t = \bar{t}_i, \bar{t}_i+1, \ldots, \bar{t}_{i+1}. \tag{2.9}$$

This, in turn, implies that there exists $c_2 \geq 2c$ so that

$$\begin{aligned} \left\| \begin{bmatrix} \vartheta(\bar{t}_{i+1}) \\ z(\bar{t}_{i+1}) \end{bmatrix} \right\| \leq c \left(\tfrac{\lambda_1+\lambda}{2}\right)^m \left\| \begin{bmatrix} \vartheta(\bar{t}_i) \\ z(\bar{t}_i) \end{bmatrix} \right\| + \\ \sum_{j=\bar{t}_i}^{\bar{t}_{i+1}-1} c_2 \left(\tfrac{\lambda_1+\lambda}{2}\right)^{\bar{t}_{i+1}-j-1} (\|r(j)\| + \|w(j)\|) + c_2 \|r(\bar{t}_{i+1})\|. \end{aligned} \tag{2.10}$$

Case 2: $\|\tilde{n}_i(t)\| > \frac{\lambda_1-\lambda}{2c} \|\vartheta(t)\|$ for some $t \in [\bar{t}_i, \bar{t}_{i+1})$.

Since $\theta^*(t) \in \mathcal{S}$ for $t \geq t_0$, we see that

$$\|\tilde{n}_i(t)\| \leq 2\|\mathcal{S}\| \|f(\vartheta(t))\| \leq 2\|f\| \|\mathcal{S}\| \|\vartheta(t)\|, \qquad t \in [\bar{t}_i, \bar{t}_{i+1}).$$

This means that

$$\begin{aligned} \psi(t+1) &= \lambda\psi(t) + \|r(t)\| + \|w(t)\| + \|\tilde{n}_i(t)\| \\ &\leq \lambda\psi(t) + \|r(t)\| + \|w(t)\| + 2\|f\| \|\mathcal{S}\| \|\vartheta(t)\| \\ &\leq \underbrace{(1 + 2c\|f\| \|\mathcal{S}\|)}_{=:\gamma_3} \psi(t) + (1 + 2c\|f\| \|\mathcal{S}\|) \|r(t)\| + \|w(t)\|, \qquad t \in [\bar{t}_i, \bar{t}_{i+1}), \end{aligned} \tag{2.11}$$

which means that

$$|\psi(t)| \leq \gamma_3^{t-\bar{t}_i} |\psi(\bar{t}_i)| + \sum_{j=\bar{t}_i}^{t-1} \gamma_3^{t-j-1} \left(\gamma_3 \|r(j)\| + \|w(j)\|\right), \qquad t = \bar{t}_i, \bar{t}_i+1, \ldots, \bar{t}_{i+1}. \tag{2.12}$$

Setting $t = \bar{t}_{i+1}$ and using (2.7) yields

$$\left\| \begin{bmatrix} \vartheta(\bar{t}_{i+1}) \\ z(\bar{t}_{i+1}) \end{bmatrix} \right\| \leq c\gamma_3^m \left\| \begin{bmatrix} \vartheta(\bar{t}_i) \\ z(\bar{t}_i) \end{bmatrix} \right\| + \sum_{j=\bar{t}_i}^{\bar{t}_{i+1}-1} c\gamma_3^{\bar{t}_{i+1}-j-1} \left(\gamma_3 \|r(j)\| + \|w(j)\|\right) + c\|r(\bar{t}_{i+1})\|$$

$$\le c\gamma_3^m \left\| \begin{bmatrix} \vartheta(\bar{t}_i) \\ z(\bar{t}_i) \end{bmatrix} \right\| +$$

$$c\gamma_3 \left(\tfrac{2\gamma_3}{\lambda_1+\lambda}\right)^m \sum_{j=\bar{t}_i}^{\bar{t}_{i+1}-1} \left(\tfrac{\lambda_1+\lambda}{2}\right)^{\bar{t}_{i+1}-j-1} (\|r(j)\| + \|w(j)\|) + c\|r(\bar{t}_{i+1})\|. \quad (2.13)$$

This completes Case 2.

At this point we combine Cases 1 and 2. We would like to analyze m intervals of length m. On the interval $[\bar{t}, \bar{t}+m^2]$, there are m subintervals of length m; furthermore, because of the choice of ϵ we have that

$$\sum_{j=\bar{t}}^{\bar{t}+m^2-1} \|\theta^*(j+1) - \theta^*(j)\| \le c_0 + \epsilon m^2 \le 2c_0.$$

It is easy to see that there are at most $N_1 := \frac{4c_0 c\|f\|}{\lambda_1-\lambda}$ subintervals which fall into the category of Case 2, with the remainder falling into the category of Case 1; it is clear from the formula for m that $m > N_1$. If we use (2.10) and (2.13) to analyze the behavior of the closed-loop system on the interval $[\bar{t}, \bar{t}+m^2]$, we end up with a crude bound of

$$\left\| \begin{bmatrix} \vartheta(\bar{t}+m^2) \\ z(\bar{t}+m^2) \end{bmatrix} \right\| \le c^m \gamma_3^{N_1 m} \left(\tfrac{\lambda_1+\lambda}{2}\right)^{m(m-N_1)} \left\| \begin{bmatrix} \vartheta(\bar{t}) \\ z(\bar{t}) \end{bmatrix} \right\| +$$

$$2m \left(\tfrac{2\gamma_3}{\lambda_1+\lambda}\right)^m (c_2\gamma_3^{m+1})^m \left(\tfrac{2}{\lambda_1+\lambda}\right)^{(m+1)m} \sum_{j=\bar{t}}^{\bar{t}+m^2-1} \left(\tfrac{\lambda_1+\lambda}{2}\right)^{\bar{t}+m^2-j-1} (\|r(j)\| + \|w(j)\|) +$$

$$c_2\|r(\bar{t}+m^2)\|. \quad (2.14)$$

From the choice of m above, it is easy to show that

$$m^2 \ln\left(\tfrac{2\lambda_1}{\lambda_1+\lambda}\right) \ge m\ln(c) + N_1 m \ln(\gamma_3) + N_1 m \ln\left(\tfrac{2}{\lambda+\lambda_1}\right);$$

this immediately implies that

$$c^m \gamma_3^{N_1 m} \left(\frac{2}{\lambda+\lambda_1}\right)^{N_1 m} \le \left(\frac{2\lambda_1}{\lambda_1+\lambda}\right)^{m^2} \quad \Leftrightarrow \quad c^m \gamma_1^{N_1 m} \left(\frac{\lambda_1+\lambda}{2}\right)^{m(m-N_1)} \le \lambda_1^{m^2}.$$

Since $\frac{\lambda_1+\lambda}{2} < \lambda_1$, it follows from (2.14) that there exists a constant γ_4 so that

$$\left\| \begin{bmatrix} \vartheta(\bar{t}+m^2) \\ z(\bar{t}+m^2) \end{bmatrix} \right\| \le \lambda_1^{m^2} \left\| \begin{bmatrix} \vartheta(\bar{t}) \\ z(\bar{t}) \end{bmatrix} \right\| + \gamma_4 \sum_{j=\bar{t}}^{\bar{t}+m^2-1} \lambda_1^{\bar{t}+m^2-j-1} \left(\|r(j)\| + \|w(j)\|\right) + \gamma_4 \|r(\bar{t}+m^2)\|. \tag{2.15}$$

Now let $\tau \ge t_0$ be arbitrary. By setting $\bar{t} = \tau, \tau+m^2, \tau+2m^2, \ldots$, in succession, it follows from (2.15) that

$$\left\| \begin{bmatrix} \vartheta(\tau+qm^2) \\ z(\tau+qm^2) \end{bmatrix} \right\| \le \lambda_1^{qm^2} \left\| \begin{bmatrix} \vartheta(\tau) \\ z(\tau) \end{bmatrix} \right\| + \\ \gamma_4 \sum_{j=\tau}^{\tau+qm^2-1} \lambda_1^{\tau+qm^2-j-1} \left(\|r(j)\| + \|w(j)\|\right) + \gamma_4 \|r(\tau+qm^2)\|, \quad q \in \mathbb{Z}^+. \tag{2.16}$$

So $\begin{bmatrix} \vartheta(t) \\ z(t) \end{bmatrix}$ is well-behaved at $t = \tau, \tau+m^2, \tau+2m^2, \ldots$, etc; we can use (2.9) of Case 1, (2.12) of Case 2 and (2.7) to prove that nothing untoward happens between these times. We conclude that there exists a constant γ_5 so that

$$\left\| \begin{bmatrix} \vartheta(t) \\ z(t) \end{bmatrix} \right\| \le \gamma_5 \lambda_1^{t-\tau} \left\| \begin{bmatrix} \vartheta(\tau) \\ z(\tau) \end{bmatrix} \right\| + \gamma_5 \sum_{j=\tau}^{t-1} \lambda_1^{t-j-1} \left(\|r(j)\| + \|w(j)\|\right) + \gamma_5 \|r(t)\|, \quad t \ge \tau. \tag{2.17}$$

Since $\tau \ge t_0$ is arbitrary, the desired bound is proven. ∎

A careful examination of the above proof reveals that $\epsilon \to 0$ as $c_0 \to 0$ and as $c_0 \to \infty$. If we do not care about the decay rate, then we can remove this drawback in the following result.

Theorem 2.2. *Suppose that the controller (2.2) provides exponential stability and a convolution bound for $\left(f, \mathcal{S}\right)$ with gain $c \ge 1$ and decay rate $\lambda \in (0,1)$. Then there exists an $\epsilon > 0$ such that for every $c_0 \ge 0$, there exist $\lambda_* \in (0,1)$ and $\gamma > 0$ so that (2.2) provides exponential stability and a convolution bound for $\left(f, \mathbb{S}(\mathcal{S}, c_0, \epsilon)\right)$ with gain γ and decay rate λ_*.*

Proof of Theorem 2.2. Suppose the controller (2.2) provides a convolution bound for (2.1) with gain $c \ge 1$ and a decay rate of λ. Fix $\lambda_1 \in (\lambda, 1)$; let $t_0 \in \mathbb{Z}$, $\vartheta_0 \in \mathbb{R}^{n_y \cdot r + n_u \cdot m}$,

$z_0 \in \mathbb{R}^{l_1}$, $\theta_0 \in \Omega$, $w \in \mathbb{S}(\mathbb{R}^r)$ and $r \in \mathbb{S}(\mathbb{R}^r)$ be arbitrary. The goal is to prove that for a small-enough ϵ, the controller (2.2) provides exponential stability and a convolution bound for $\big(f, \mathbb{S}(\mathcal{S}, c_0, \epsilon)\big)$ for every $c_0 \geq 0$. So at this point we will analyze the closed-loop system for an arbitrary $\epsilon > 0$, $c_0 \geq 0$, and $\theta^* \in \mathbb{S}(\mathcal{S}, c_0, \epsilon)$.

To proceed, let $\bar{t} \geq t_0$ be arbitrary. For $m \in \mathbb{N}$, we will first analyze closed-loop behavior on intervals of length m; define a sequence $\{\bar{t}_i\}$ by

$$\bar{t}_i = \bar{t} + im, \qquad i \in \mathbb{Z}^+.$$

We can rewrite the time-varying plant as

$$y(t+1) = \theta^*(\bar{t}_i)^\top f(\vartheta(t)) + w(t) + \underbrace{\Big[\theta^*(t) - \theta^*(\bar{t}_i)\Big]^\top f(\vartheta(t))}_{=:\tilde{n}_i(t)}, \qquad t \in [\bar{t}_i, \bar{t}_{i+1}).$$

On the interval $[\bar{t}_i, \bar{t}_{i+1}]$, we regard the plant as time-invariant, but with an extra disturbance: so we obtain

$$\left\| \begin{bmatrix} \vartheta(t) \\ z(t) \end{bmatrix} \right\| \leq c\lambda^{t-\bar{t}_i} \left\| \begin{bmatrix} \vartheta(\bar{t}_i) \\ z(\bar{t}_i) \end{bmatrix} \right\| + \sum_{j=\bar{t}_i}^{t-1} c\lambda^{t-j-1}(\|r(j)\| + \|w(j)\| + \|\tilde{n}_i(j)\|) + c\|r(t)\|,$$
$$t \in [\bar{t}_i, \bar{t}_{i+1}],\ i \in \mathbb{Z}^+. \tag{2.18}$$

Using the same idea as in the proof of Theorem 2.1, we define the difference equation

$$\psi(t+1) = \lambda\psi(t) + \|r(t)\| + \|w(t)\| + \|\tilde{n}_i(t)\|, \qquad t \in [\bar{t}_i, \bar{t}_{i+1})$$

with

$$\psi(\bar{t}_i) = \left\| \begin{bmatrix} \vartheta(\bar{t}_i) \\ z(\bar{t}_i) \end{bmatrix} \right\|;$$

it follows that

$$\left\| \begin{bmatrix} \vartheta(t) \\ z(t) \end{bmatrix} \right\| \leq c\psi(t) + c\|r(t)\|, \qquad t \in [\bar{t}_i, \bar{t}_{i+1}]. \tag{2.19}$$

Case 1: $\|\tilde{n}_i(t)\| \leq \frac{\lambda_1 - \lambda}{2c}\|\vartheta(t)\|$ for all $t \in [\bar{t}_i, \bar{t}_{i+1})$.

Arguing in an identical manner to the proof of Theorem 2.1, we obtain the following two bounds:

$$|\psi(t)| \leq \left(\tfrac{\lambda_1+\lambda}{2}\right)^{t-\bar{t}_i} |\psi(\bar{t}_i)| + \sum_{j=\bar{t}_i}^{t-1} \left(\tfrac{\lambda_1+\lambda}{2}\right)^{t-j-1} (2\|r(j)\| + \|w(j)\|), \qquad t = \bar{t}_i, \bar{t}_i+1, \dots, \bar{t}_{i+1}; \tag{2.20}$$

this, in turn, implies that there exists $c_2 > c$ so that

$$\left\| \begin{bmatrix} \vartheta(\bar{t}_{i+1}) \\ z(\bar{t}_{i+1}) \end{bmatrix} \right\| \leq c \left(\tfrac{\lambda_1+\lambda}{2}\right)^m \left\| \begin{bmatrix} \vartheta(\bar{t}_i) \\ z(\bar{t}_i) \end{bmatrix} \right\| + \sum_{j=\bar{t}_i}^{\bar{t}_{i+1}-1} c_2 \left(\tfrac{\lambda_1+\lambda}{2}\right)^{\bar{t}_{i+1}-j-1} (\|r(j)\| + \|w(j)\|) + c_2\|r(\bar{t}_{i+1})\|. \tag{2.21}$$

Case 2: $\|\tilde{n}_i(t)\| > \frac{\lambda_1-\lambda}{2c}\|\vartheta(t)\|$ for some $t \in [\bar{t}_i, \bar{t}_{i+1})$.

Arguing in an identical manner to the proof of Theorem 2.1, we obtain the following two bounds: there exists $\gamma_3 > 0$ so that

$$|\psi(t)| \leq \gamma_3^{t-\bar{t}_i} |\psi(\bar{t}_i)| + \sum_{j=\bar{t}_i}^{t-1} \gamma_3^{t-j-1} (\gamma_3\|r(j)\| + \|w(j)\|), \qquad t = \bar{t}_i, \bar{t}_i+1, \dots, \bar{t}_{i+1}, \tag{2.22}$$

$$\begin{aligned} \left\| \begin{bmatrix} \vartheta(\bar{t}_{i+1}) \\ z(\bar{t}_{i+1}) \end{bmatrix} \right\| &\leq c\gamma_3^m \left\| \begin{bmatrix} \vartheta(\bar{t}_i) \\ z(\bar{t}_i) \end{bmatrix} \right\| + \sum_{j=\bar{t}_i}^{\bar{t}_{i+1}-1} c\gamma_3^{\bar{t}_{i+1}-j-1} (\gamma_3\|r(j)\| + \|w(j)\|) + c\|r(\bar{t}_{i+1})\| \\ &\leq c\gamma_3^m \left\| \begin{bmatrix} \vartheta(\bar{t}_i) \\ z(\bar{t}_i) \end{bmatrix} \right\| + \\ &\qquad c_2\gamma_3 \left(\tfrac{2\gamma_3}{\lambda_1+\lambda}\right)^m \sum_{j=\bar{t}_i}^{\bar{t}_{i+1}-1} \left(\tfrac{\lambda_1+\lambda}{2}\right)^{\bar{t}_{i+1}-j-1} (\|r(j)\| + \|w(j)\|) + c_2\|r(\bar{t}_{i+1})\|. \end{aligned} \tag{2.23}$$

This completes Case 2.

At this point we combine Cases 1 and 2. We would like to analyze $\bar{N} \in \mathbb{N}$ intervals of length m; for now we let $\bar{N}$ be free. We see that

$$\sum_{j=\bar{t}}^{\bar{t}+m\bar{N}-1} \|\theta^*(j+1) - \theta^*(j)\| \leq c_0 + \epsilon m\bar{N}.$$

Let N_1 denote the number of intervals of the form $[\bar{t}_i, \bar{t}_{i+1})$ which lie in $[\bar{t}, \bar{t} + m\bar{N}]$ which fall into Case 2; it is easy to see that N_1 satisfies

$$\begin{aligned} N_1 \frac{\lambda_1 - \lambda}{2c} &\leq \left(c_0 + \epsilon m\bar{N}\right)\|f\| \\ \Rightarrow N_1 &\leq \left(\frac{2c\|f\|}{\lambda_1 - \lambda}\right)c_0 + \left(\frac{2c\|f\|}{\lambda_1 - \lambda}\right)\epsilon m\bar{N}; \end{aligned} \tag{2.24}$$

observe that N_1 depends on both c_0 and ϵ. Using (2.21) and (2.23) we obtain

$$\begin{aligned} \left\|\begin{bmatrix}\vartheta(\bar{t}+m\bar{N})\\ z(\bar{t}+m\bar{N})\end{bmatrix}\right\| &\leq c^{\bar{N}}\left(\tfrac{\lambda_1+\lambda}{2}\right)^{m(\bar{N}-N_1)}\gamma_3^{mN_1}\left\|\begin{bmatrix}\vartheta(\bar{t})\\ z(\bar{t})\end{bmatrix}\right\| + \\ & 2\bar{N}\left(\tfrac{2\gamma_3}{\lambda_1+\lambda}\right)^{\bar{N}}(c_2\gamma_3^{m+1})^{\bar{N}}\left(\tfrac{2}{\lambda_1+\lambda}\right)^{(m+1)\bar{N}} \times \sum_{j=\bar{t}}^{\bar{t}+m\bar{N}-1}\left(\tfrac{\lambda_1+\lambda}{2}\right)^{\bar{t}+m\bar{N}-j-1}(\|r(j)\|+\|w(j)\|) \\ & + c_2\|r(\bar{t}_{(q+1)m})\|. \end{aligned} \tag{2.25}$$

At this point, we will choose quantities m, ϵ and $\bar{N}$, in that order, so that the key gain

$$c^{\bar{N}}\left(\tfrac{\lambda_1+\lambda}{2}\right)^{m(\bar{N}-N_1)}\gamma_3^{mN_1} < 1.$$

First of all, we apply the bound on N_1 given in (2.24) to this key gain:

$$\begin{aligned} c^{\bar{N}}\left(\tfrac{\lambda_1+\lambda}{2}\right)^{m(\bar{N}-N_1)}\gamma_3^{mN_1} &= \left[c\left(\tfrac{\lambda_1+\lambda}{2}\right)^{m}\right]^{\bar{N}}\left[\left(\tfrac{2\gamma_3}{\lambda_1+\lambda}\right)^{N_1}\right]^{m} \\ &\leq \left[c\left(\tfrac{\lambda_1+\lambda}{2}\right)^{m}\right]^{\bar{N}}\left[\left(\tfrac{2\gamma_3}{\lambda_1+\lambda}\right)^{\left[\left(\frac{2c\|f\|}{\lambda_1-\lambda}\right)c_0+\left(\frac{2c\|f\|}{\lambda_1-\lambda}\right)\epsilon m\bar{N}\right]}\right]^{m}. \end{aligned} \tag{2.26}$$

Now choose m so that $c\left(\frac{\lambda_1+\lambda}{2}\right)^m =: \lambda_2 < 1$, i.e. any $m > \frac{\ln(c)}{\ln(2)-\ln(\lambda_1+\lambda)}$. So rewriting (2.26), we now obtain

$$c^{\bar{N}}\left(\tfrac{\lambda_1+\lambda}{2}\right)^{m(\bar{N}-N_1)}\gamma_3^{mN_1} \leq \left[\left(\tfrac{2\gamma_3}{\lambda_1+\lambda}\right)^{\left(\frac{2c\|f\|}{\lambda_1-\lambda}\right)c_0 m}\right]\left(\left[\left(\tfrac{2\gamma_3}{\lambda_1+\lambda}\right)^{\left(\frac{2c\|f\|}{\lambda_1-\lambda}\right)\epsilon m^2}\right]\times\lambda_2\right)^{\bar{N}}.$$

Now observe that

$$\lim_{\epsilon\to 0}\left[\left(\tfrac{2\gamma_3}{\lambda_1+\lambda}\right)^{\left(\frac{2c\|f\|}{\lambda_1-\lambda}\right)\epsilon m^2}\right] = 1,$$

so now choose $\epsilon > 0$ so that

$$\underbrace{\left[\left(\tfrac{2\gamma_3}{\lambda_1+\lambda} \right)^{\left(\frac{2c\|f\|}{\lambda_1-\lambda} \right) \epsilon m^2} \right] \times \lambda_2}_{=:\lambda_3} < 1;$$

notice that ϵ is independent of c_0. With this choice we now have

$$c^{\bar{N}} \left(\tfrac{\lambda_1+\lambda}{2} \right)^{m(\bar{N}-N_1)} \gamma_3^{mN_1} \leq \left[\left(\tfrac{2\gamma_3}{\lambda_1+\lambda} \right)^{\left(\frac{2c\|f\|}{\lambda_1-\lambda} \right) c_0 m} \right] \times \lambda_3^{\bar{N}}.$$

Last of all, now choose $\bar{N}$ so that

$$\underbrace{\left[\left(\tfrac{2\gamma_3}{\lambda_1+\lambda} \right)^{\left(\frac{2c\|f\|}{\lambda_1-\lambda} \right) c_0 m} \right] \times \lambda_3^{\bar{N}}}_{=:\lambda_4} < 1;$$

any $\bar{N} > \frac{2cc_0 m\|f\|[\ln(2\gamma_3)-\ln(\lambda_1+\lambda)]}{(\lambda-\lambda_1)\ln(\lambda_3)}$ will do. Observe that $\bar{N}$ depends on c_0.

So incorporating all of the above, there exists $\gamma_4 > 0$ (which clearly depends on c_0 via $\bar{N}$) so that we can rewrite (2.25) as

$$\left\| \begin{bmatrix} \vartheta(\bar{t}+m\bar{N}) \\ z(\bar{t}+m\bar{N}) \end{bmatrix} \right\| \leq \lambda_4 \left\| \begin{bmatrix} \vartheta(\bar{t}) \\ z(\bar{t}) \end{bmatrix} \right\| + \gamma_4 \sum_{j=\bar{t}}^{\bar{t}+m\bar{N}-1} \left(\tfrac{\lambda_1+\lambda}{2} \right)^{\bar{t}+m\bar{N}-j-1} (\|r(j)\| + \|w(j)\|) + \gamma_4 \|r(\bar{t}+m\bar{N})\|. \tag{2.27}$$

Now let $\tau \geq t_0$ be arbitrary. By setting $\bar{t} = \tau, \tau+m\bar{N}, \tau+2m\bar{N}, \ldots$, in succession, with $\lambda_5 := \max\left\{ \lambda_4^{\frac{1}{m\bar{N}}}, \frac{\lambda_1+\lambda}{2} \right\}$ (which clearly depends on c_0 via $\bar{N}$) it follows from (2.27) that

$$\left\| \begin{bmatrix} \vartheta(\tau+q\bar{N}m) \\ z(\tau+q\bar{N}m) \end{bmatrix} \right\| \leq \lambda_5^{q\bar{N}m} \left\| \begin{bmatrix} \vartheta(\tau) \\ z(\tau) \end{bmatrix} \right\| + \gamma_4 \sum_{j=\tau}^{\tau+q\bar{N}m-1} \lambda_5^{\tau+q\bar{N}m-j-1} (\|r(j)\| + \|w(j)\|) + \gamma_4 \|r(\tau+q\bar{N}m)\|, \qquad q \in \mathbb{Z}^+. \tag{2.28}$$

So $\begin{bmatrix} \vartheta(t) \\ z(t) \end{bmatrix}$ is well-behaved at $t = \tau, \tau+m\bar{N}, \tau+2m\bar{N}, \ldots$, etc; we can use (2.20) of Case 1, (2.22) of Case 2 and (2.19) to prove that nothing untoward happens between these times. We conclude that there exists a constant γ_5 so that

$$\left\| \begin{bmatrix} \vartheta(t) \\ z(t) \end{bmatrix} \right\| \leq \gamma_5 \lambda_5^{t-\tau} \left\| \begin{bmatrix} \vartheta(\tau) \\ z(\tau) \end{bmatrix} \right\| + \gamma_5 \sum_{j=\tau}^{t-1} \lambda_5^{t-j-1} \left(\|r(j)\| + \|w(j)\|\right) + \gamma_5 \|r(t)\|, \quad t \geq \tau. \tag{2.29}$$

Since $\tau \geq t_0$ is arbitrary, the desired bound is proven. ■

2.4 Tolerance to Unmodelled Dynamics

We now consider the time-varying plant (2.4) with the term $d_\Delta(t) \in \mathbb{R}^r$ added to represent unmodelled dynamics:

$$y(t+1) = \theta^*(t)^\top f\Big(\vartheta(t)\Big) + w(t) + d_\Delta(t), \quad \vartheta(t_0) = \vartheta_0. \tag{2.30}$$

Here we consider (a generalized version of) a class of unmodelled dynamics which is common in the adaptive control literature—see [34] and [46]. 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

$$m(t+1) = \beta m(t) + \beta |g(\vartheta(t))|, \qquad m(t_0) = m_0 \tag{2.31a}$$
$$\|d_\Delta(t)\| \leq \mu m(t) + \mu |g(\vartheta(t))|, \qquad t \geq t_0. \tag{2.31b}$$

It turns out that this model subsumes classical additive uncertainty, multiplicative uncertainty, and uncertainty in a coprime factorization, with side constraints on the pole locations (less than β in magnitude) as well as strict causality; see [46] for a more detailed explanation. We will now show that if the controller (2.2) provides exponential stability and a convolution bound for $\Big(f, \mathcal{S}(\mathcal{S}, c_0, \epsilon)\Big)$, then a degree of tolerance to unmodelled dynamics can be proven.

Theorem 2.3. *Suppose that the controller (2.2) provides exponential stability and a convolution bound for $\big(f, \mathbb{S}(\mathcal{S}, c_0, \epsilon)\big)$ with a gain c_1 and decay rate $\lambda_1 \in (0,1)$. Then for every $\beta \in (0,1)$ and $\lambda_2 \in (\max\{\lambda_1, \beta\}, 1)$, there exist $\bar{\mu} > 0$ and $c_2 > 0$ so that for every $\theta^* \in \mathbb{S}(\mathcal{S}, c_0, \epsilon)$, $\mu \in (0, \bar{\mu})$, $t_0 \in \mathbb{Z}$, $\vartheta_0 \in \mathbb{R}^{n_y \cdot r + n_u \cdot m}$, $z_0 \in \mathbb{R}^{l_1}$, $\theta_0 \in \Omega \subset \mathbb{R}^{l_2}$, and $w, r \in \mathbb{S}(\mathbb{R}^r)$, when the controller (2.2) is applied to the plant (2.30) with d_Δ satisfying (2.31), the following holds:*

$$\left\| \begin{bmatrix} \vartheta(t) \\ z(t) \\ m(t) \end{bmatrix} \right\| \leq c_2 \lambda_2^{t-t_0} \left\| \begin{bmatrix} \vartheta_0 \\ z_0 \\ m_0 \end{bmatrix} \right\| + \sum_{j=t_0}^{t-1} c_2 \lambda_2^{t-j-1} (\|r(j)\| + \|w(j)\|) + c_2 \|r(t)\|, \quad t \geq t_0. \tag{2.32}$$

Remark 2.5. *The following proof is based, in part, on the proof of Theorem 3 of [46], which deals with a much simpler setup.*

Proof of Theorem 2.3. Fix $\beta \in (0,1)$ and $\lambda_2 \in (\max\{\lambda_1, \beta\}, 1)$ and let $\theta^* \in \mathbb{S}(\mathcal{S}, c_0, \epsilon)$, $t_0 \in \mathbb{Z}$, $\vartheta_0 \in \mathbb{R}^{n_y \cdot r + n_u \cdot m}$, $z_0 \in \mathbb{R}^{l_1}$, $\theta_0 \in \Omega \subset \mathbb{R}^{l_2}$, $w \in \mathbb{S}(\mathbb{R}^r)$ and $r \in \mathbb{S}(\mathbb{R}^r)$ be arbitrary. So by hypothesis:

$$\left\| \begin{bmatrix} \vartheta(t) \\ z(t) \end{bmatrix} \right\| \leq c_1 \lambda_1^{t-\tau} \left\| \begin{bmatrix} \vartheta(\tau) \\ z(\tau) \end{bmatrix} \right\| + \sum_{j=\tau}^{t-1} c_1 \lambda_1^{t-j-1} (\|r(j)\| + \|w(j)\| + \|d_\Delta(j)\|) + c_1 \|r(t)\|,$$
$$t \geq \tau \geq t_0. \tag{2.33}$$

To convert this inequality to an equality, we consider the associated difference equations

$$\tilde{\vartheta}(t+1) = \lambda_1 \tilde{\vartheta}(t) + c_1 \|r(t)\| + c_1 \|w(t)\| + c_1 \mu \tilde{m}(t) + c_1 \mu \|g\| \tilde{\vartheta}(t), \qquad \tilde{\vartheta}(t_0) = c_1 \left\| \begin{bmatrix} \vartheta_0 \\ z_0 \end{bmatrix} \right\|,$$

together with the difference equation based on (2.31a):

$$\tilde{m}(t+1) = \beta \tilde{m}(t) + \beta \|g\| \tilde{\vartheta}(t), \quad \tilde{m}(t_0) = |m_0|.$$

Using induction together with (2.33), (2.31a), and (2.31b), we can prove that

$$\left\| \begin{bmatrix} \vartheta(t) \\ z(t) \end{bmatrix} \right\| \leq \tilde{\vartheta}(t) + c_1 \|r(t)\|, \tag{2.34a}$$
$$|m(t)| \leq \tilde{m}(t), \qquad t \geq t_0. \tag{2.34b}$$

If we combine the difference equations for $\tilde{\vartheta}(t)$ and $\tilde{m}(t)$, we obtain

$$\begin{bmatrix} \tilde{\vartheta}(t+1) \\ \tilde{m}(t+1) \end{bmatrix} = \underbrace{\begin{bmatrix} \lambda_1 + c_1\|g\|\mu & c_1\mu \\ \beta\|g\| & \beta \end{bmatrix}}_{=:A_{\mathrm{cl}}(\mu)} \begin{bmatrix} \tilde{\vartheta}(t) \\ \tilde{m}(t) \end{bmatrix} + \begin{bmatrix} c_1 \\ 0 \end{bmatrix} \left(\|r(t)\| + \|w(t)\|\right), \qquad t \geq t_0. \tag{2.35}$$

Now we see that

$$A_{\mathrm{cl}}(\mu) \to \begin{bmatrix} \lambda_1 & 0 \\ \beta\|g\| & \beta \end{bmatrix}$$

as $\mu \to 0$, and this matrix has eigenvalues of $\{\lambda_1, \beta\}$ which are both less that $\lambda_2 < 1$. Using a standard Lyapunov argument, it is easy to prove that there exist $\bar{\mu} > 0$ and $\gamma_1 > 0$ such that for all $\mu \in (0, \bar{\mu}]$, we have

$$\left\| A_{\mathrm{cl}}(\mu)^k \right\| \leq \gamma_1 \lambda_2^k, \qquad k \geq 0;$$

if we use this in (2.35) and then apply the bound in (2.34), it follows that there exists a constant γ_2 so that

$$\left\| \begin{bmatrix} \vartheta(t) \\ z(t) \\ m(t) \end{bmatrix} \right\| \leq \gamma_2 \lambda_2^{t-t_0} \left\| \begin{bmatrix} \vartheta_0 \\ z_0 \\ m_0 \end{bmatrix} \right\| + \sum_{j=t_0}^{t-1} c_1 \gamma_1 \lambda_2^{t-j-1} \left(\|r(j)\| + \|w(j)\|\right) + c_1\|r(t)\|, \quad t \geq t_0 \tag{2.36}$$

as desired. ■

2.5 Conclusion

In this chapter we have shown that for a class of nonlinear plant and controller combinations, if linear-like properties of exponential stability and a convolution bound on the closed-loop behavior can be proven, then tolerance to small parameter time-variations and a small amount of unmodelled dynamics follows immediately. This result is applicable to various adaptive control paradigms in a modular fashion; it allows one to focus on the nominal plant in the analysis knowing that robustness will come for free. Indeed, this chapter will be utilized as a powerful technical tool in subsequent results of this thesis to prove robustness of the proposed adaptive controllers. Last of all, this result has the potential for use in other non-adaptive contexts.

Chapter 3

The Projection Algorithm

In this brief chapter we present the parameter estimation algorithm that will be used as part of the adaptive controller—we consider the original **Projection Algorithm**.

3.1 The Ideal Projection Algorithm

Although we will eventually use the estimation algorithm in the context of dynamic systems, we consider the following more general model:

$$y(t+1) = \phi(t)^\top \theta^* + w(t), \qquad t \in \mathbb{Z}, \tag{3.1}$$

with $y(t) \in \mathbb{R}$ as the output measurement, $\phi(t) \in \mathbb{R}^p$ as the regressor (data) vector, $\theta^* \in \mathbb{R}^p$ as the *unknown* parameter vector, and $w(t) \in \mathbb{R}$ as the noise term.

Given $y(t+1)$, $\phi(t)$, and an estimate of θ^* denoted $\hat{\theta}(t)$, at time t, define the prediction error associated with (3.1) by

$$e(t+1) := y(t+1) - \phi(t)^\top \hat{\theta}(t); \tag{3.2}$$

this is a measure of the error in the estimate. A common way to obtain the next parameter

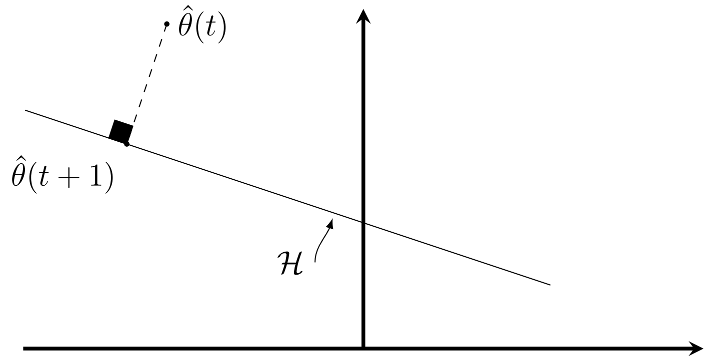


Figure 3.1: Geometric picture of the Projection Algorithm.

estimate is to solve the following optimization problem:

$$\arg\min_{\theta} \tfrac{1}{2}\|\theta - \hat{\theta}(t)\|^2$$

$$\text{subject to} \quad y(t+1) = \phi(t)^\top \theta.$$

Solving this optimization problem by the Lagrange multipliers method yields the **original ideal Projection Algorithm**:

$$\hat{\theta}(t+1) = \begin{cases} \hat{\theta}(t) & \text{if } \phi(t) = 0 \\ \hat{\theta}(t) + \dfrac{\phi(t)}{\|\phi(t)\|^2} e(t+1) & \text{otherwise.} \end{cases} \tag{3.3}$$

The geometric interpretation of the algorithm in (3.3) is as follows: given $y(t+1)$ and $\phi(t)$, the possible values for θ^* in the model (3.1) lie in the hyperplane

$$\mathcal{H} = \left\{ \theta \in \mathbb{R}^p \,:\, y(t+1) = \phi(t)^\top \theta \right\};$$

then, we choose the next update to be the one closest to $\hat{\theta}(t)$, i.e. we choose $\hat{\theta}(t+1)$ to be the one that minimizes $\frac{1}{2}\|\hat{\theta}(t+1) - \hat{\theta}(t)\|^2$ and lies in $\mathcal{H}$. The estimate update $\hat{\theta}(t+1)$ is merely the projection of $\hat{\theta}(t)$ onto the hyperplane $\mathcal{H}$, hence the name "Projection Algorithm". You can view this interpretation visually for the case when $\theta^* \in \mathbb{R}^2$ in Figure 3.1, where $\mathcal{H}$ is a line.

If $\phi(t)$ is close to zero, numerical issues may arise, so it is the norm in the literature, e.g. [21] and [20], to replace (3.3) with the following classical algorithm (with $0 < \alpha < 2$ and $\beta > 0$):

$$\hat{\theta}(t+1) = \hat{\theta}(t) + \frac{\alpha\phi(t)}{\beta + \|\phi(t)\|^2} e(t+1). \tag{3.4}$$

This is widely used and plays a role in many discrete-time adaptive control approaches; however, a careful look at (3.4) shows that gain of the update law is small if $\phi(t)$ is small, which is the reason why the closed-loop behavior in the adaptive control context is asymptotic rather than exponential (in general) when an estimator of this form is used—see [43], [44] and [46].

3.2 Projection Onto a Convex Set

For a compact and convex set $\mathcal{S} \subset \mathbb{R}^p$, let the function

$$\operatorname*{Proj}_{\mathcal{S}} \{\cdot\} \, : \, \mathbb{R}^p \mapsto \mathcal{S}$$

denote the projection onto the set $\mathcal{S}$; because the set $\mathcal{S}$ is closed and convex, we know that the function $\operatorname*{Proj}_{\mathcal{S}}$ is well-defined; some examples will be discussed shortly.

If we know that θ^* belongs to a set $\mathcal{S} \subset \mathbb{R}^p$, then we can also constrain estimate updates to that set. We use the algorithm in (3.3) accompanied by projection onto $\mathcal{S}$ to ensure that the estimate remains in $\mathcal{S}$ for all time. To this end, with $\hat{\theta}(t_0) \in \mathcal{S}$, for $t \geq t_0$ we set

$$\check{\theta}(t+1) = \begin{cases} \hat{\theta}(t) & \text{if } \phi(t) = 0 \\ \hat{\theta}(t) + \dfrac{\phi(t)}{\|\phi(t)\|^2} e(t+1) & \text{otherwise,} \end{cases} \tag{3.5a}$$

$$\hat{\theta}(t+1) = \operatorname*{Proj}_{\mathcal{S}} \left\{ \check{\theta}(t+1) \right\}. \tag{3.5b}$$

See Figure 3.2 to visualize (3.5b) for an arbitrary closed and convex set $\mathcal{S}$; the function $\operatorname*{Proj}_{\mathcal{S}}$ has the nice property that, for every $\check{\theta} \in \mathbb{R}^p$ and every $\theta^* \in \mathcal{S}$, we have

$$\left\| \operatorname*{Proj}_{\mathcal{S}}\{\check{\theta}\} - \theta^* \right\| \leq \left\| \check{\theta} - \theta^* \right\|,$$

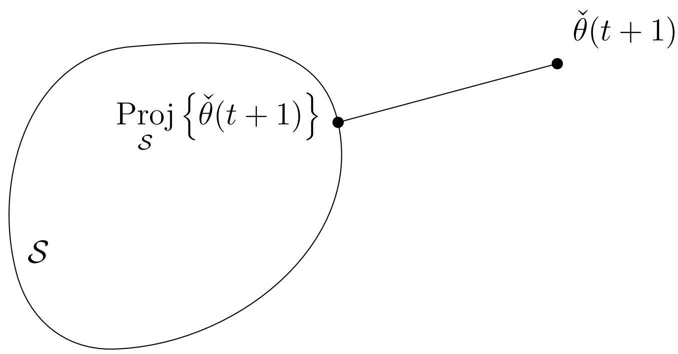


Figure 3.2: Projection onto a convex set.

i.e. projecting $\check{\theta}$ onto $\mathcal{S}$ never makes it further away from θ^* (in the Euclidean-norm sense).

Remark 3.1. *Here, we present a couple of examples for the projection operation of* $\underset{\mathcal{S}}{\text{Proj}}\{\check{\theta}\}$ *applied to a vector* $\check{\theta} \in \mathbb{R}^p$*.*

- *Let us denote the elements of a vector* $\check{\theta} \in \mathbb{R}^p$ *as follows:* $\check{\theta} = \begin{bmatrix}\check{\theta}_1 & \check{\theta}_2 & \cdots & \check{\theta}_p\end{bmatrix}^\top$*. Let us consider the case of a hyper-rectangle: there exists constants* $\underline{\theta}_j \leq \bar{\theta}_j$*,* $j = 1, 2, \ldots, p$*, such that*
$$\mathcal{S} = \left\{\check{\theta} \in \mathbb{R}^p \,:\, \underline{\theta}_j \leq \check{\theta}_j \leq \bar{\theta}_j,\ j = 1, 2, \ldots, p\right\}.$$
We can apply the function $\underset{\mathcal{S}}{\text{Proj}}$ *in an element-by-element basis: with*
$$\hat{\theta}_j := \begin{cases} \bar{\theta}_j & \text{if } \check{\theta}_j > \bar{\theta}_j \\ \check{\theta}_j & \text{if } \underline{\theta}_j \leq \check{\theta}_j \leq \bar{\theta}_j, \\ \underline{\theta}_j & \text{if } \check{\theta}_j < \underline{\theta}_j \end{cases}$$
for every $j = 1, 2, \ldots, p$*, it turns out that*
$$\underset{\mathcal{S}}{\text{Proj}}\left\{\check{\theta}\right\} = \begin{bmatrix}\hat{\theta}_1 \\ \hat{\theta}_2 \\ \vdots \\ \hat{\theta}_p\end{bmatrix}.$$

- *Now consider the case when the set $\mathcal{S} \subset \mathbb{R}^p$ is a hypersphere defined by a center $\theta_c \in \mathcal{S}$ and a radius R; then the projection of $\check{\theta}$ onto $\mathcal{S}$ can be defined as [37]:*

$$\operatorname*{Proj}_{\mathcal{S}}\left\{\check{\theta}\right\} = \begin{cases} \check{\theta} & \text{if } \|\check{\theta} - \theta_c\| \le R \\ \theta_c + R\dfrac{\check{\theta} - \theta_c}{\|\check{\theta} - \theta_c\|} & \text{if } \|\check{\theta} - \theta_c\| > R. \end{cases}$$

- *For a more general form of the projection operator, one can refer to [64].*

3.3 A Revised Estimation Algorithm

One may be concerned that the original problem of dividing by a number close to zero remains. Also we have discussed that the classical algorithm in (3.4) can lead to loss of exponential stability and a bounded noise gain. So as proposed in [44] and [46], a middle ground is presented: we turn-off the estimation if it is clear that the noise is swamping the estimation error[1]. In particular, when $\theta^* \in \mathcal{S}$ and $\theta(t) \in \mathcal{S}$, by examining (3.2), we see that

$$\begin{aligned} e(t+1) &= y(t+1) - \phi(t)^\top \hat{\theta}(t) \\ &= \phi(t)^\top \left[\theta^* - \hat{\theta}(t)\right] + w(t), \end{aligned}$$

which means that

$$|e(t+1)| \le 2\|\mathcal{S}\|\|\phi(t)\| + |w(t)|.$$

Therefore, if

$$|e(t+1)| > 2\|\mathcal{S}\|\|\phi(t)\|,$$

then the update to $\hat{\theta}(t)$, namely

$$\|\check{\theta}(t+1) - \hat{\theta}(t)\| = \frac{|e(t+1)|}{\|\phi(t)\|},$$

[1]This is different than the norm in literature when the estimator is turned-off if the prediction error is small, e.g. estimators with deadzone—e.g. [21], [34] and [41].

will be greater than $2\|\mathcal{S}\|$; this means that the noise may be overwhelming the estimation update. Motivated by this, with $\delta \in (0, \infty]$, let us replace (3.5a) with

$$\check{\theta}(t+1) = \begin{cases} \hat{\theta}(t) + \dfrac{\phi(t)}{\|\phi(t)\|^2} e(t+1) & \text{if } |e(t+1)| < (2\|\mathcal{S}\| + \delta)\, \|\phi(t)\| \\ \hat{\theta}(t) & \text{otherwise;} \end{cases} \tag{3.6}$$

in the case of $\delta = \infty$, we will adopt the understanding that $\infty \times 0 = 0$, in which case the above formula collapses into the original ideal one in (3.3). In the case of $\delta < \infty$, we can be assured the update term is bounded above by $2\|\mathcal{S}\| + \delta$, which should alleviate concerns about having an infinite gain. To have a more compact notation, define $\rho_\delta : \mathbb{R}^p \times \mathbb{R} \mapsto \{0, 1\}$ by

$$\rho_\delta\left(\phi(t), e(t+1)\right) := \begin{cases} 1 & \text{if } |e(t+1)| < (2\|\mathcal{S}\| + \delta)\, \|\phi(t)\| \\ 0 & \text{otherwise;} \end{cases}$$

this leads to the "vigilant" estimator:

$$\check{\theta}(t+1) = \hat{\theta}(t) + \rho_\delta\left(\phi(t), e(t+1)\right) \frac{\phi(t)}{\|\phi(t)\|^2} e(t+1) \tag{3.7a}$$

$$\hat{\theta}(t+1) = \operatorname*{Proj}_{\mathcal{S}} \left\{ \check{\theta}(t+1) \right\}. \tag{3.7b}$$

3.4 The Main Estimation Algorithm

In this section, we present a more general version of the estimator (3.7) to be used throughout this thesis. In several cases in the following chapters, some modifications are incorporated into the original-projection-algorithm based estimator of (3.7). To this end, with $\phi_m(t) \in \mathbb{R}^q$ where $q \geq p$ and

$$\|\phi(t)\| \leq \|\phi_m(t)\|, \qquad t \in \mathbb{Z}, \tag{3.8}$$

we replace the algorithm in (3.7) as follows: with $\rho_\delta : \mathbb{R}^q \times \mathbb{R} \mapsto \{0, 1\}$ defined by

$$\rho_\delta\left(\phi_m(t), e(t+1)\right) := \begin{cases} 1 & \text{if } |e(t+1)| < (2\|\mathcal{S}\| + \delta)\, \|\phi_m(t)\| \\ 0 & \text{otherwise,} \end{cases}$$

we set

$$\check{\theta}(t+1) = \hat{\theta}(t) + \rho_\delta\left(\phi_m(t), e(t+1)\right)\frac{\phi(t)}{\|\phi_m(t)\|^2}e(t+1) \tag{3.9a}$$

$$\hat{\theta}(t+1) = \operatorname*{Proj}_{\mathcal{S}}\left\{\check{\theta}(t+1)\right\}. \tag{3.9b}$$

Observe that the quantity ϕ_m has replaced ϕ in the denominator of the estimator update and in the associated definition of ρ_δ. In the following chapters, definitions for ϕ_m will be introduced as needed.

In the following we list properties of the estimation algorithm (3.9). These properties are similar to ones found in Propositions 1 and 3 of [46]; the difference arises when $\phi_m(t) \neq \phi(t)$. Define the parameter estimation error

$$\tilde{\theta}(t) := \hat{\theta}(t) - \theta^*.$$

First, we present a property of $\tilde{\theta}(t)$, which will be essential to the analysis of the closed-loop behavior; this Proposition is dealing with the case when the estimator algorithm in (3.9) is applied to the model in (3.1) when θ^* lies in a compact and convex set $\mathcal{S} \subset \mathbb{R}^p$. It is a generalization of the well-known result for the classical estimator (3.4).

Proposition 3.1. *For every $t_0 \in \mathbb{Z}$, $t_2 > t_1 \geq t_0$, $\phi \in \mathbb{S}(\mathbb{R}^p)$, $\phi_m \in \mathbb{S}(\mathbb{R}^q)$ satisfying (3.8), $\hat{\theta}(t_0) \in \mathcal{S}$, $\theta^* \in \mathcal{S}$ and $w \in \boldsymbol{\ell}_\infty$, when the estimation algorithm in (3.9) is applied to the model (3.1), the following holds:*

$$\|\tilde{\theta}(t_2)\|^2 \leq \|\tilde{\theta}(t_1)\|^2 + \sum_{j=t_1}^{t_2-1} \rho_\delta(\phi_m(j), e(j+1))\left[-\frac{1}{2}\frac{e(j+1)^2}{\|\phi_m(j)\|^2} + 2\frac{w(j)^2}{\|\phi_m(j)\|^2}\right]. \tag{3.10}$$

Proof of Proposition 3.1. Let $t_0 \in \mathbb{Z}$, $t_2 > t_1 \geq t_0$, $\phi \in \mathbb{S}(\mathbb{R}^p)$, $\phi_m \in \mathbb{S}(\mathbb{R}^q)$ satisfying (3.8), $\hat{\theta}(t_0) \in \mathcal{S}$, $\theta^* \in \mathcal{S}$ and $w \in \boldsymbol{\ell}_\infty$ be arbitrary.

Next, define $\check{\tilde{\theta}}(t) := \check{\theta}(t) - \theta^*$. When $\rho_\delta(\phi_m(t), e(t+1)) = 0$ we have $\hat{\theta}(t+1) = \hat{\theta}(t)$ by (3.9), which implies that

$$\|\tilde{\theta}(t+1)\|^2 = \|\tilde{\theta}(t)\|^2. \tag{3.11}$$

On the other hand, when $\rho_\delta(\phi_m(t), e(t+1)) = 1$, by (3.9) and (3.8) we have

$$\begin{aligned}
\check{\tilde{\theta}}(t+1) &= \tilde{\theta}(t) + \frac{\phi(t)}{\|\phi_m(t)\|^2} e(t+1) \\
\Rightarrow \|\check{\tilde{\theta}}(t+1)\|^2 &= \|\tilde{\theta}(t)\|^2 + \frac{\|\phi(t)\|^2 |e(t+1)|^2}{\|\phi_m(t)\|^4} + 2\frac{\tilde{\theta}(t)^\top \phi(t) e(t+1)}{\|\phi_m(t)\|^2} \\
&\leq \|\tilde{\theta}(t)\|^2 + \frac{|e(t+1)|^2}{\|\phi_m(t)\|^2} + 2\frac{\tilde{\theta}(t)^\top \phi(t) e(t+1)}{\|\phi_m(t)\|^2}.
\end{aligned} \tag{3.12}$$

From (3.2) and (3.1), we have

$$\begin{aligned}
e(t+1) &= y(t+1) - \hat{\theta}(t)^\top \phi(t) \\
&= \theta^{*\top} \phi(t) - \hat{\theta}(t)^\top \phi(t) + w(t) \\
&= -\tilde{\theta}(t)^\top \phi(t) + w(t).
\end{aligned}$$

If we use the above to find a representation for $\tilde{\theta}(t)^\top \phi(t)$ in (3.12) we obtain

$$\begin{aligned}
\|\check{\tilde{\theta}}(t+1)\|^2 &\leq \|\tilde{\theta}(t)\|^2 + \frac{|e(t+1)|^2}{\|\phi_m(t)\|^2} + 2\frac{[w(t) - e(t+1)]e(t+1)}{\|\phi_m(t)\|^2} \\
&= \|\tilde{\theta}(t)\|^2 - \frac{|e(t+1)|^2}{\|\phi_m(t)\|^2} + \frac{2w(t)e(t+1)}{\|\phi_m(t)\|^2} \\
&\leq \|\tilde{\theta}(t)\|^2 - \frac{|e(t+1)|^2}{2\|\phi_m(t)\|^2} + \frac{2w(t)^2}{\|\phi_m(t)\|^2}.
\end{aligned}$$

(The last step uses the fact that for $a, b \geq 0$, we have $-a^2 + 2ab \leq -\frac{1}{2}a^2 + 2b^2$). Since projection does not make the parameter estimate worse (i.e. $\|\tilde{\theta}(t+1)\| \leq \|\check{\tilde{\theta}}(t+1)\|$), it follows that

$$\|\tilde{\theta}(t+1)\|^2 \leq \|\tilde{\theta}(t)\|^2 - \frac{|e(t+1)|^2}{2\|\phi_m(t)\|^2} + \frac{2w(t)^2}{\|\phi_m(t)\|^2}. \tag{3.13}$$

If we combine the above bound for the case of $\rho_\delta(\cdot,\cdot) = 1$ with (3.11) for the case of $\rho_\delta(\cdot,\cdot) = 0$, and iterate, then we obtain (3.10). ■

Next, we provide the following result that provides a bound on the difference between parameter estimates at two different points in time; this property holds when the algorithm (3.9) is applied to (3.1) whether θ^* lies in $\mathcal{S}$ or not.

Proposition 3.2. *For every $t_0 \in \mathbb{Z}$, $t_2 > t_1 \geq t_0$, $\phi \in \mathbb{S}(\mathbb{R}^p)$, $\phi_m \in \mathbb{S}(\mathbb{R}^q)$ satisfying (3.8), $\hat{\theta}(t_0) \in \mathcal{S}$, $\theta^* \in \mathbb{R}^p$ and $w \in \boldsymbol{\ell}_\infty$, when the estimation algorithm in (3.9) is applied to the model (3.1), the following holds:*

$$\|\hat{\theta}(t_2) - \hat{\theta}(t_1)\| \leq \sum_{j=t_1}^{t_2-1} \rho_\delta(\phi_m(j), e(j+1)) \frac{|e(j+1)|}{\|\phi_m(j)\|}. \tag{3.14}$$

Proof of Proposition 3.2. Let $t_0 \in \mathbb{Z}$, $t_2 > t_1 \geq t_0$, $\phi \in \mathbb{S}(\mathbb{R}^p)$, $\phi_m \in \mathbb{S}(\mathbb{R}^q)$ satisfying (3.8), $\hat{\theta}(t_0) \in \mathcal{S}$, $\theta^* \in \mathbb{R}^p$ and $w \in \boldsymbol{\ell}_\infty$ be arbitrary.

For the estimator (3.9), projection does not make the parameter estimate worse; for $t \geq t_0$, it follows from (3.9) that if $\rho_\delta(\phi_m(t), e(t+1)) = 0$, then $\hat{\theta}(t+1) = \hat{\theta}(t)$, so

$$\|\hat{\theta}(t+1) - \hat{\theta}(t)\| = 0,$$

and if $\rho_\delta(\phi_m(t), e(t+1))(t) = 1$, then

$$\begin{aligned}
\|\hat{\theta}(t+1) - \hat{\theta}(t)\| &\leq \|\check{\theta}(t+1) - \hat{\theta}(t)\| \\
&\leq \left\| \frac{\phi(t)e(t+1)}{\|\phi_m(t)\|^2} \right\| \\
&\leq \frac{\|\phi(t)\||e(t+1)|}{\|\phi_m(t)\|^2} \\
&\leq \frac{|e(t+1)|}{\|\phi_m(t)\|}.
\end{aligned}$$

We conclude that (3.14) follows by iteration. ■

Before proceeding with the rest of the thesis, we would like to point out that, for ease of notation, we introduce a compact notation of the function ρ_δ: for $t \in \mathbb{Z}$ we define

$$\rho(t) := \rho_\delta\left(\phi_m(t), e(t+1)\right).$$

In each chapter, specific details of the estimation algorithm will be presented accordingly.

Chapter 4

One-Step-Ahead Adaptive Control of First-order Plants

4.1 Introduction

In the recent paper by Miller [43], an approach is provided which guarantees a linear-like convolution bound on the closed-loop behavior in the context of the first-order one-step-ahead adaptive control paradigm. The requirement of convexity on the set of uncertainty plays a crucial role in obtaining these nice closed-loop properties. Since convexity is a very restrictive requirement, the main objective of this chapter is to extend the approach by replacing the assumption of a convex and compact uncertainty set with the assumption of a compact uncertainty set. This chapter will present itself as a showcase to illustrate the approach proposed in the whole thesis[1].

4.2 The Setup

Here we consider the first order system

$$y(t+1) = ay(t) + bu(t) + w(t), \qquad y(t_0) = y_0, \quad t \geq t_0, \tag{4.1}$$

[1] A version of this chapter was published as a conference paper in [68].

where $y(t) \in \mathbb{R}$ is the output, $u(t) \in \mathbb{R}$ is the control input, $w(t) \in \mathbb{R}$ is the noise (or disturbance); define

$$\phi(t) := \begin{bmatrix} y(t) \\ u(t) \end{bmatrix} \quad \text{and} \quad \theta^* := \begin{bmatrix} a \\ b \end{bmatrix}.$$

We assume that θ^* is unknown and belongs to a closed and bounded (compact) set $\mathcal{S} \subset \mathbb{R}^2$ satisfying a controllability assumption: $\begin{bmatrix} a \\ 0 \end{bmatrix} \notin \mathcal{S}$ for every $a \in \mathbb{R}$. Here we have an exogenous reference signal and the control objective is to track it asymptotically while providing a strong notion of closed-loop stability.

As discussed earlier, the property of convexity on the set of uncertainty is shown to play a crucial role in getting nice closed-loop properties. However, here we impose no such assumption. If the set of admissible parameters is not convex, the standard trick in adaptive control is to replace it with its closed convex hull. Unfortunately, often that set contains uncontrollable models (i.e. $b = 0$ in case of first-order plants). Here the key idea is to "cover" 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 two convex sets.

Proposition 4.1. *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 $\mathcal{S} \subset \mathcal{S}_1 \cup \mathcal{S}_2$.*

Proof of Proposition 4.1. For a given $\mathcal{S}$, define

$$\mathcal{S}_1 := \text{convex hull of } \left\{ \begin{bmatrix} a \\ b \end{bmatrix} \in \mathcal{S} : b > 0 \right\},$$

$$\mathcal{S}_2 := \text{convex hull of } \left\{ \begin{bmatrix} a \\ b \end{bmatrix} \in \mathcal{S} : b < 0 \right\}.$$

The result follows immediately. ∎

Remark 4.1. *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 4.1 also holds if we define*

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

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.

If $\mathcal{S}_1$ and $\mathcal{S}_2$ are large, it may be beneficial to have more than two, but smaller, convex sets. At this point we assume that

$$\mathcal{S} \subset \bigcup_{i=1}^{m} \mathcal{S}_i$$

and each set $\mathcal{S}_i$ is compact and convex and satisfies $\begin{bmatrix} a \\ 0 \end{bmatrix} \notin \mathcal{S}_i$ for every $a \in \mathbb{R}$. Before proceeding, for each $\theta^* \in \mathcal{S}_i$, $i = 1, 2, \ldots, m$, we define

$$i^*(\theta^*) = \min\left\{i \in \{1, 2, \ldots, m\} : \theta^* \in \mathcal{S}_i\right\};$$

when there is no ambiguity, we will drop the argument and simply write i^*.

4.2.1 Parameter Estimation

For each $\mathcal{S}_i$ and $\hat{\theta}_i(t_0) \in \mathcal{S}_i$, we design a projection-algorithm based estimator which generates an estimate $\hat{\theta}_i(t) \in \mathcal{S}_i$ at each $t > t_0$. The associated prediction error is defined as

$$e_i(t+1) = y(t+1) - \phi(t)^\top \hat{\theta}_i(t). \tag{4.2}$$

We apply the simplest form of the general algorithm presented in (3.9) of Chapter 3, where we set $\phi_m = \phi$ and $\delta = \infty$; observe that in this case

$$\rho_\delta(\phi(t), e_i(t+1)) = 1 \quad \Leftrightarrow \quad \phi(t) \neq 0.$$

The parameter estimation algorithm is as follows:

$$\check{\theta}_i(t+1) = \begin{cases} \hat{\theta}_i(t) & \text{if } \phi(t) = 0 \\ \hat{\theta}_i(t) + \dfrac{\phi(t)}{\|\phi(t)\|^2}\, e_i(t+1) & \text{otherwise;} \end{cases} \tag{4.3a}$$

$$\hat{\theta}_i(t+1) = \operatorname*{Proj}_{\mathcal{S}_i}\Big\{\check{\theta}_i(t+1)\Big\}. \tag{4.3b}$$

Define for each i the parameter estimation error $\tilde{\theta}_i(t) := \hat{\theta}_i(t) - \theta^*$. The following Proposition lists a property of the estimation algorithm (4.3) which follows directly from Proposition 3.1 of Chapter 3, where we set $\phi_m = \phi$ and $\delta = \infty$. Obviously, we do not know i^*.

Proposition 4.2. *For every $t_0 \in \mathbb{Z}$, $y_0 \in \mathbb{R}$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ $(i = 1, 2, \ldots, m)$, $\theta^* \in \mathcal{S}$, $w \in \boldsymbol{\ell}_\infty$, when the estimation algorithm in (4.3) is applied to the plant (4.1), the following holds:*

$$\|\tilde{\theta}_{i^*}(t)\|^2 \leq \|\tilde{\theta}_{i^*}(t_0)\|^2 - \frac{1}{2}\sum_{j=t_0,\phi(j)\neq 0}^{t-1} \frac{|e_{i^*}(j+1)|^2}{\|\phi(j)\|^2} + 2\sum_{j=t_0,\phi(j)\neq 0}^{t-1} \frac{|w(j)|^2}{\|\phi(j)\|^2}, \qquad t > t_0.$$

4.2.2 The Switching Controller

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

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

Let y^* be the reference signal to be tracked. We assume that it is known one step ahead, i.e. we know $y^*(t+1)$ 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:

$$u(t) = -\frac{\hat{a}_i(t)}{\hat{b}_i(t)} y(t) + \frac{1}{\hat{b}_i(t)} y^*(t+1),$$

which ensures that $y^*(t+1) = \phi(t)^\top \hat{\theta}_i(t)$. Here, of course, we do not know which $\mathcal{S}_i$ contains θ^*. In fact, θ^* may lie in more than one such set.

Let us define the index set

$$\mathcal{I}^* = \{1, 2, \ldots, m\}.$$

To this end, we define a switching signal $\sigma : \mathbb{Z} \mapsto \mathcal{I}^*$ which decides which controller to use at any given point in time, i.e. we set

$$u(t) = -\frac{\hat{a}_{\sigma(t)}(t)}{\hat{b}_{\sigma(t)}(t)} y(t) + \frac{1}{\hat{b}_{\sigma(t)}(t)} y^*(t+1). \tag{4.4}$$

To proceed, we define the tracking error ε by

$$\varepsilon(t) := y(t) - y^*(t).$$

Let us analyze the relationship between the tracking error and the prediction error when the above control is applied:

$$\begin{aligned} \varepsilon(t+1) &= y(t+1) - y^*(t+1) \\ &= y(t+1) - \phi(t)^\top \hat{\theta}_{\sigma(t)}(t) \\ &= e_{\sigma(t)}(t+1), \qquad t \geq t_0. \end{aligned} \tag{4.5}$$

So the choice of σ at time t affects the tracking error ε at time $t+1$.

What remains is to show how to choose $\sigma(t)$ to obtain the desired properties. We first present the main result of $m = 2$. Then after that we briefly discuss the case of $m > 2$.

4.3 The Main Result: The Case of $m = 2$

We begin with the case of two uncertainty sets, i.e. we have $\mathcal{I}^* = \{1, 2\}$. Here we adopt the following simple switching rule: with an initial condition of $\sigma(t_0) = \sigma_0$,

$$\sigma(t) = \operatorname*{argmin}_{i \in \mathcal{I}^*} |e_i(t)|, \qquad t > t_0, \tag{4.6}$$

i.e. we choose the model with the minimum prediction error. This rule is memoryless and is a function only of signals at the same instant. For the case when $|e_1(t)| = |e_2(t)|$, we (somewhat arbitrarily) select $\sigma(t)$ to be 1.

Before presenting the main result of this chapter, we first show that the simple logic in (4.6) yields a very desirable closed-loop property.

Lemma 4.1. *Consider the plant (4.1) with $m = 2$, and suppose that the adaptive controller consisting of the parameter estimator (4.3), the control law (4.4), and the switching rule (4.6) is applied. Then for every $t_0 \in \mathbb{Z}$, $y_0 \in \mathbb{R}$, $\sigma_0 \in \{1, 2\}$, $\theta^* \in \mathcal{S}$, and $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ $(i = 1, 2)$ and $y^*, w \in \boldsymbol{\ell}_\infty$, for every $j \in \{1, 2\}$ and $t \geq t_0 + 1$ we have that*

(a) $|\varepsilon(t)| \leq |e_j(t)|$ *or*

(b) $|\varepsilon(t + 1)| \leq |e_j(t + 1)|$.

Proof of Lemma 4.1. Fix $t_0 \in \mathbb{Z}$, $y_0 \in \mathbb{R}$, $\sigma_0 \in \{1, 2\}$, $\theta^* \in \mathcal{S}$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ $(i = 1, 2)$, and $y^*, w \in \boldsymbol{\ell}_\infty$, and let $j \in \{1, 2\}$ and $t \geq t_0 + 1$ be arbitrary.

Let $\bar{\jmath}$ be the element of $\{1, 2\}$ which is not j. Suppose that (b) fails to hold; in view of (4.5) it must be that $\sigma(t) = \bar{\jmath}$; from (4.6) this means that $|e_{\bar{\jmath}}(t)| \leq |e_j(t)|$. Since $\varepsilon(t) \in \{e_1(t), e_2(t)\}$, we conclude that $|\varepsilon(t)| \leq |e_j(t)|$, i.e. (a) holds. ■

In the above we do not make any claim that $\theta^* \in \mathcal{S}_{\sigma(t)}$ at any time; it only makes a statement about the size of the prediction error. Quite surprisingly, it turns out that this is enough to ensure that closed-loop stability is attained.

Now, we present the main result of this chapter.

Theorem 4.1. *Consider the plant (4.1) with $\mathcal{I}^* = \{1, 2\}$ and suppose that the adaptive controller consisting of the parameter estimator (4.3), the control law (4.4), and the switching rule (4.6) is applied. For every $\lambda \in (0, 1)$, there exists a constant $\gamma > 0$ such that for every $t_0 \in \mathbb{Z}$, $y_0 \in \mathbb{R}$, $\sigma_0 \in \mathcal{I}^*$, $\theta^* \in \mathcal{S}$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ $(i \in \mathcal{I}^*)$, and $y^*, w \in \boldsymbol{\ell}_\infty$, the closed-loop system satisfies*

$$\|\phi(t)\| \leq \gamma\lambda^{t-t_0}|y_0| + \sum_{j=t_0}^{t-1} \gamma\lambda^{t-1-j}(|w(j)| + |y^*(j + 1)|) + \gamma|y^*(t + 1)|, \qquad t \geq t_0. \tag{4.7}$$

The above result shows that the closed-loop system experiences linear-like behavior. There is a uniform exponential decay bound on the effect of the initial condition, and a convolution sum bound on the effect of the exogenous signals. If the initial condition is zero, there is a bounded gain on the map from the exogenous signals (the noise and reference signal) to ϕ in every p-norm; in classical adaptive control this is rarely the case. This is analogous to the result in [43] which deals with one convex uncertainty set and a single estimator. Furthermore, while the choice in (4.6) seems obvious, as far as we are aware there is no proof of stability in the literature for the situation in which the classical estimation algorithm (3.4) is used in conjunction with the control law (4.4).

Before proving the main result, define the constants

$$\bar{\mathbf{s}} := \max_i \|\mathcal{S}_i\|,$$

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

Remark 4.2. *The following proof will include a separate analysis for the case when no noise entering the system, i.e. $w = 0$, and the case when noise is entering. It is clear that the noise-free case is just a special case of the noisy case, but it is included in this chapter to help the reader understand the proof technique. In subsequent results of the thesis, proofs will only include the more general and complicated case of when noise enters the system.*

Proof of Theorem 4.1:

The proof is a significant extension of that of the main result of [43]. It will be given for the case when no noise enters the system, followed by the case with noise.

Fix $\lambda \in (0, 1)$. Let $t_0 \in \mathbb{Z}$, $y_0 \in \mathbb{R}$, $\sigma_0 \in \mathcal{I}^*$, $\theta^* \in \mathcal{S}$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ $(i \in \mathcal{I}^*)$, and $y^*, w \in \boldsymbol{\ell}_\infty$ be arbitrary.

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 (4.4) we obtain the general bound

$$\|\phi(t)\| \leq c_1|y(t)| + c_2|y^*(t+1)|; \tag{4.8}$$

if we define $c_3 := \max\{\bar{a} + \bar{b}\bar{f}, \bar{b}\bar{g}\}$ from the plant equation (4.1) we have the crude bound

$$|y(t+1)| \leq c_3|y(t)| + c_3|y^*(t+1)| + |w(t)|. \tag{4.9}$$

Case 1: $w(t) = 0$ for all $t \geq t_0$.

In this part, the proof has several steps. First, we will analyze the behavior for two consecutive instants. Then, we will consider the whole time horizon.

Using the fact that $\|\tilde{\theta}_{i^*}(t_0)\| \leq 2\|\mathcal{S}_{i^*}\|$, from Proposition 4.2 we have that:

$$\|\tilde{\theta}_{i^*}(t)\|^2 \leq \|\tilde{\theta}_{i^*}(t_0)\|^2 - \frac{1}{2} \sum_{j=t_0,\phi(j)\neq 0}^{t-1} \frac{|e_{i^*}(j+1)|^2}{\|\phi(j)\|^2}$$

$$\Rightarrow \sum_{j=t_0,\phi(j)\neq 0}^{t-1} \frac{|e_{i^*}(j+1)|^2}{\|\phi(j)\|^2} \leq 2\|\tilde{\theta}_{i^*}(t_0)\|^2 \leq 8\|\mathcal{S}_{i^*}\|^2 \leq \underbrace{8\bar{\mathbf{s}}^2}_{=:\, c_4}, \quad t \geq t_0 + 1. \tag{4.10}$$

For $k \geq t_0$, define

$$\alpha_k := \begin{cases} \dfrac{|e_{i^*}(k+1)|}{\|\phi(k)\|} & \text{if } \phi(k) \neq 0 \\ 0 & \text{otherwise.} \end{cases} \tag{4.11}$$

For $\phi(j) \neq 0$, we have

$$|e_{i^*}(j+1)| = \alpha_j \|\phi(j)\|. \tag{4.12}$$

For $\phi(j) = 0$, we have $y(j) = u(j) = 0$; from (4.1) we conclude that $y(j+1) = 0$, and from (4.2) we conclude that $e_i(j+1) = 0$ for all $i \in \mathcal{I}^*$, which means that (4.12) holds for this case as well. Using (4.8), we get

$$|e_{i^*}(j+1)| \leq \alpha_j (c_1 |y(j)| + c_2 |y^*(j+1)|). \tag{4.13}$$

Motivated by Lemma 4.1, now we will analyze closed-loop behavior on two consecutive instants of time. Let $j \in \mathbb{Z}^+$ be arbitrary; as $i^* \in \mathcal{I}^*$, from Lemma 4.1 we have that either

$$|\varepsilon(t_0 + 2j + 1)| \leq |e_{i^*}(t_0 + 2j + 1)|$$

or

$$|\varepsilon(t_0 + 2j + 2)| \leq |e_{i^*}(t_0 + 2j + 2)|.$$

If we combine these with (4.13), use the fact that $\varepsilon(t) = y(t) - y^*(t)$, note from (4.10) that $\alpha_k \leq \sqrt{c_4}$, and define $\bar{\alpha}_{t_0+2j} := \max\{\alpha_{t_0+2j}, \alpha_{t_0+2j+1}\}$, we can conclude that either

$$|y(t_0 + 2j + 1)| \leq c_1 \bar{\alpha}_{t_0+2j} |y(t_0 + 2j)| + (1 + c_2 c_4^{\frac{1}{2}}) |y^*(t_0 + 2j + 1)|$$

or

$$|y(t_0+2j+2)| \le c_1\bar{\alpha}_{t_0+2j}|y(k_0+2j+1)| + (1+c_2c_4^{\frac{1}{2}})|y^*(t_0+2j+2)|.$$

Combining this with the crude bound (4.9) and defining $c_5 := \max\{1, c_1c_3\}$ and $c_6 := 2c_3 + c_1c_3c_4^{\frac{1}{2}} + c_2c_4^{\frac{1}{2}} + c_2c_3c_4^{\frac{1}{2}} + 1$, we see that in either case,

$$|y(t_0+2j+2)| \le c_5\bar{\alpha}_{t_0+2j}|y(t_0+2j)| + c_6(|y^*(t_0+2j+1)| + |y^*(t_0+2j+2)|), \quad j \in \mathbb{Z}^+. \tag{4.14}$$

Now we examine behavior across the whole time horizon. Observe from (4.10) that

$$\sum_{j=0}^{\infty} \bar{\alpha}_{t_0+2j}^2 \le \sum_{k=t_0}^{\infty} \alpha_k^2 \le c_4.$$

Now define

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

We now utilize the inequality of arithmetic and geometric means:

$$\begin{aligned}
\prod_{j=q}^{p-1} \bar{\alpha}_{t_0+2j} &\le \left[\frac{1}{p-q}\sum_{j=q}^{p-1}\bar{\alpha}_{t_0+2j}^2\right]^{\frac{p-q}{2}} \\
&\le \left[\frac{c_4}{p-q}\right]^{\frac{p-q}{2}} \\
&= \left[\left(\frac{c_4}{p-q}\right)^{\frac{1}{2}}\right]^{p-q}, \qquad 0 \le q < p.
\end{aligned} \tag{4.15}$$

With $\bar{k} := \left\lceil \left(\frac{c_4}{\lambda_1}\right)^2 \right\rceil$, we have

$$\frac{c_4}{\bar{k}} \le \lambda_1^2;$$

so it easy to see that

$$\left[\left(\frac{c_4}{k}\right)^{\frac{1}{2}}\right]^k \le \lambda_1^k, \qquad k \ge \bar{k}. \tag{4.16}$$

Since $\frac{c_4}{k}$ decreases as $k \geq 1$ increases, if we define $c_7 := \max\left\{1, c_4^{\frac{\bar{k}}{2}}\right\}$, then

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

as well. If we combine this with (4.16), from (4.15) we conclude that

$$\prod_{j=q}^{p-1} \bar{\alpha}_{t_0+2j} \leq c_7 \lambda_1^{p-q}, \qquad 0 \leq q < p. \tag{4.17}$$

Then by the definition of λ_1 we obtain

$$\begin{aligned}\prod_{j=q}^{p-1} [c_5 \bar{\alpha}_{t_0+2j}] &\leq c_7 \lambda_1^{p-q} c_5^{p-q} \\ &\leq c_7 \lambda^{2(p-q)}, \qquad 0 \leq q < p.\end{aligned} \tag{4.18}$$

Now we solve the difference inequality (4.14) recursively and apply the above bound (4.18): we obtain

$$|y(t_0+2j)| \leq c_7 \lambda^{2j} |y(t_0)| + \sum_{l=0}^{j-1} c_7 c_6 \lambda^{2(j-l-1)} (|y^*(t_0+2l+1)| + |y^*(t_0+2l+2)|), \qquad j \in \mathbb{Z}^+,$$

which simplifies to

$$|y(t_0+2j)| \leq c_7 \lambda^{2j} |y(t_0)| + \sum_{l=0}^{2j-1} \frac{c_7 c_6}{\lambda} \lambda^{2j-l-1} |y^*(t_0+l+1)|, \qquad j \in \mathbb{Z}^+. \tag{4.19}$$

We can use (4.9) 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

$$|y(t)| \leq \bar{\gamma}_1 \lambda^{t-t_0} |y(t_0)| + \sum_{j=t_0}^{t-1} \bar{\gamma}_1 \lambda^{t-j-1} |y^*(j+1)|, \qquad t \geq t_0. \tag{4.20}$$

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

We now analyze the case when there is noise entering the system; this is more complicated since $\|\tilde{\theta}_{i^*}(t)\|^2$ is no longer monotonically decreasing. Motivated by Case 1, also define

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

Following [43] and [46], we partition the timeline into two parts: one in which $w(\cdot)$ is small versus $\phi(\cdot)$ and one where it is not. Before proceeding, define

$$\nu := \left(\frac{\lambda_1}{4}\right)^2.$$

Let us now define two sets in relation to size of the noise $w(\cdot)$:

$$S_{good} = \left\{ j \geq t_0 : \phi(j) \neq 0 \text{ and } \frac{|w(j)|^2}{\|\phi(j)\|^2} < \nu \right\},$$

$$S_{bad} = \left\{ j \geq t_0 : \phi(j) = 0 \text{ or } \frac{|w(j)|^2}{\|\phi(j)\|^2} \geq \nu \right\};$$

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

Now we partition the time index $\{j \in \mathbb{Z} : j \geq t_0\}$ into intervals which oscillate between S_{good} and S_{bad}. We can clearly define a (possibly infinite) sequence of intervals of the form $[k_l, k_{l+1})$ 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_{good} or S_{bad}; and
iii) if $k_{l+1} \neq \infty$ and $[k_l, k_{l+1})$ belongs to S_{good} then $[k_{l+1}, k_{l+2})$ belongs to S_{bad} and vice versa.

Case 2a: $[k_l, k_{l+1})$ belongs to S_{bad}.

Let $j \in [k_l, k_{l+1})$ be arbitrary. So we have $\|\phi(j)\| = 0$ or $\frac{|w(j)|^2}{\|\phi(j)\|^2} \geq \nu$, so in either case $\|\phi(j)\| \leq \frac{1}{\sqrt{\nu}}|w(j)|$. If we define $c_8 := \frac{1}{\sqrt{\nu}}$ and utilize the definition of $\phi(j)$ we conclude that

$$|y(j)| \leq c_8 |w(j)|.$$

Also, from (4.1) we have

$$|y(j)| \leq \underbrace{\left((\bar{a} + \bar{b})\frac{1}{\sqrt{\nu}} + 1\right)}_{=: c_9} |w(j-1)|, \qquad j = k_l + 1, k_l + 2, \ldots, k_{l+1}.$$

We conclude that for $j \in [k_l, k_{l+1})$, we have

$$|y(j)| \leq \begin{cases} c_8|w(j)| & j = k_l \\ c_9|w(j-1)| & j = k_l+1, k_l+2, \ldots, k_{l+1}. \end{cases} \tag{4.21}$$

Case 2b: $[k_l, k_{l+1})$ belongs to S_{good}.

Using the same notation as in Case 1, we define

$$\alpha_k := \frac{|e_{i^*}(k+1)|}{\|\phi(k)\|};$$

for $j \in \mathbb{Z}^+$ so that $k_l + 2j + 1 < k_{l+1}$, we define

$$\bar{\alpha}_{k_l+2j} := \max\{\alpha_{k_l+2j}, \alpha_{k_l+2j+1}\}.$$

From Proposition 4.2 we have

$$\|\tilde{\theta}_{i^*}(\bar{k})\|^2 \leq \|\tilde{\theta}_{i^*}(\underline{k})\|^2 - \frac{1}{2}\sum_{j=\underline{k}}^{\bar{k}-1} \frac{|e_i(j+1)|^2}{\|\phi(j)\|^2} + 2\sum_{j=\underline{k}}^{\bar{k}-1} \frac{|w(j)|^2}{\|\phi(j)\|^2}, \qquad k_l \leq \underline{k} < \bar{k} \leq k_{l+1}, \tag{4.22}$$

which yields

$$\sum_{k=\underline{k}}^{\bar{k}-1} \alpha_k^2 \leq 8\bar{\mathbf{s}}^2 + 4(\bar{k} - \underline{k})\nu, \qquad k_l \leq \underline{k} < \bar{k} \leq k_{l+1},$$

or in other words:

$$\sum_{j=q}^{p-1} \bar{\alpha}_{k_l+2j}^2 \leq 8\bar{\mathbf{s}}^2 + 4[(k_l+2p) - (k_l+2q)]\nu = 8\bar{\mathbf{s}}^2 + 8(p-q)\nu,$$
$$\text{for all } q, p \in \mathbb{Z}^+ \text{ s.t. } k_l \leq k_l + 2q < k_l + 2p \leq k_{l+1}. \tag{4.23}$$

Note from the above that $\alpha_k \leq \sqrt{8\bar{\mathbf{s}}^2 + 4\nu} := c_{10}$.

If we now analyze the closed-loop system as in the noise-free case, we end up with a version of (4.14) with the noise now included: there exists a constant $c_{11} := 2c_3 + c_1c_3c_{10}^{\frac{1}{2}} + c_2c_{10}^{\frac{1}{2}} + c_2c_3c_{10}^{\frac{1}{2}} + c_1c_{10}^{\frac{1}{2}} + 1$ so that

$$|y(k_l+2j+2)| \leq c_5\bar{\alpha}_{k_l+2j}|y(k_l+2j)| + c_{11}\Big(|y^*(k_l+2j+1)| + |y^*(k_l+2j+2)| +$$

$$|w(k_l+2j)| + |w(k_l+2j+1)|\Big), \qquad j \in \mathbb{Z}^+ \text{ s.t. } k_l + 2j + 1 < k_{l+1}. \tag{4.24}$$

By similar analysis like in Case 1, we use the inequality of arithmetic and geometric means; to this end, from (4.23) and incorporating the definition of ν:

$$\begin{aligned}\prod_{j=q}^{p-1} \bar{\alpha}_{k_l+2j} &\le \left[\frac{1}{p-q}\sum_{j=q}^{p-1} \bar{\alpha}^2_{k_l+2j}\right]^{\frac{p-q}{2}} \\ &\le \left[\frac{8\bar{\mathbf{s}}^2}{p-q} + 8\nu\right]^{\frac{p-q}{2}} \\ &= \left[\left(\frac{8\bar{\mathbf{s}}^2}{p-q} + \frac{\lambda_1^2}{2}\right)^{\frac{1}{2}}\right]^{p-q}, \\ &\qquad \text{for all } q, p \in \mathbb{Z}^+ \text{ s.t. } k_l \le k_l + 2q < k_l + 2p \le k_{l+1}.\end{aligned} \tag{4.25}$$

With $\bar{k} := \left\lceil \left(\frac{4\bar{\mathbf{s}}}{\lambda_1}\right)^2 \right\rceil$, we have

$$\frac{8\bar{\mathbf{s}}^2}{\bar{k}} \le \frac{\lambda_1^2}{2},$$

which means that

$$\left[\left(\frac{8\bar{\mathbf{s}}^2}{k} + \frac{\lambda_1^2}{2}\right)^{\frac{1}{2}}\right]^{k} \le \lambda_1^k, \qquad k \ge \bar{k}.$$

Then in a similar manner to that of Case 1, if we define $c_{12} := (8\bar{\mathbf{s}}^2 + 1)^{\frac{\bar{k}}{2}}$, it is easy to see that

$$\prod_{j=q}^{p-1} \bar{\alpha}_{k_l+2j} \le c_{12}\lambda_1^{p-q}, \qquad \text{for all } q, p \in \mathbb{Z}^+, \text{ s.t. } k_l \le k_l + 2q < k_l + 2p \le k_{l+1}. \tag{4.26}$$

Then by the definition of λ_1 we obtain

$$\begin{aligned}\prod_{j=q}^{p-1} [c_5\bar{\alpha}_{k_l+2j}] &\le c_{12}\lambda_1^{p-q}c_5^{p-q} \\ &\le c_{12}\lambda^{2(p-q)}, \qquad \text{for all } q, p \in \mathbb{Z}^+, \text{ s.t. } k_l \le k_l + 2q < k_l + 2p \le k_{l+1}.\end{aligned} \tag{4.27}$$

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 (4.24) and obtain a bound which is valid on $t = k_l, \ldots, k_{l+1} - 1$ 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 $t = k_l, \ldots, k_{l+1}$. So in any case, we now proceed to solve (4.24) iteratively and apply the bound in (4.27). 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

$$|y(t)| \leq \bar{\gamma}_2 \lambda^{t-k_l} |y(k_l)| + \sum_{j=k_l}^{t-1} \bar{\gamma}_2 \lambda^{t-j-1} (|y^*(j+1)| + |w(j)|), \quad t = k_l, k_l + 1, \ldots, k_{l+1} - 1. \tag{4.28}$$

Note that (4.28) does not apply for $t = k_{l+1}$; so to conclude Case 2b, define $\bar{\gamma}_3 := c_3 \max\{1, \frac{\bar{\gamma}_2}{\lambda}\}$ and utilizing (4.9) to obtain a bound accounting for the extra step yields

$$|y(t)| \leq \bar{\gamma}_3 \lambda^{t-k_l} |y(k_l)| + \sum_{j=k_l}^{t-1} \bar{\gamma}_3 \lambda^{t-j-1} (|y^*(j+1)| + |w(j)|), \quad t = k_l, k_l + 1, \ldots, k_{l+1}. \tag{4.29}$$

Finally, we will combine the results of Case 2a and Case 2b to find a general bound on y. Before proceeding, define

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

Claim 4.1. *The following bound holds:*

$$|y(t)| \leq \bar{\gamma}_4 \lambda^{t-t_0} |y(t_0)| + \sum_{j=t_0}^{t-1} \bar{\gamma}_4 \lambda^{t-j-1} (|y^*(j+1)| + |w(j)|), \qquad t \geq t_0. \tag{4.30}$$

Proof of Claim 4.1. If $[k_0, k_1) = [t_0, k_1) \subset S_{good}$, then (4.30) is true for $t \in [k_0, k_1]$ by (4.29). If $[k_0, k_1) \subset S_{bad}$, then from (4.21) we have

$$|y(j)| \leq \begin{cases} |y(k_0)| = |y(t_0)| & j = k_0 \\ c_9 |w(j-1)| & j = k_0 + 1, k_0 + 2, \ldots, k_1. \end{cases}$$

which means that (4.30) holds on $[k_0, k_1]$ for this case as well.

We now use induction: suppose that (4.30) 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 $t \in [k_l, k_{l+1}) \subset S_{bad}$, then from (4.21) we see that

$$|y(t)| \leq c_9 |w(t-1)|, \; t = k_l + 1, k_l + 2, \ldots, k_{l+1},$$

so (4.30) clearly holds on $(k_l, k_{l+1}]$. On the other hand, if $[k_l, k_{l+1}) \subset S_{good}$, then $k_l - 1 \in S_{bad}$; from (4.21) we have

$$|y(k_l)| \leq c_9 |w(k_l - 1)|.$$

Using (4.29) to analyze the behavior on $t \in [k_l, k_{l+1}]$, we have

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

which implies that (4.30) holds. □

At this point we have bounds on $y(\cdot)$ for both cases with noise and without. To combine the bounds (4.20) and (4.30), define $\bar{\gamma}_5 := \max\{\bar{\gamma}_1, \bar{\gamma}_4\}$. Then the overall bound is given by

$$|y(t)| \leq \bar{\gamma}_5 \lambda^{t-t_0} |y(t_0)| + \sum_{j=t_0}^{t-1} \bar{\gamma}_5 \lambda^{t-j-1} (|y^*(j+1)| + |w(j)|), \qquad t \geq t_0. \tag{4.31}$$

To conclude the proof of Theorem 4.1, we need a bound on $u(\cdot)$: using (4.4) we obtain

$$|u(t)| \leq \bar{f} |y(t)| + \bar{g} |y^*(t+1)|;$$

so by substituting (4.31) into the above, we get

$$|u(t)| \leq \bar{f} \bar{\gamma}_5 \lambda^{t-t_0} |y(t_0)| + \sum_{j=t_0}^{t-1} \bar{f} \bar{\gamma}_5 \lambda^{t-j-1} |w(j)| + \sum_{j=t_0}^{t} \left(\frac{\bar{f}\bar{\gamma}_5}{\lambda} + \bar{g} \right) \lambda^{t-j} |y^*(j+1)|, \qquad t \geq t_0. \tag{4.32}$$

By combining (4.32) and (4.31) and defining $\gamma := \max\{\bar{\gamma}_5, \bar{f}\bar{\gamma}_5, \frac{\bar{f}\bar{\gamma}_5}{\lambda} + \bar{g}\}$, we conclude the proof. ■

4.4 The Case of $m \geq 2$

Now we consider the case of $m > 2$ uncertainty sets. As mentioned earlier, it may be beneficial for performance to have more than two sets. Unfortunately, although the rule in (4.6) is a well-defined rule in this case, and it works in all of our simulations, we have been unable to prove that it will work. In particular, a potential problem is that the algorithm could oscillate between two bad choices, and never (or rarely) choose the correct one; it is not clear that Lemma 4.1 would hold. Instead, we propose a modified version of (4.6). At each point in time we have an admissible set $\mathcal{I}(t)$: we initialize $\mathcal{I}(t_0) = \mathcal{I}^*$, and we obtain $\mathcal{I}(t+1)$ from $\mathcal{I}(t)$ by removing all $j \in \mathcal{I}(t)$ satisfying

$$|\varepsilon(t+1)| \leq |e_j(t+1)|,$$

clearly $j = \sigma(t)$ satisfies this bound, but more j's may as well; if this results in $\mathcal{I}(t+1)$ being empty, then we **reset** $\mathcal{I}(t+1)$ to be $\mathcal{I}^*$. This *Switching Algorithm*[2] is summarized as follows: with $\sigma(t_0) = \sigma_0$ and $\mathcal{I}(t_0) = \mathcal{I}^*$:

$$\hat{\mathcal{I}}(t) = \{i \in \mathcal{I}^* \ : \ |e_i(t+1)| < |\varepsilon(t+1)|\}, \tag{4.33a}$$

$$\mathcal{I}(t+1) = \begin{cases} \mathcal{I}^* & \text{if } \mathcal{I}(t) \cap \hat{\mathcal{I}}(t) = \varnothing \\ \mathcal{I}(t) \cap \hat{\mathcal{I}}(t) & \text{otherwise,} \end{cases} \tag{4.33b}$$

$$\sigma(t+1) = \operatorname*{argmin}_{i \in \mathcal{I}(t+1)} |e_i(t+1)|, \qquad t \geq t_0. \tag{4.33c}$$

Remark 4.3. *We define the* **index set reset times** *as those $k \geq t_0$ for which $\mathcal{I}(k) = \mathcal{I}^*$.*

Remark 4.4. *In computing the* argmin *in the RHS of* (4.33c)*, it could very well that there are more values $i \in \mathcal{I}(t+1)$ which achieves the minimum. In such a case, we (somewhat arbitrarily) choose the smallest such index.*

Lemma 4.2. *Consider the plant* (4.1) *for which $m \geq 2$, and suppose that the adaptive controller consisting of the parameter estimator* (4.3)*, the control law* (4.4)*, and the switching algorithm* (4.33) *is applied. Then for every $t_0 \in \mathbb{Z}$, $y_0 \in \mathbb{R}$, $\sigma_0 \in \mathcal{I}^*$, $\theta^* \in \mathcal{S}$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ $(i \in \mathcal{I}^*)$, $y^*, w \in \boldsymbol{\ell}_\infty$, if $\underline{k}$ and $\bar{k}$ are two consecutive index set reset times, there exists a $\hat{k} \in [\underline{k}, \bar{k})$ such that:*

$$|\varepsilon(\hat{k}+1)| \leq |e_{i^*}(\hat{k}+1)|. \tag{4.34}$$

[2] This algorithm is reminiscent of the localization-based algorithm of [88],[87]; though, in that work there are no resets (to $\mathcal{I}^*$).

Remark 4.5. *Lemma 4.2 says that, between every two index set resets, there is an instant for which the tracking error is equal to, or smaller than, the prediction error associated with the correct index.*

Proof of Lemma 4.2. Let $\theta^* \in \mathcal{S}$, $t_0 \in \mathbb{Z}$, $\sigma_0 \in \mathcal{I}^*$, $y_0 \in \mathbb{R}$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ $(i \in \mathcal{I}^*)$, and $w, y^* \in \boldsymbol{\ell}_\infty$ be arbitrary. Let $\bar{k}$ and $\underline{k}$ be two consecutive reset times.

We prove (4.34) by contradiction; assume that

$$|\varepsilon(j+1)| > |e_{i^*}(j+1)|, \quad \text{for all } j \in [\underline{k}, \bar{k}). \tag{4.35}$$

Then, according to (4.33a), we should have

$$i^* \in \hat{\mathcal{I}}(j), \qquad j \in [\underline{k}, \bar{k}). \tag{4.36}$$

We know by the definition of index resets that for all $j \in (\underline{k}, \bar{k})$ we have $\mathcal{I}(j) \neq \mathcal{I}^*$, which means that by (4.33b)

$$\mathcal{I}(j) = \mathcal{I}(j-1) \cap \hat{\mathcal{I}}(j-1), \qquad j \in (\underline{k}, \bar{k});$$

then by induction we see that

$$\mathcal{I}(j) = \mathcal{I}(\underline{k}) \cap \hat{\mathcal{I}}(\underline{k}) \cap \hat{\mathcal{I}}(\underline{k}+1) \cap \cdots \cap \hat{\mathcal{I}}(j-2) \cap \hat{\mathcal{I}}(j-1), \qquad j \in (\underline{k}, \bar{k}).$$

But $\mathcal{I}(\underline{k}) = \mathcal{I}^*$, so using (4.36) in the above, we see that

$$i^* \in \mathcal{I}(j), \qquad j \in [\underline{k}, \bar{k}) \tag{4.37}$$

as well. So according to this and to (4.36) we have $i^* \in \mathcal{I}(\bar{k}-1) \cap \hat{\mathcal{I}}(\bar{k}-1)$. However, we know by the definition of index resets and (4.33b) that $\mathcal{I}(\bar{k}-1) \cap \hat{\mathcal{I}}(\bar{k}-1) = \varnothing$, which is a contradiction, so it must be that (4.35) does not hold. ■

We now present the following main result for the case when we have more than two estimators.

Theorem 4.2. *Consider the plant in (4.1) with $\mathcal{I}^* = \{1, 2, \ldots, m\}$ and suppose that the adaptive controller consisting of the parameter estimator (4.3), the control law (4.4), and the switching algorithm (4.33) is applied. For every $\lambda \in (0,1)$, there exists a constant $\bar{\gamma} > 0$ such that for every $t_0 \in \mathbb{Z}$, $y_0 \in \mathbb{R}$, $\sigma_0 \in \mathcal{I}^*$, $\theta^* \in \mathcal{S}$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ ($i \in \mathcal{I}^*$) and $y^*, w \in \boldsymbol{\ell}_\infty$, the closed-loop system satisfies*

$$\|\phi(t)\| \leq \bar{\gamma}\lambda^{t-t_0}|y_0| + \sum_{j=t_0}^{t-1} \bar{\gamma}\lambda^{t-1-j}(|y^*(j+1)| + |w(j)|) + \bar{\gamma}|y^*(t+1)|, \qquad t \geq t_0. \tag{4.38}$$

Proof of Theorem 4.2. We apply an analysis similar to that of the proof of Theorem 4.1; instead of analyzing two consecutive instants, we analyze intervals between index set resets. We apply Lemma 4.2; we further utilize the fact that the maximum length between any consecutive resets is not more than m. We omit the details here and refer the reader to the main result of Chapter 7 that includes a similar proof. ■

4.5 Robustness Results

Here we show that we can leverage the fact that a convolution bound holds in the case of a fixed plant parameter to prove that a convolution bound (with larger constants) also holds if we allow time-variation and/or unmodelled dynamics. To proceed, we consider a time-varying version of the plant (4.1) along with the term $d_\Delta(t) \in \mathbb{R}$ added to represent the unmodelled dynamics:

$$\begin{aligned} y(t+1) &= a(t)y(t) + b(t)u(t) + w(t) + d_\Delta(t) \\ &= \underbrace{\begin{bmatrix} y(t) \\ u(t) \end{bmatrix}^\top}_{=\phi(t)^\top} \underbrace{\begin{bmatrix} a(t) \\ b(t) \end{bmatrix}}_{=:\theta^*(t)} + w(t) + d_\Delta(t), \quad t \in \mathbb{Z}; \end{aligned} \tag{4.39}$$

as discussed in Chapter 2, we assume that d_Δ satisfies

$$\mathfrak{w}(t+1) = \beta\mathfrak{w}(t) + \beta\|\phi(t)\|, \qquad \mathfrak{w}(t_0) = \mathfrak{w}_0 \tag{4.40a}$$

$$|d_\Delta(t)| \leq \mu\mathfrak{w}(t) + \mu\|\phi(t)\|, \qquad t \geq t_0. \tag{4.40b}$$

The following result addresses the problem when $\mathcal{I}^* = \{1, 2\}$, i.e. $m = 2$. Naturally, the same result is true for when $m > 2$ when the switching algorithm (4.33) is applied.

Theorem 4.3. *Suppose that the adaptive controller (4.3), (4.4) and (4.6) is applied to the time-varying plant (4.39) with d_Δ satisfying (4.40). Then, for every $\beta \in (0,1)$ and $\bar{c}_0 \geq 0$, there exist $\bar{\epsilon} > 0$, $\mu > 0$, $\tilde{\lambda} \in (\beta, 1)$ and $\tilde{\gamma} > 0$ such that for every $t_0 \in \mathbb{Z}$, $y_0 \in \mathbb{R}$, $\sigma_0 \in \mathcal{I}^*$, $\theta^* \in \mathbb{S}(\mathcal{S}, \bar{c}_0, \bar{\epsilon})$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ ($i \in \mathcal{I}^*$), and $w, y^* \in \boldsymbol{\ell}_\infty$, the following holds:*

$$\left\| \begin{bmatrix} \phi(t) \\ \mathfrak{w}(t) \end{bmatrix} \right\| \leq \tilde{\gamma} \tilde{\lambda}^{t-t_0} \left\| \begin{bmatrix} \phi(t_0) \\ \mathfrak{w}_0 \end{bmatrix} \right\| + \sum_{j=t_0}^{t-1} \tilde{\gamma} \tilde{\lambda}^{t-j-1} (|w(j)| + |y^*(j+1)|) + \tilde{\gamma} |y^*(t+1)|, \quad t \geq t_0.$$

Proof of Theorem 4.3. We observe here that the plant (4.39) and the controller (4.3), (4.4) and (4.6) fit into the paradigm of Chapter 2: we set

$$\begin{aligned} \vartheta(t) = \phi(t) &= \begin{bmatrix} y(t) \\ u(t) \end{bmatrix}, \\ f(\vartheta(\cdot)) &= \phi(\cdot), \\ z(t) &= \varnothing, \\ \hat{\theta}(t) &= \begin{bmatrix} \hat{\theta}_1(t) \\ \hat{\theta}_2(t) \end{bmatrix}, \\ r(t) &= y^*(t+1), \\ \Omega &= \mathcal{S}_1 \times \mathcal{S}_2. \end{aligned}$$

In Theorem 4.1 it is proven the controller (4.3), (4.4) and (4.6) provides a convolution bound for (4.1). Then, by Theorems 2.2, 2.2 and 2.3 we immediately see that the same is true in the presence of time-variation and/or unmodelled dynamics. ∎

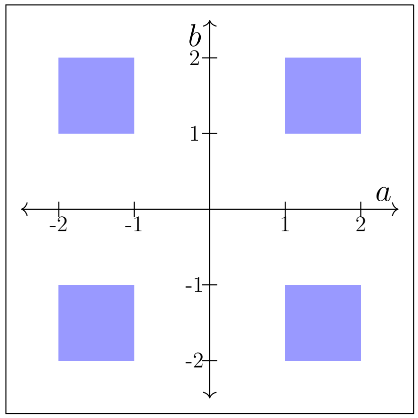


Figure 4.1: Uncertainty set $\mathcal{S}$ (shaded area).

4.6 Simulation Examples

In this section, simulation examples are provided to illustrate the results of this chapter. Consider the time-varying plant

$$y(t+1) = \underbrace{\begin{bmatrix} y(t) \\ u(t) \end{bmatrix}^\top}_{=:\phi(t)^\top} \underbrace{\begin{bmatrix} a(t) \\ b(t) \end{bmatrix}}_{=:\theta^*(t)} + w(t),$$

with $\theta^*(t)$ belonging to the uncertainty set $\mathcal{S}$:

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

which can be visualized in Figure 4.1. Hence, every admissible model is unstable, and the sign of the input gain b is unknown.

4.6.1 Example 1

Here the plant parameters are varying as follows:

$$a(t) = \begin{cases} -\frac{3}{2} - \frac{1}{2}\sin(\frac{1}{20}t), & 51 \le t \le 100, 151 \le t \le 200 \\ \frac{3}{2} + \frac{1}{2}\sin(\frac{1}{20}t), & \text{otherwise}, \end{cases}$$

$$b(t) = \begin{cases} -\frac{3}{2} - \frac{1}{2}\cos(\frac{1}{15}t), & 101 \le t \le 150, 151 \le t \le 200 \\ \frac{3}{2} + \frac{1}{2}\cos(\frac{1}{15}t), & \text{otherwise.} \end{cases}$$

In the first approach, we define two convex sets by convexifying the 1st and 2nd quadrants and the 3rd and 4th quadrants, respectively, yielding

$$\mathcal{S}_1 := \left\{ \begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{R}^2 : a \in [-2, 2], b \in [1, 2] \right\}, \tag{4.41a}$$

$$\mathcal{S}_2 := \left\{ \begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{R}^2 : a \in [-2, 2], b \in [-2, -1] \right\}. \tag{4.41b}$$

We will apply the control law (4.4) using estimates from (4.3) and $\sigma(t)$ determined by (4.6). We set $\hat{\theta}_1(0) = [1.5\ 1.5]^\top, \hat{\theta}_2(0) = [-1.5\ \ -1.5]^\top, \sigma_0 = 2, y_0 = -1$, the reference $y^*(\cdot)$ to be a unit-amplitude square wave of period 65, and noise to be $w(t) = \frac{1}{20}\sin(5t)$. Figures 4.2 and 4.3 display the results. We see that the controller does a reasonable job, even though the switching occasionally chooses the wrong model. Large transient may ensue, but on average the adaptive controller provides good tracking.

As mentioned earlier, it may be beneficial to have more than two convex sets. So in the second approach, we define four convex sets in the following natural way:

$$\mathcal{S}_1 := \left\{ \begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{R}^2 : a \in [1, 2], b \in [1, 2] \right\},$$
$$\mathcal{S}_2 := \left\{ \begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{R}^2 : a \in [1, 2], b \in [-2, -1] \right\},$$
$$\mathcal{S}_3 := \left\{ \begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{R}^2 : a \in [-2, -1], b \in [1, 2] \right\},$$
$$\mathcal{S}_4 := \left\{ \begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{R}^2 : a \in [-2, -1], b \in [-2, -1] \right\}.$$

We will apply the control law (4.4) using estimates from (4.3) and $\sigma(t)$ determined by (4.33). We set $\hat{\theta}_1(0) = [1.5\ \ 1.5]^\top, \hat{\theta}_2(0) = [1.5\ \ -1.5]^\top, \hat{\theta}_3(0) = [-1.5\ \ 1.5]^\top, \hat{\theta}_4(0) = [-1.5\ \ -1.5]^\top, \sigma_0 = 2$. As above we set $y_0 = -1$, the reference $y^*(\cdot)$ to be a unit-amplitude square wave of period 65, and noise to be $w(t) = \frac{1}{20}\sin(5t)$. Figures 4.4 and 4.5 display the results. We see the controller does a good job of tracking, and with smaller transients than in the first approach. Furthermore, the estimator does a fairly good job of tracking the time-varying parameter. Both examples illustrate that the approach handles time-variation and occasional jumps.

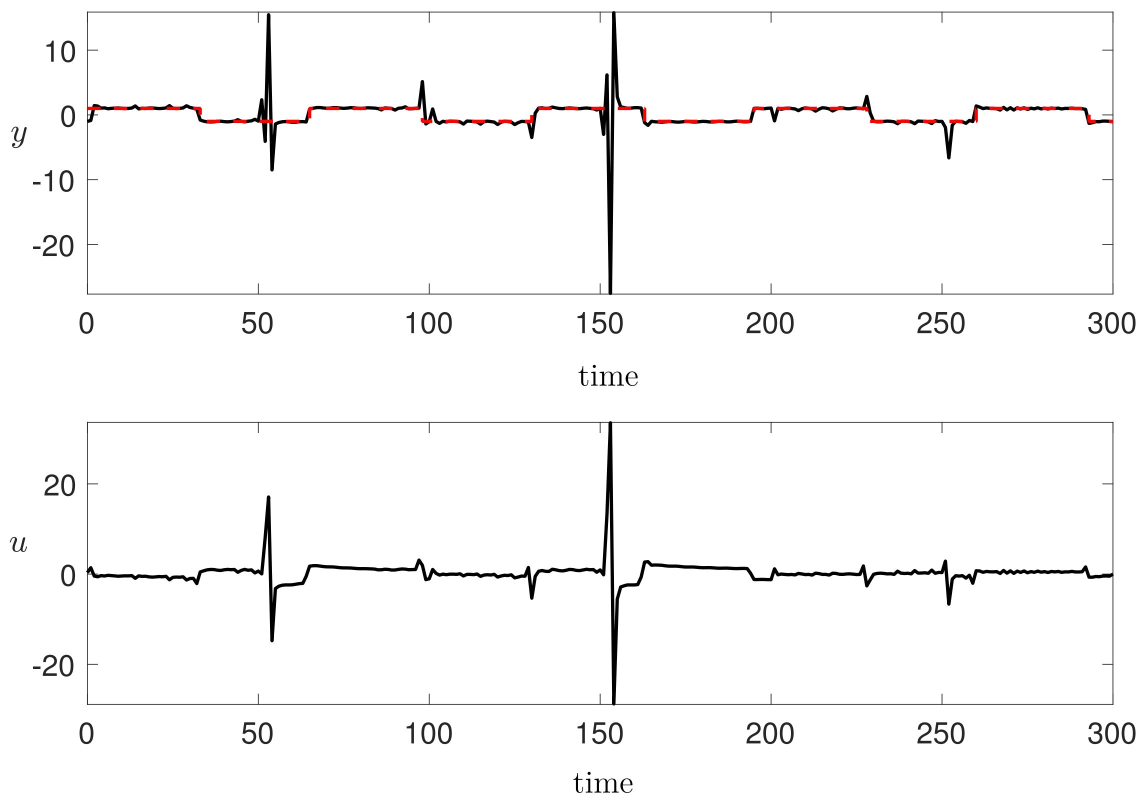


Figure 4.2: The upper plot shows both the reference (dashed) and the output (solid); the lower plot shows control input.

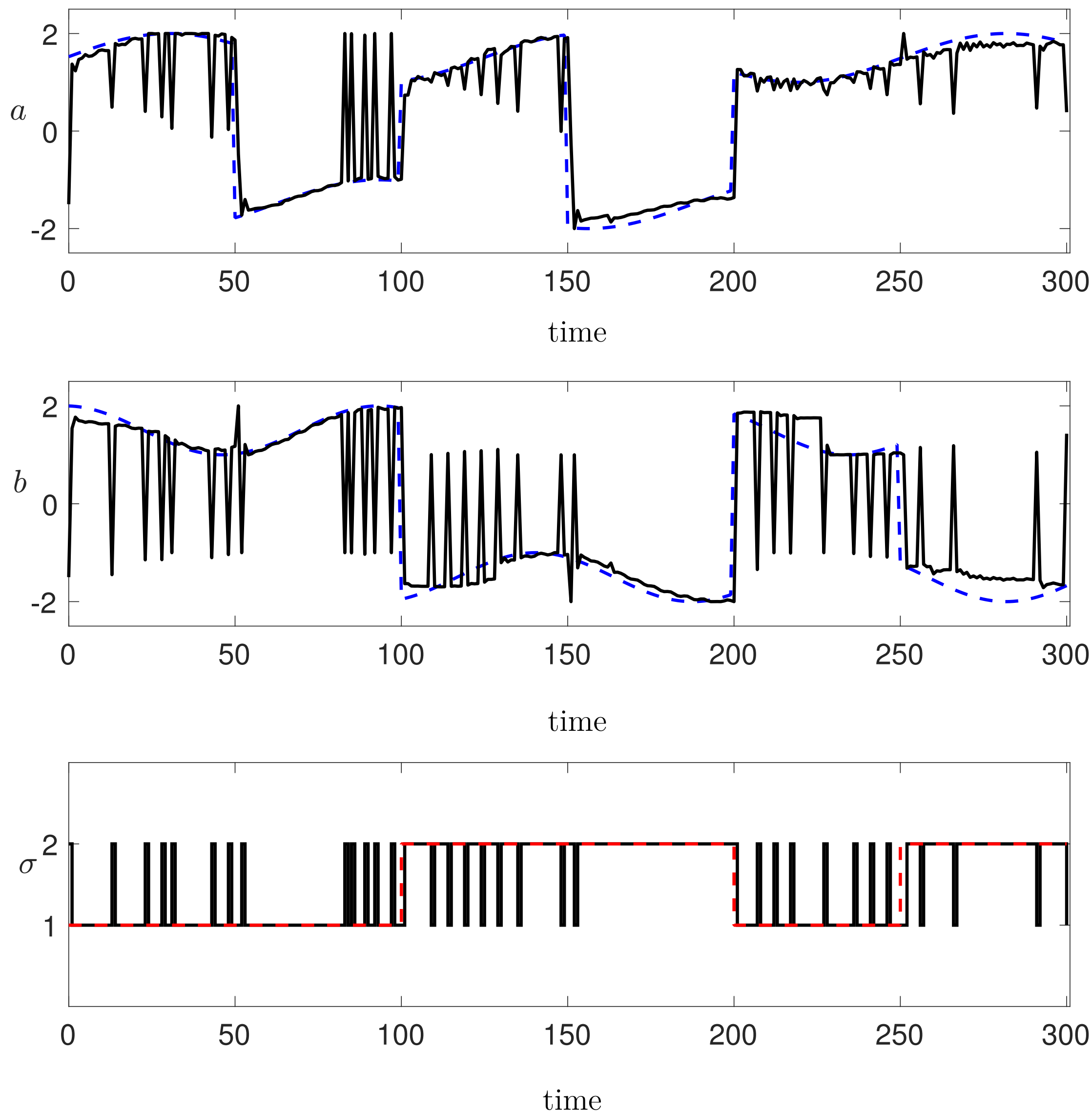


Figure 4.3: The upper two plots show the parameter estimates $\hat{\theta}_{\sigma(t)}(t)$, (solid) and actual parameters $\theta^*(t)$ (dashed); the bottom plot shows the switching signal (solid) and the correct index (dashed).

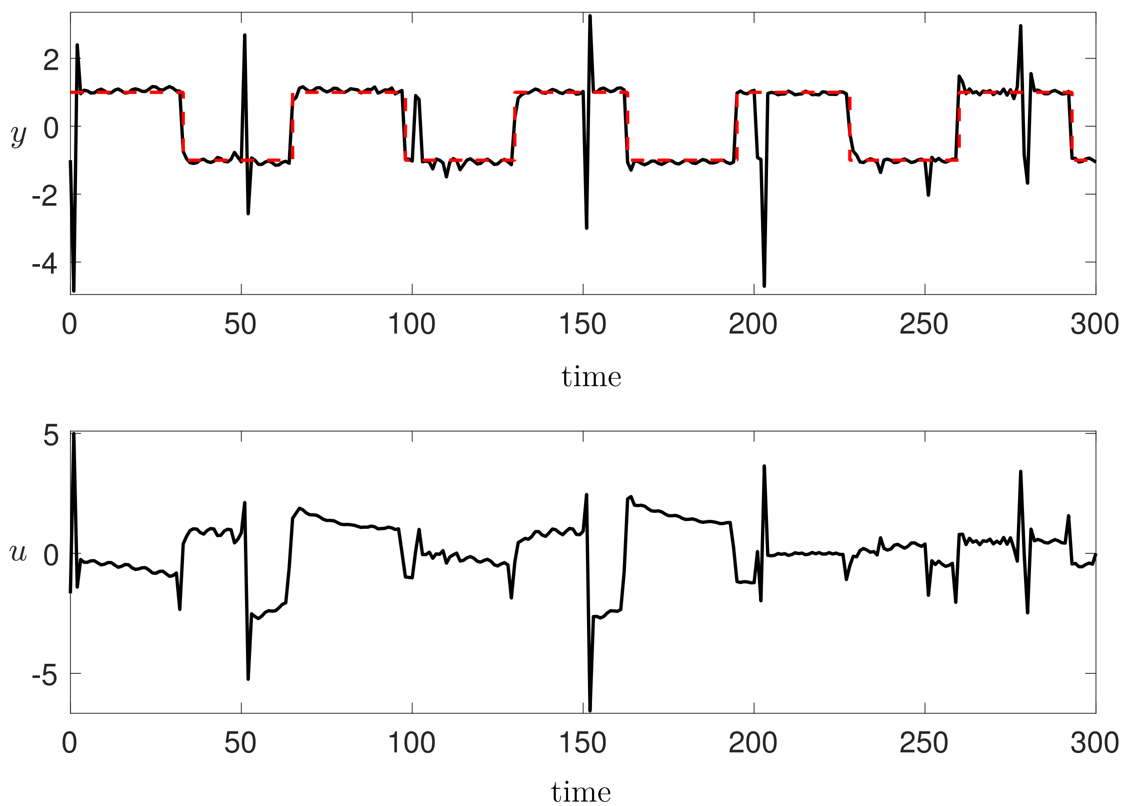


Figure 4.4: The upper plot shows both the reference (dashed) and the output (solid); the lower plot shows control input.

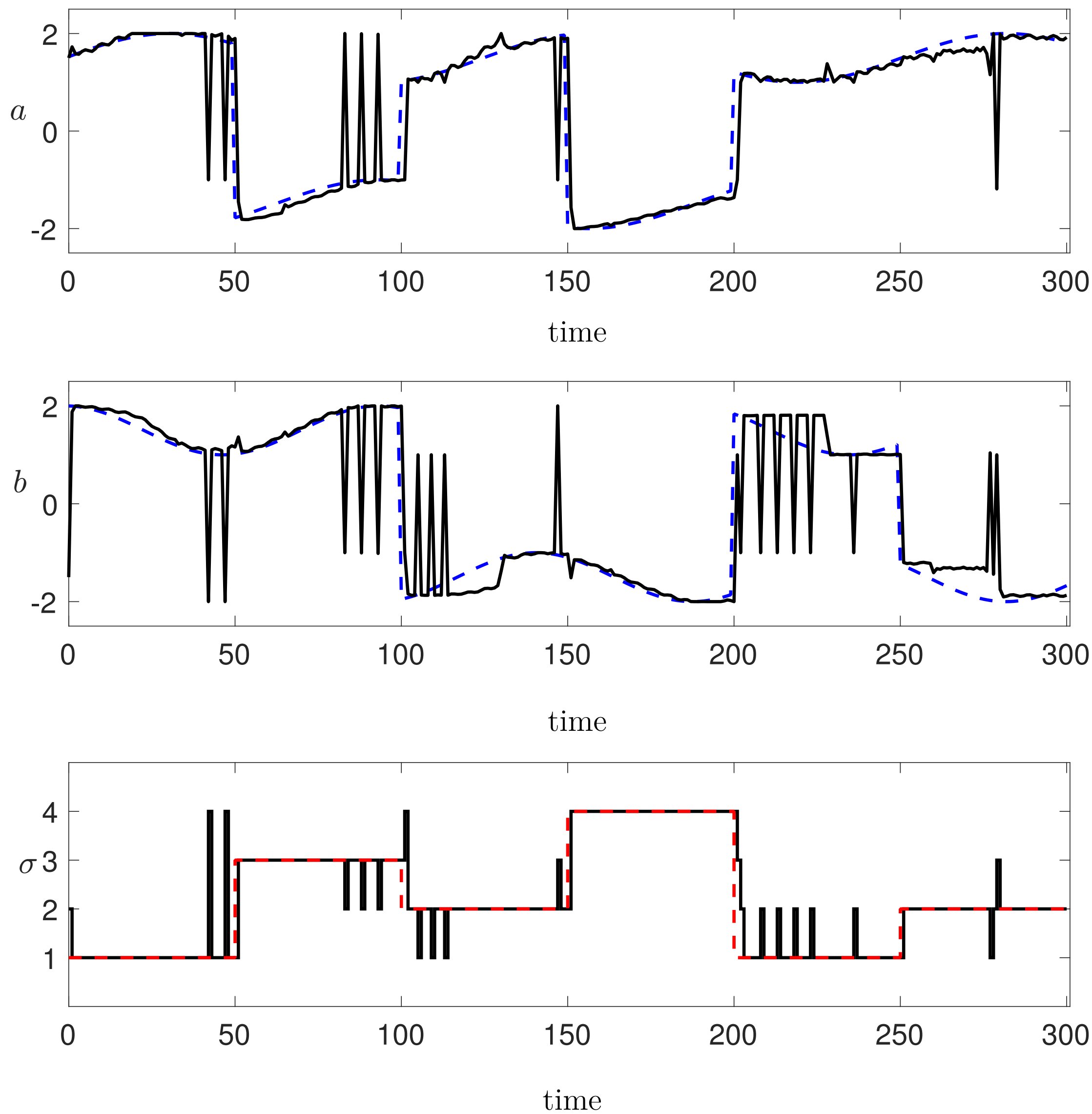


Figure 4.5: The upper two plots show the parameter estimates $\hat{\theta}_{\sigma(t)}(t)$, (solid) and actual parameters $\theta^*(t)$ (dashed); the bottom plot shows the switching signal (solid) and the correct index (dashed).

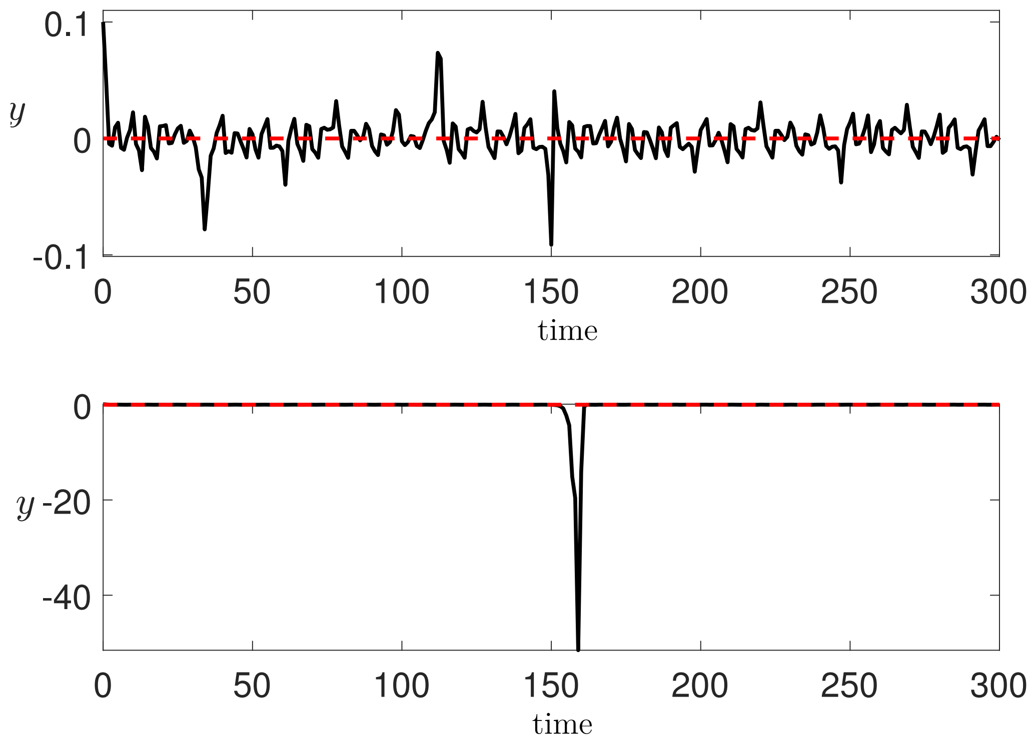


Figure 4.6: The upper plot shows the output when the original-projection-algorithm based estimator is applied; the bottom plot shows the output when the classical estimator is applied.

4.6.2 Example 2

Finally, we compare the closed-loop performance when using the original-projection-algorithm based estimator in (4.3) with the performance when using the classical estimator in (3.4) (suitably modified to have projection onto $\mathcal{S}_i$). We consider again the case of having two convex sets as defined in (4.41). We apply the adaptive controller in both cases by coupling the estimator with control law (4.4) and switching rule (4.33); in applying estimator (3.4), we choose $\alpha = \beta = 1$. Here we suppose that the plant parameters are as follows:

$$a(t) = 1.5, \qquad \text{and} \qquad b(t) = \begin{cases} 1 & t \leq 150 \\ -1 & t > 150. \end{cases}$$

We set $\hat{\theta}_1(0) = [1.5\ 1.5]^\top, \hat{\theta}_2(0) = [-1.5\ -1.5]^\top, \sigma_0 = 1, y_0 = 0.1$, the reference $y^*(\cdot)$ to be zero, and noise to be $w(t) = \frac{1}{100}\sin(5t)$. See the results in Figure 4.6. You can see that, in this example, as the control input gain changes its sign, our approach provides a smaller transient, i.e. better performance, than when using the classical estimator.

4.7 Conclusion

In this chapter, we have considered the first-order case with unknown plant parameters belonging to a closed and bounded uncertainty set. We designed a one-step-ahead adaptive controller; no assumption on convexity of the uncertainty set is imposed. A parameter estimation process is run by having multiple parallel estimators with each operating on a compact and convex set. A switching algorithm is used to determine which parameters are used in the controller. The corresponding one-step-ahead adaptive controller guarantees linear-like convolution bounds on the closed loop behavior, which confers exponential stability and a bounded noise gain. Hence, we have extended the approach of [43] which imposes a convexity requirement and uses a single estimator to the case where there is no convexity requirement and where we use multiple estimators. On the other hand, in the absence of noise, we are unable to prove that we obtain asymptotic tracking; however, if switching eventually stops then that will be the case. There are several possible approaches to alleviate this problem:

- we may be able to design such a switching algorithm which has some memory.
- if we want to track set-points, as in Morse's Supervisory Control approach, we can adopt the technique of Chapter 7.

Chapter 5

Adaptive Control of a Class of Nonlinear Systems

5.1 Introduction

The goal of this chapter is to build on the results on the adaptive control of first-order linear systems of Chapter 4 and extend them to the *nonlinear* setting. A natural nonlinear model in first-order case is

$$y(t+1) = \theta^\top \varphi\Big(y(t)\Big) + bu(t) + w(t), \tag{5.1}$$

with a known nonlinear function $\varphi : \mathbb{R} \to \mathbb{R}^q$ and unknown parameters $\theta \in \mathbb{R}^q$ and $b \in \mathbb{R}$. Indeed for this case, the desired linear-like properties is presented in our conference paper [70] for the case of a known sign of b; the details are provided there but not explicitly included here. Instead, here we will extend the approach to a class of possibly high order nonlinear systems for which (5.1) is a special case.

In general, it turns out that the adaptive control of nonlinear discrete-time systems is more challenging than in the continuous-time setting; this is, in part, due to difficulties associated with the analysis of sampled-data nonlinear systems [32]. A standard approach to achieve tracking in the context of adaptive control is to first put the input-output system into the d-step-ahead predictor form (see [21]). However, this is hard to do for nonlinear systems. To facilitate the approach, we assume that the nonlinear system is already in predictor form: indeed, we adopt the class of nonlinear n^{th}-order plants of the form

$$y(t+1) = \theta^\top \varphi\Big(y(t), y(t-1), \ldots, y(t-n+1)\Big) + bu(t) + w(t) \tag{5.2}$$

with $\varphi : \mathbb{R}^n \to \mathbb{R}^q$ a known nonlinear function and with $\theta \in \mathbb{R}^q$ and $b \in \mathbb{R}$ unknown parameters. To motivate this class of plant models, we provide the following physical example.

Example 5.1. Consider a simple one-link manipulator with an attached load; see the system pictured in Figure 5.1. The corresponding *continuous-time* dynamic system is

$$(J + 4ML^2)\ddot{\mathbf{q}} + (m + 2M)gL\sin\Big(\mathbf{q}\Big) = \tau \tag{5.3}$$

with $\mathbf{q}$ as the output angle (measured as shown in Figure 5.1) and τ as the control torque input. We have J as the moment inertia of the link about the origin, m as the mass of the link, L as the length of the link, M as the mass of the load, and g as the gravitational acceleration constant. We can rewrite the above system as

$$\frac{d^2\mathbf{q}(t)}{dt^2} = -\frac{(m+2M)gL}{(J+4ML^2)}\sin\Big(\mathbf{q}(t)\Big) + \frac{1}{(J+4ML^2)}\tau(t).$$

Instead of applying standard discretization using the Euler method, we elect to approximate the 2^{nd} derivative as follows:

$$\frac{d^2\mathbf{q}(t)}{dt^2} \simeq \frac{\mathbf{q}(t+h) - 2\mathbf{q}(t) + \mathbf{q}(t-h)}{h^2},$$

with h as the sampling period. If we define $y(t) := \mathbf{q}(th)$ and $u(t) := \tau(th)$, we end up with the following discrete-time second-order system which approximates (5.3):

$$y(t+1) = 2y(t) - \frac{(m+2M)gLh^2}{(J+4ML^2)}\sin\Big(y(t)\Big) - y(t-1) + \frac{h^2}{(J+4ML^2)}u(t), \tag{5.4}$$

matching the form in (5.2) with

$$\theta = \begin{bmatrix} 2 \\ -\frac{(m+2M)gLh^2}{(J+4ML^2)} \\ -1 \end{bmatrix}, \quad \varphi\Big(y(t), y(t-1)\Big) = \begin{bmatrix} y(t) \\ \sin\Big(y(t)\Big) \\ y(t-1) \end{bmatrix}, \quad b = \frac{h^2}{(J+4ML^2)}.$$

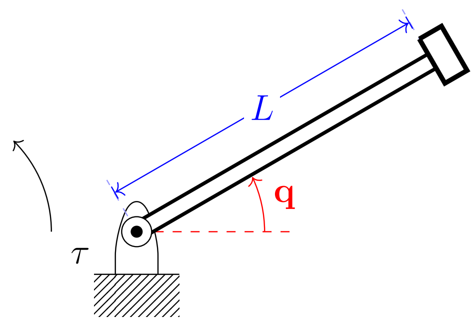


Figure 5.1: One-link manipulator with load.

Remark 5.1. *Adaptive control of systems of the form in (5.2) is analyzed in Guo's paper [22]; desirable linear-like closed-loop properties are not obtained.*

Remark 5.2. *In the literature, a lot of work has been done on adaptive control of nonlinear discrete-time systems, e.g. see [32], [12], [73], [85], [86], [84], [19], [18], [83], [89], [13]; almost exclusively they consider systems with relative degree n, i.e. with no zero dynamics, and in general, exponential stability and a bounded gain on the noise are not proven. In contrast, here we consider another class of nonlinear discrete-time systems, and we are able to prove all the desirable linear-like properties proven in the previous chapter.*

In the following we consider a class of nonlinear plants (a version of (5.2) which allows for measurement noise) and show, under suitable assumptions, how to carry out adaptive control so that we obtain not only exponential stability and a bounded gain on the noise, but also a convolution bound on the effect of the exogenous inputs. First, we consider plants with a known sign of the control/input gain; the second part of the chapter will consider the case when that sign is unknown.

5.2 The Setup

Here we consider a class of n^{th}-order discrete-time nonlinear plants described by

$$x(t+1) = \sum_{i=1}^{p-1} \theta_i \varphi_i\Big(x(t), x(t-1), \cdots, x(t-n+1)\Big) + \theta_p u(t) + w(t) \tag{5.5a}$$

$$y(t) = x(t) + v(t). \tag{5.5b}$$

In (5.5) we have $y(t) \in \mathbb{R}$ as the output, $u(t) \in \mathbb{R}$ as the control input, $x(t) \in \mathbb{R}$ as an internal variable, $w(t) \in \mathbb{R}$ as the process noise, and $v(t) \in \mathbb{R}$ as the measurement noise. At time t, one needs all of $x(t), x(t-1), \ldots, x(t-n+1)$ to form $x(t+1)$, so the natural systems theoretic notion of a state is

$$\mathcal{X}(t) := \begin{bmatrix} x(t) \\ x(t-1) \\ \vdots \\ x(t-n+1) \end{bmatrix};$$

here we will adopt the notation

$$\varphi_i\Big(\mathcal{X}(t)\Big) := \varphi_i\Big(x(t), x(t-1), \cdots, x(t-n+1)\Big).$$

For each $i = 1, 2, \ldots, p-1$, $\varphi_i : \mathbb{R}^n \to \mathbb{R}$ is a *known* nonlinear function. With the function $\varphi : \mathbb{R}^n \to \mathbb{R}^{p-1}$ defined by

$$\varphi\Big(\mathcal{X}\Big) := \begin{bmatrix} \varphi_1\Big(\mathcal{X}\Big) \\ \varphi_2\Big(\mathcal{X}\Big) \\ \vdots \\ \varphi_{p-1}\Big(\mathcal{X}\Big) \end{bmatrix}$$

and the unknown parameter vector defined by

$$\theta^* := \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_p \end{bmatrix},$$

we can rewrite the plant in a more compact form as

$$x(t+1) = \theta^{*\top} \begin{bmatrix} \varphi\Big(\mathcal{X}(t)\Big) \\ u(t) \end{bmatrix} + w(t) \tag{5.6a}$$

$$y(t) = x(t) + v(t). \tag{5.6b}$$

It is assumed that θ^* lies in a known set $\mathcal{S}^* \subset \mathbb{R}^p$; we impose the following assumption on the set.

Assumption 5.1. The set $\mathcal{S}^*$ is closed and bounded (compact), and for each $\theta^* \in \mathcal{S}^*$, the p^{th} element θ_p is non-zero.

The boundedness requirement is reasonable in practical situations; it is used here to prove uniform bounds and decay rates on the closed-loop behavior.

We also impose some conditions on the nonlinear function φ.

Assumption 5.2. The function φ satisfies the following conditions:

1. $\varphi(\mathbf{0}_n) = \mathbf{0}_{p-1}$, and
2. φ is globally Lipschitz continuous: there exists a constant $K > 0$ such that

$$\left\|\varphi\big(\mathcal{X}\big) - \varphi\big(\tilde{\mathcal{X}}\big)\right\| \leq K\|\mathcal{X} - \tilde{\mathcal{X}}\|, \quad \text{for all } \mathcal{X}, \tilde{\mathcal{X}} \in \mathbb{R}^n;$$

   we denote the smallest such constant by c_φ.

This assumption restricts the family of nonlinear functions to those that vanish at zero and are Lipschitz continuous. These restrictions are common in the literature.

We have an exogenous reference signal $y^*(\cdot)$ and the objective is to track it asymptotically while stabilizing the closed-loop system; we assume that we know y^* one step ahead in time, i.e. we know $y^*(t+1)$ at time t. We are interested in analyzing the corresponding one-step-ahead control law when the plant parameters are unknown. We first present the case when the sign of θ_p is known; afterwards, we will present the case of a switching control law when the sign of θ_p is not known.

Before proceeding, we present some definitions and observations that will be used in the rest of the chapter. For ease of notation define

$$\mathcal{Y}(t) := \begin{bmatrix} y(t) \\ y(t-1) \\ \vdots \\ y(t-n+1) \end{bmatrix} \tag{5.7}$$

and define the following data vector:

$$\phi(t) := \begin{bmatrix} \varphi\big(\mathcal{Y}(t)\big) \\ u(t) \end{bmatrix}. \tag{5.8}$$

We see from the definition of $\phi(\cdot)$ and by Assumption 5.2 that

$$\|\phi(t)\| \leq \left\|\varphi\big(\mathcal{Y}(t)\big)\right\| + |u(t)| \leq c_\varphi \|\mathcal{Y}(t)\| + |u(t)|. \tag{5.9}$$

With $t \geq t_0$, observe from (5.6) that

$$\begin{aligned} y(t+1) &= \theta^{*\top} \begin{bmatrix} \varphi\big(\mathcal{X}(t)\big) \\ u(t) \end{bmatrix} + w(t) + v(t+1) \\ &= \theta^{*\top}\phi(t) + w(t) + v(t+1) + \underbrace{\theta^{*\top} \begin{bmatrix} \varphi\big(\mathcal{X}(t)\big) - \varphi\big(\mathcal{Y}(t)\big) \\ 0 \end{bmatrix}}_{=:\Delta(t)}. \end{aligned} \tag{5.10}$$

Note that

$$|\Delta(t)| \leq \|\theta^*\| \left\|\varphi\big(\mathcal{X}(t)\big) - \varphi\big(\mathcal{Y}(t)\big)\right\|;$$

so by Assumption 5.2, from the definition of $\mathcal{X}(\cdot)$, $\mathcal{Y}(\cdot)$ and the fact that $y(j) = x(j) + v(j)$, we obtain

$$\begin{aligned} |\Delta(t)| &\leq \|\theta^*\| c_\varphi \left\| \begin{bmatrix} v(t) \\ v(t-1) \\ \vdots \\ v(t-n+1) \end{bmatrix} \right\| \\ &\leq \|\mathcal{S}^*\| c_\varphi \sum_{j=1}^{n} |v(t-j+1)|. \end{aligned} \tag{5.11}$$

Also, define a combined noise signal that will be used throughout the chapter:

$$\begin{aligned} \bar{w}(t) &:= \sqrt{w(t)^2 + \sum_{j=0}^{n} v(t-j+1)^2} \\ \Rightarrow \bar{w}(t)^2 &= w(t)^2 + \sum_{j=0}^{n} v(t-j+1)^2; \end{aligned} \tag{5.12}$$

note that from the above we can easily confirm that

$$\begin{aligned} |w(t)| &\leq |\bar{w}(t)| \\ |v(t)| &\leq |\bar{w}(t)| \\ |v(t-1)| &\leq |\bar{w}(t)| \end{aligned}$$

$$\vdots$$
$$|v(t-n+1)| \le |\bar{w}(t)|,$$

which means that

$$|w(t)| + \sum_{j=0}^{n} |v(t-j+1)| \le (n+2)|\bar{w}(t)|. \tag{5.13}$$

5.3 The Case of the Sign of θ_p being Known

In this section, we impose an extra assumption on the set of admissible parameters.

Assumption 5.3. For each $\theta^* \in \mathcal{S}^*$, the sign of the p^{th} element θ_p is always the same.

The knowledge of the sign of the input gain θ_p is a common assumption in adaptive control, e.g. see [21]. To proceed, we use a parameter estimator together with a one-step-ahead adaptive control law. To facilitate estimation, it is convenient for the set of admissible parameters to be convex, so at this point let $\mathcal{S} \subset \mathbb{R}^p$ be any **convex and compact set** containing $\mathcal{S}^*$ for which the p^{th} element is never zero—the convex hull of $\mathcal{S}^*$ would do.

5.3.1 Parameter Estimation

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

$$e(t+1) := y(t+1) - \hat{\theta}(t)^\top \phi(t); \tag{5.14}$$

this is a measure of the error in $\hat{\theta}(t)$, since it is zero if $\hat{\theta}(t) = \theta^*$ and $w = v = 0$. Here we will be using a version of the original-projection-algorithm based algorithm presented in (3.9) of Chapter 3. It turns out for the approach to work, the denominator of the estimator update needs to be set carefully. Notice that the data vector $\phi(t)$ does not directly have access to plant outputs $\mathcal{Y}(\cdot)$; we only have $\varphi\Big(\mathcal{Y}(\cdot)\Big)$ instead. To this end, first define

$$\tilde{\phi}(t) := \begin{bmatrix} \phi(t) \\ \mathcal{Y}(t) \end{bmatrix}.$$

With $\delta \in (0, \infty]$ and with $\|\tilde{\phi}(t)\|$ replacing $\|\phi(t)\|$, we define $\rho : \mathbb{Z} \mapsto \{0, 1\}$ by

$$\rho(t) := \begin{cases} 1 & \text{if } |e(t+1)| < (2\|\mathcal{S}\| + \delta)\|\tilde{\phi}(t)\| \\ 0 & \text{otherwise;} \end{cases} \tag{5.15}$$

with initial condition $\hat{\theta}(t_0) \in \mathcal{S}$, for $t \geq t_0$ estimator updates are computed by

$$\check{\theta}(t+1) = \hat{\theta}(t) + \rho(t)\frac{\phi(t)}{\|\tilde{\phi}(t)\|^2}e(t+1) \tag{5.16a}$$

$$\hat{\theta}(t+1) = \operatorname*{Proj}_{\mathcal{S}}\left\{\check{\theta}(t+1)\right\}. \tag{5.16b}$$

Remark 5.3. *Observe that we can rewrite the plant model (5.6) as*

$$y(t+1) = \begin{bmatrix} \theta^* \\ \mathbf{0}_n \end{bmatrix}^\top \underbrace{\begin{bmatrix} \phi(t) \\ \mathcal{Y}(t) \end{bmatrix}}_{\tilde{\phi}(t)} + w(t) + v(t+1) + \Delta(t), \tag{5.17}$$

so the use of $\tilde{\phi}(t)$ in the estimator is not as odd as it might first seem. Indeed, if we were to apply the standard approach of Chapter 3 to (5.17), we would attempt to estimate $\begin{bmatrix} \theta^* \\ \mathbf{0}_n \end{bmatrix}$ which we would then project onto $\mathcal{S} \times \{\mathbf{0}_n\}$. The method of (5.16) is a simplified version, where we do not attempt to estimate the bottom part of $\begin{bmatrix} \theta^* \\ \mathbf{0}_n \end{bmatrix}$ since we already know it is zero.

To proceed, define the parameter estimation error

$$\tilde{\theta}(t) := \hat{\theta}(t) - \theta^*.$$

In the following result we present a property of the estimator; this Proposition follows directly from Proposition 3.1 of Chapter 3, where we set $\phi_m = \tilde{\phi}$.

Proposition 5.1. *There exists a constant $\bar{c} > 0$ such that for every $t_0 \in \mathbb{Z}$, $\mathcal{X}(t_0) \in \mathbb{R}^n$, $\hat{\theta}(t_0) \in \mathcal{S}$, $\theta^* \in \mathcal{S}^*$, and $v, w \in \boldsymbol{\ell}_\infty$, when the estimator (5.16) is applied to the plant (5.5), the following holds:*

$$\left\|\tilde{\theta}(t)\right\|^2 \leq \left\|\tilde{\theta}(\tau)\right\|^2 + \sum_{j=\tau}^{t-1} \rho(j) \left[-\frac{1}{2} \frac{e(j+1)^2}{\|\tilde{\phi}(j)\|^2} + \bar{c} \frac{\bar{w}(j)^2}{\|\tilde{\phi}(j)\|^2} \right], \quad t > \tau \geq t_0. \tag{5.18}$$

Remark 5.4. *We will see in the proof that we can set*

$$\bar{c} = 6 \max\left\{1, c_\varphi^2 \|\mathcal{S}^*\|^2 n\right\}.$$

Proof of Proposition 5.1. Let $t_0 \in \mathbb{Z}$, $t > \tau \geq t_0$, $\mathcal{X}(t_0) \in \mathbb{R}^n$, $\hat{\theta}(t_0) \in \mathcal{S}$, $\theta^* \in \mathcal{S}^*$, and $v, w \in \boldsymbol{\ell}_\infty$ be arbitrary.

Applying Proposition 3.1 from Chapter 3 to (5.10) which is an equivalent version of the plant equation (5.5), we obtain

$$\left\|\tilde{\theta}(t)\right\|^2 \leq \left\|\tilde{\theta}(\tau)\right\|^2 + \sum_{j=\tau}^{t-1} \rho(j) \left[-\frac{1}{2} \frac{e(j+1)^2}{\|\tilde{\phi}(j)\|^2} + \frac{2\left[w(j) + v(j+1) + \Delta(j)\right]^2}{\|\tilde{\phi}(j)\|^2} \right],$$
$$t > \tau \geq t_0. \tag{5.19}$$

Observe that

$$\left[w(j) + v(j+1) + \Delta(j)\right]^2 \leq 3\left[w(j)^2 + v(j+1)^2 + \Delta(j)^2\right]; \tag{5.20}$$

using the bound on $\Delta(\cdot)$ in (5.11), we can see that

$$\Delta(j)^2 \leq \|\mathcal{S}^*\|^2 c_\varphi^2 n \sum_{j=1}^{n} |v(t-j+1)|^2.$$

Substituting the above into (5.20) and using the notation defined in (5.12), we see that

$$\left[w(j) + v(j+1) + \Delta(j)\right]^2 \leq 3 \max\left\{1, c_\varphi^2 \|\mathcal{S}^*\|^2 n\right\} \bar{w}(j)^2.$$

Substitute this into (5.19) and define constant $\bar{c} := 6 \max\left\{1, c_\varphi^2 \|\mathcal{S}^*\|^2 n\right\}$ to conclude the proof. ∎

5.3.2 The Control Law

We partition $\hat{\theta}(t)$ in a natural way as

$$\hat{\theta}(t) =: \begin{bmatrix} \hat{\theta}_1(t) \\ \hat{\theta}_2(t) \\ \vdots \\ \hat{\theta}_p(t) \end{bmatrix} = \begin{bmatrix} \hat{\bar{\theta}}(t) \\ \hat{\theta}_p(t) \end{bmatrix}, \qquad \text{with} \quad \hat{\bar{\theta}}(t) := \begin{bmatrix} \hat{\theta}_1(t) \\ \hat{\theta}_2(t) \\ \vdots \\ \hat{\theta}_{p-1}(t) \end{bmatrix}.$$

The natural choice for the one-step-ahead adaptive control law is given by

$$u(t) = \frac{1}{\hat{\theta}_p(t)} \left[y^*(t+1) - \hat{\bar{\theta}}(t)^\top \varphi\big(\mathcal{Y}(t)\big) \right], \quad t \geq t_0; \tag{5.21}$$

note that if θ^* were known (i.e. $\hat{\theta}(t) = \theta^*$) and no noise were entering the system, this control law will ensure that $y(t+1) = y^*(t+1)$.

Let the tracking error be defined by

$$\varepsilon(t) := y(t) - y^*(t);$$

from (5.21) we see that

$$y^*(t+1) = \hat{\theta}(t)^\top \phi(t);$$

combining this with (5.14) yields

$$e(t+1) = \varepsilon(t+1), \qquad t \geq t_0. \tag{5.22}$$

Before proceeding, define

$$c_\theta := \max \left\{ \frac{1}{|\theta_p|} : \theta^* \in \mathcal{S} \right\}.$$

5.3.3 The Main Result

We now prove that the proposed adaptive controller has very desirable properties.

Theorem 5.1. *Suppose that the adaptive controller (5.16) and (5.21) is applied to the nonlinear plant (5.5). Then for every $\lambda \in (0,1)$ and $\delta \in (0,\infty]$, there exists a constant $\gamma > 0$ so that, for every $t_0 \in \mathbb{Z}$, $\mathcal{X}(t_0) \in \mathbb{R}^n$, $\hat{\theta}(t_0) \in \mathcal{S}$, $\theta^* \in \mathcal{S}^*$, and $v, w, y^* \in \boldsymbol{\ell}_\infty$, the following holds:*

$$\left\| \begin{bmatrix} \mathcal{Y}(t) \\ u(t) \end{bmatrix} \right\| \leq \gamma \lambda^{t-\tau} \left\| \mathcal{Y}(\tau) \right\| + \sum_{j=\tau}^{t-1} \gamma \lambda^{t-\tau-1} \left(|y^*(j+1)| + |\bar{w}(j)| \right) + \gamma |y^*(t+1)|, \qquad t > \tau \geq t_0. \tag{5.23}$$

Remark 5.5. *The above result shows that the closed-loop system experiences linear-like behavior: there is a uniform exponential decay bound on the effect of the initial condition, and there is a convolution bound on the effect of the exogenous signals. This implies that the system has a bounded gain (from w, v and y^* to y) in every p-norm: in particular, for $p = \infty$, it follows from (5.23) that there exists $c > 0$ such that*

$$\begin{aligned} \left\| \begin{bmatrix} \mathcal{Y}(t) \\ u(t) \end{bmatrix} \right\| &\leq \gamma \lambda^{t-t_0} \| \mathcal{Y}(t_0) \| + \tfrac{\gamma}{1-\lambda} \sup_{j \geq t_0} \left(|y^*(j+1)| + |\bar{w}(j)| \right), \\ &\leq \frac{c\gamma}{1-\lambda} \left(\lambda^{t-t_0} \| \mathcal{X}(t_0) \| + \|y^*\|_\infty + \|w\|_\infty + \|v\|_\infty \right), \qquad t \geq t_0. \end{aligned}$$

Hence, if $w, v, y^* \in \boldsymbol{\ell}_\infty$, then $y, x, u \in \boldsymbol{\ell}_\infty$, so ε, e lie in $\boldsymbol{\ell}_\infty$ as well; all signals in the closed-loop system are uniformly bounded.

Before presenting the proof of Theorem 5.1, we provide a crude bound on the closed-loop behavior.

Proposition 5.2. *Suppose that the adaptive controller (5.16) and (5.21) is applied to the plant (5.5). Then for every $p \geq 0$, there exists a constant $\gamma_1 > 0$ so that, for every $t_0 \in \mathbb{Z}$, $t \geq t_0$, $\mathcal{X}(t_0) \in \mathbb{R}^n$, $\hat{\theta}(t_0) \in \mathcal{S}$, $\theta^* \in \mathcal{S}^*$, and $v, w, y^* \in \boldsymbol{\ell}_\infty$, the following holds:*

$$\| \mathcal{Y}(t+p) \| \leq \gamma_1 \| \mathcal{Y}(t) \| + \gamma_1 \sum_{j=0}^{p-1} \left(|\bar{w}(t+j)| + |y^*(t+j+1)| \right). \tag{5.24}$$

Proof. See Appendix B. ∎

Proof of Theorem 5.1:

Fix $\lambda \in (0,1)$ and $\delta \in (0,\infty]$. Let $t_0 \in \mathbb{Z}$, $\mathcal{X}(t_0) \in \mathbb{R}^n$, $\hat{\theta}(t_0) \in \mathcal{S}$, $\theta^* \in \mathcal{S}^*$, and $v, w, y^* \in \boldsymbol{\ell}_\infty$ be arbitrary.

Step 1: Construct a useful difference equation.

From the definition of the tracking error we have

$$y(t+1) = \varepsilon(t+1) + y^*(t+1);$$

from (5.22), we know that the tracking error equals the prediction error, so we conclude that

$$y(t+1) = e(t+1) + y^*(t+1). \tag{5.25}$$

Now define this important quantity which appears in the estimator:

$$\alpha(t) := \rho(t)\frac{e(t+1)}{\|\tilde{\phi}(t)\|^2}\tilde{\phi}(t)^\top; \tag{5.26}$$

it is clear that

$$\alpha(t)\tilde{\phi}(t) = \rho(t)e(t+1). \tag{5.27}$$

Now observe that

$$e(t+1) = \rho(t)e(t+1) + \underbrace{[1-\rho(t)]e(t+1)}_{=:\eta(t)};$$

if we use this together with (5.27) and substitute the result into (5.25), then we obtain

$$y(t+1) = \alpha(t)\tilde{\phi}(t) + \eta(t) + y^*(t+1). \tag{5.28}$$

Step 2: Analyze the difference equation (5.28).

By the definition of $\mathcal{Y}$, it is easy to see that

$$\|\mathcal{Y}(t+n)\| = \left\| \begin{bmatrix} y(t+n) \\ y(t+n-1) \\ \vdots \\ y(t+1) \end{bmatrix} \right\| \le \sum_{j=0}^{n-1} |y(t+j+1)|;$$

so using (5.28), we obtain

$$\|\mathcal{Y}(t+n)\| \le \sum_{j=0}^{n-1} \left(\|\alpha(t+j)\| \|\tilde{\phi}(t+j)\| + |\eta(t+j)| + |y^*(t+j+1)| \right). \tag{5.29}$$

From (5.21) and Assumptions 5.2, we obtain

$$\begin{aligned} |u(t)| &\le c_\theta \left(\left\| \hat{\bar{\theta}}(t) \right\| \left\| \varphi\big(\mathcal{Y}(t)\big) \right\| + |y^*(t+1)| \right) \\ &\le c_\theta \|\mathcal{S}\| \|\varphi\big(\mathcal{Y}(t)\big)\| + c_\theta |y^*(t+1)| \\ &\le c_\theta c_\varphi \|\mathcal{S}\| \|\mathcal{Y}(t)\| + c_\theta |y^*(t+1)|; \end{aligned} \tag{5.30}$$

we now combine this with (5.9) to obtain a bound on $\phi(t)$:

$$\|\phi(t)\| \le \underbrace{c_\varphi(1 + c_\theta \|\mathcal{S}\|)}_{=: c_1} \|\mathcal{Y}(t)\| + c_\theta |y^*(t+1)|. \tag{5.31}$$

If we incorporate this with the definition of $\tilde{\phi}$ and define $c_2 := 1 + c_1$, then we obtain

$$\begin{aligned} \|\tilde{\phi}(t)\| &\le \|\phi(t)\| + \|\mathcal{Y}(t)\| \\ &\le c_2 \|\mathcal{Y}(t)\| + c_\theta |y^*(t+1)|. \end{aligned} \tag{5.32}$$

Now if we use this bound on $\tilde{\phi}$ in (5.29), then we obtain

$$\begin{aligned} \|\mathcal{Y}(t+n)\| \le \sum_{j=0}^{n-1} \Big(& c_2 \|\alpha(t+j)\| \|\mathcal{Y}(t+j)\| + |\eta(t+j)| + \\ & (1 + c_\theta \|\alpha(t+j)\|) |y^*(t+j+1)| \Big). \end{aligned} \tag{5.33}$$

Now let us obtain a bound on $|\eta(t+j)|$ in terms of $|\bar{w}(t+j)|$.

Claim 5.1. *There exists a constant c_3 such that*

$$|\eta(t)| \le c_3 |\bar{w}(t)|, \qquad t \ge t_0. \tag{5.34}$$

Proof of Claim 5.1. If $\rho(t) = 1$, then $\eta(t) = 0$, so (5.34) clearly holds. Now suppose that $\rho(t) = 0$. Then $\eta(t) = e(t+1)$, and from the definition of ρ given in (5.15) we see that

$$|e(t+1)| \ge (2\|\mathcal{S}\| + \delta) \|\tilde{\phi}(t)\|. \tag{5.35}$$

If we combine the prediction error definition in (5.14), the formula for y given in equation (5.10) and the bound on $\Delta(\cdot)$ given in equation (5.11), then we see that

$$
\begin{aligned}
|e(t+1)| &\leq \|\mathcal{S}^*\|\|\phi(t)\| + |w(t)| + |v(t+1)| + |\Delta(t)| + \|\hat{\theta}(t)\|\|\phi(t)\| \\
&\leq 2\|\mathcal{S}\|\|\phi(t)\| + |w(t)| + |v(t+1)| + c_\varphi\|\mathcal{S}^*\| \left(\sum_{j=1}^{n} |v(t-j+1)| \right) \\
&\leq 2\|\mathcal{S}\|\|\phi(t)\| + (1 + c_\varphi\|\mathcal{S}^*\|)(n+2)|\bar{w}(t)| \\
&\leq 2\|\mathcal{S}\|\|\tilde{\phi}(t)\| + (1 + c_\varphi\|\mathcal{S}^*\|)(n+2)|\bar{w}(t)|, \quad t \geq t_0.
\end{aligned} \tag{5.36}
$$

Combining this inequality with (5.35) yields

$$
\begin{aligned}
&2\|\mathcal{S}\|\|\tilde{\phi}(t)\| + (1 + c_\varphi\|\mathcal{S}^*\|)(n+2)|\bar{w}(t)| \geq (2\|\mathcal{S}\| + \delta)\|\tilde{\phi}(t)\| \\
&\Rightarrow \|\tilde{\phi}(t)\| \leq \frac{(1 + c_\varphi\|\mathcal{S}^*\|)(n+2)}{\delta}|\bar{w}(t)|;
\end{aligned}
$$

finally, if we combine the above with (5.36), then we obtain

$$
\Rightarrow |\eta(t)| = |e(t+1)| \leq \underbrace{(1 + c_\varphi\|\mathcal{S}^*\|)(n+2)\left(\frac{2\|\mathcal{S}\|}{\delta} + 1 \right)}_{=:c_3} |\bar{w}(t)|,
$$

so we conclude that (5.34) holds. □

Applying (5.34) into (5.33), we obtain

$$
\begin{aligned}
\|\mathcal{Y}(t+n)\| \leq \sum_{j=0}^{n-1} \Big[& c_2\|\alpha(t+j)\|\|\mathcal{Y}(t+j)\| + \\
& (1 + c_\theta\|\alpha(t+j)\|)\, |y^*(t+j+1)| + c_3|\bar{w}(t+j)| \Big].
\end{aligned} \tag{5.37}
$$

Now let us define

$$
\tilde{\alpha}_n(t) := \max_{j=0,1,\ldots,n-1} \|\alpha(t+j)\|, \qquad t \geq t_0. \tag{5.38}
$$

This means that we can rewrite (5.37) as

$$
\|\mathcal{Y}(t+n)\| \leq c_2\tilde{\alpha}_n(t) \sum_{j=0}^{n-1} \|\mathcal{Y}(t+j)\| +
$$

$$\sum_{j=0}^{n-1}\left[\left(1+c_\theta\tilde{\alpha}_n(t)\right)|y^*(t+j+1)|+c_3|\bar{w}(t+j)|\right]. \tag{5.39}$$

It follows from Proposition 5.2 (applied for $p=1,2,\ldots,n-1$) that there exists a constant c_4 so that the following holds:

$$\sum_{j=0}^{n-1}\|\mathcal{Y}(t+j)\|\le c_4\|\mathcal{Y}(t)\|+c_4\sum_{j=0}^{n-2}(|y^*(t+j+1)|+|\bar{w}(t+j)|);$$

after substituting this into (5.39) and simplifying it follows that there exists a constant c_5 such that

$$\begin{aligned}\|\mathcal{Y}(t+n)\| &\le c_2\tilde{\alpha}_n(t)\left(c_4\|\mathcal{Y}(t)\|+c_4\sum_{j=0}^{n-2}(|y^*(t+j+1)|+|\bar{w}(t+j)|)\right)+\\ &\qquad \sum_{j=0}^{n-1}\Big[(1+c_\theta\tilde{\alpha}_n(t))\,|y^*(t+j+1)|+c_3|\bar{w}(t+j)|\Big]\\ &\le c_5\tilde{\alpha}_n(t)\|\mathcal{Y}(t)\|+c_5\left(1+\tilde{\alpha}_n(t)\right)\sum_{j=0}^{n-1}\left(|y^*(t+j+1)|+|\bar{w}(t+j)|\right). \end{aligned}\tag{5.40}$$

This difference inequality governs the closed-loop system's behavior.

Step 3: Analyze the first-order difference inequality (5.40).

Now we analyze the closed-loop behavior on the whole timeline. First, define

$$\lambda_1=\frac{\lambda^n}{\max\{1,c_5\}}\in(0,1)$$

and

$$\nu:=\frac{\lambda_1^2}{4n^2\bar{c}}\quad\in(0,1). \tag{5.41}$$

To proceed, let $\tau\ge t_0$ be arbitrary. We now partition the timeline into two parts: one in which $\bar{w}(\cdot)$ is small versus $\tilde{\phi}(\cdot)$ and one where it is not. With ν defined above, we define

$$S_{good}=\left\{j\ge\tau:\tilde{\phi}(j)\ne 0\text{ and }\frac{|\bar{w}(j)|^2}{\|\tilde{\phi}(j)\|^2}<\nu\right\},$$

$$S_{bad} = \left\{ j \geq \tau : \tilde{\phi}(j) = 0 \text{ or } \frac{|\bar{w}(j)|^2}{\|\tilde{\phi}(j)\|^2} \geq \nu \right\};$$

clearly $\{j \in \mathbb{Z} : \ j \geq \tau\} = S_{good} \cup S_{bad}$. We can clearly define a (possibly infinite) sequence of intervals of the form $[k_l, k_{l+1})$ which satisfy:
(i) $k_0 = \tau$ serves as the initial instant of the first interval;
(ii) $[k_l, k_{l+1})$ either belongs to S_{good} or S_{bad}; and
(iii) if $k_{l+1} \neq \infty$ and $[k_l, k_{l+1})$ belongs to S_{good} then $[k_{l+1}, k_{l+2})$ belongs to S_{bad}, and vice versa.

Now we analyze the behavior during each interval.

Case 1: $[k_l, k_{l+1})$ lies in S_{bad}.

Let $j \in [k_l, k_{l+1})$ be arbitrary. In this case, $\frac{\bar{w}(j)^2}{\|\tilde{\phi}(j)\|^2} \geq \nu$ or $\|\tilde{\phi}(j)\| = 0$; in either case by the definition of $\tilde{\phi}(\cdot)$ we get

$$\|\phi(j)\|^2 + \|\mathcal{Y}(j)\|^2 = \|\tilde{\phi}(j)\|^2 \leq \frac{1}{\nu}|\bar{w}(j)|^2,$$

which implies that

$$\|\phi(j)\| \leq \frac{1}{\sqrt{\nu}}|\bar{w}(j)|, \qquad j \in [k_l, k_{l+1}), \tag{5.42a}$$

and

$$\|\mathcal{Y}(j)\| \leq \frac{1}{\sqrt{\nu}}|\bar{w}(j)|, \qquad j \in [k_l, k_{l+1}). \tag{5.42b}$$

We would like to obtain a bound similar to that in (5.42b) for $\|\mathcal{Y}(k_{l+1})\|$. To proceed, first observe that, based on the definition of $\mathcal{Y}(t)$, we have that

$$\|\mathcal{Y}(t+1)\| \leq \|\mathcal{Y}(t)\| + |y(t+1)|. \tag{5.43}$$

Second of all, from (5.10) and (5.11), we obtain

$$|y(j+1)| \leq \|\mathcal{S}^*\|\|\phi(j)\| + |w(j)| + |v(j+1)| + \|\mathcal{S}^*\|c_\varphi \sum_{q=1}^{n} |v(j-q+1)|;$$

using the bound in (5.42a) and the definition of $\bar{w}$ in (5.12) into the above, with $c_6 :=$

$\frac{\|\mathcal{S}^*\|}{\sqrt{\nu}} + (1 + c_\varphi \|\mathcal{S}^*\|)\,(n+2)$, we see that the following holds:

$$|y(j+1)| \leq c_6 |\bar{w}(j)|, \quad j \in [k_l, k_{l+1}).$$

If we consider this with (5.43), we see that

$$\|\mathcal{Y}(k_{l+1})\| \leq \|\mathcal{Y}(k_{l+1}-1)\| + c_6 |\bar{w}(k_{l+1}-1)|;$$

using (5.42b) to obtain a bound on $\|\mathcal{Y}(k_{l+1}-1)\|$, we conclude that

$$\|\mathcal{Y}(k_{l+1})\| \leq \left(\tfrac{1}{\sqrt{\nu}} + c_6\right) |\bar{w}(k_{l+1}-1)|.$$

So this combined with (5.42b) can be written compactly as

$$\|\mathcal{Y}(j)\| \leq \begin{cases} c_6 |\bar{w}(j)|, & j = k_l, k_l+1, \ldots, k_{l+1}-1 \\ \left(\tfrac{1}{\sqrt{\nu}} + c_6\right) |\bar{w}(j-1)|, & j = k_{l+1}. \end{cases} \tag{5.44}$$

Case 2: $[k_l, k_{l+1})$ lies in S_{good}.

First suppose that $k_{l+1} - k_l \leq 2n$; then by Proposition 5.2 it can be easily proven that there exists a constant $\bar{\gamma}_1$ so that

$$\|\mathcal{Y}(t)\| \leq \bar{\gamma}_1 \lambda^{t-k_l} \|\mathcal{Y}(k_l)\| + \bar{\gamma}_1 \sum_{j=k_l}^{t-1} \lambda^{t-j-1} \left(|y^*(j+1)| + |\bar{w}(j)|\right), \quad t \in [k_l, k_{l+1}]. \tag{5.45}$$

Now suppose that $k_{l+1} - k_l > 2n$. Let $j \in [k_l, k_{l+1})$ be arbitrary. In this case, $\tilde{\phi}(j) \neq 0$ and

$$\frac{\bar{w}(j)^2}{\|\tilde{\phi}(j)\|^2} < \nu. \tag{5.46}$$

By Proposition 5.1, it follows that for any $t_2 > t_1 \geq t_0$:

$$\sum_{j=t_1}^{t_2-1} \rho(j) \frac{|e(j+1)|^2}{\|\tilde{\phi}(j)\|^2} \leq 2\|\tilde{\theta}(t_1)\|^2 + 2\bar{c} \sum_{j=t_1}^{t_2-1} \rho(j) \frac{|\bar{w}(j)|^2}{\|\tilde{\phi}(j)\|^2}. \tag{5.47}$$

If we now use the bound in (5.46) which holds on $[k_l, k_{l+1})$, together with the fact that

$\|\tilde{\theta}(k_l)\| \leq 2\|\mathcal{S}\|$, we see that

$$\sum_{j=k_l}^{j-1} \rho(j) \frac{|e(j+1)|^2}{\|\tilde{\phi}(j)\|^2} \leq 8\|\mathcal{S}\|^2 + 2\bar{c}\nu\,(j-k_l)\,, \quad j \in [k_l, k_{l+1}). \tag{5.48}$$

Now we turn to the update equation (5.40); with $\tau \geq t_0$, we can rewrite this equation as

$$\begin{aligned}
&\|\mathcal{Y}(\tau+(j+1)n)\| \leq c_5 \tilde{\alpha}_n(\tau+jn)\|\mathcal{Y}(\tau+jn)\| \\
&\qquad + c_5\,(1+\tilde{\alpha}_n(\tau+jn)) \underbrace{\sum_{q=0}^{n-1} (|y^*(\tau+jn+q+1)| + |\bar{w}(\tau+jn+q)|)}_{=:\tilde{w}(\tau+jn)}, \quad j \in \mathbb{Z}^+.
\end{aligned} \tag{5.49}$$

Now we analyze the square sum of $\tilde{\alpha}_n(\cdot)$; using the definition of $\tilde{\alpha}_n$ given in (5.38), the definition of α given in (5.26), and the Cauchy-Schwarz property, we see that

$$\begin{aligned}
\sum_{j=q}^{p-1} \tilde{\alpha}_n(\tau+jn)^2 &= \sum_{j=q}^{p-1} \left(\max_{k=0,1,\ldots,n-1} \|\alpha(\tau+jn+k)\| \right)^2 \\
&\leq \sum_{j=q}^{p-1} \left(\sum_{k=0}^{n-1} \|\alpha(\tau+jn+k)\| \right)^2 \\
&\leq n \sum_{j=q}^{p-1} \sum_{k=0}^{n-1} \|\alpha(\tau+jn+k)\|^2 \\
&= n \sum_{j=q}^{p-1} \sum_{k=0}^{n-1} \rho(\tau+jn+k) \frac{|e(\tau+jn+k+1)|^2}{\|\tilde{\phi}(\tau+jn+k)\|^2} \\
&= n \sum_{j=\tau+qn}^{\tau+pn-1} \rho(j) \frac{|e(j+1)|^2}{\|\tilde{\phi}(j)\|^2}, \qquad 0 \leq q < p.
\end{aligned} \tag{5.50}$$

Before proceeding, from (5.50) and (5.48) we obtain an upper bound on $\tilde{\alpha}_n$ on $[k_l, k_{l+1})$:

$$\begin{aligned}
\tilde{\alpha}_n(\tau+jn) &\leq \left(n \sum_{q=\tau+jn}^{\tau+(j+1)n-1} \rho(q) \frac{|e(q+1)|^2}{\|\tilde{\phi}(q)\|^2} \right)^{\frac{1}{2}} \\
&\leq \left(8n\|\mathcal{S}\|^2 + 2\bar{c}\nu n^2 \right)^{\frac{1}{2}} =: c_7,
\end{aligned}$$

$$\text{for } j \in \mathbb{Z}^+ \text{ s.t. } [\tau + jn, \tau + (j+1)n) \subset [k_l, k_{l+1}); \tag{5.51}$$

we can then use this to bound the second occurrence of $\tilde{\alpha}_n$ in (5.49) to yield

$$\|\mathcal{Y}(\tau + (j+1)n)\| \leq c_5 \tilde{\alpha}_n(\tau + jn) \|\mathcal{Y}(\tau + jn)\| + \underbrace{c_5 (1 + c_7)}_{=:c_8} \tilde{w}(\tau + jn),$$
$$j \in \mathbb{Z}^+ \text{ s.t. } [\tau + jn, \tau + (j+1)n) \subset [k_l, k_{l+1}). \tag{5.52}$$

We now utilize the inequality of arithmetic and geometric means.

Claim 5.2. *There exists a constant $\bar{\gamma}_2 > 1$ such that*

$$\prod_{j=q}^{p-1} \tilde{\alpha}_n(\tau + jn) \leq \bar{\gamma}_2 \lambda_1^{p-q},$$
$$\textit{for } q, p \in \mathbb{Z}^+ \textit{ s.t. } k_l \leq \tau + qn < \tau + pn \leq k_{l+1}. \tag{5.53}$$

Proof of Claim 5.2. Let $q, p \in \mathbb{Z}^+$ be arbitrary such that $k_l \leq \tau + qn < \tau + pn \leq k_{l+1}$. By the fact that $\tilde{\alpha}_n(\cdot) \geq 0$, we obtain

$$\prod_{j=q}^{p-1} \tilde{\alpha}_n(\tau + jn) \leq \left[\frac{1}{p-q} \sum_{j=q}^{p-1} \tilde{\alpha}_n(\tau + jn)^2 \right]^{\frac{p-q}{2}}. \tag{5.54}$$

Substituting (5.48) into (5.50) yields

$$\begin{aligned} \sum_{j=q}^{p-1} \tilde{\alpha}_n(\tau + jn)^2 &\leq n \sum_{j=\tau+qn}^{\tau+pn-1} \rho(j) \frac{|e(j+1)|^2}{\|\tilde{\phi}(j)\|^2} \\ &\leq 8n\|\mathcal{S}\|^2 + 2\bar{c}\nu n^2 (p-q); \end{aligned} \tag{5.55}$$

applying this into (5.54), we obtain

$$\prod_{j=q}^{p-1} \tilde{\alpha}_n(\tau + jn) \leq \left[\frac{8n\|\mathcal{S}\|^2}{p-q} + 2\bar{c}\nu n^2 \right]^{\frac{p-q}{2}}. \tag{5.56}$$

So it is enough to prove that there exists a constant $\bar{\gamma}_2$ so that

$$\Bigg(\underbrace{\left[\frac{8n\|\mathcal{S}\|^2}{k} + 2\bar{c}\nu n^2\right]^{\frac{1}{2}}}_{=:g(k)}\Bigg)^k \leq \bar{\gamma}_2\lambda_1^k, \qquad k > 0.$$

We can easily show that with $\bar{k} := 16n \times \left\lceil \left(\frac{\|\mathcal{S}\|}{\lambda_1}\right)^2 \right\rceil$, we have

$$\frac{8n\|\mathcal{S}\|^2}{\bar{k}} \leq \frac{\lambda_1^2}{2},$$

which means that, by the choice of ν in (5.41), we see that

$$g(k)^k \leq \lambda_1^k, \qquad k \geq \bar{k}.$$

Since $g(k)$ decreases as $k \geq 1$ increases, we conclude that if we define $\bar{\gamma}_2 := \max\left\{1, \left(\frac{g(1)}{\lambda_1}\right)^{\bar{k}}\right\}$, then

$$g(k)^k \leq \bar{\gamma}_2\lambda_1^k, \qquad k = 1, 2, \ldots, \bar{k},$$

as well, so the claim holds. □

First of all, using the bound in (5.53) and the definition of λ_1, we obtain

$$\begin{aligned}\prod_{j=q}^{p-1} [c_5\tilde{\alpha}_n(\tau + jn)] &\leq \bar{\gamma}_2\lambda_1^{p-q}c_5^{p-q} \\ &\leq \bar{\gamma}_2\lambda^{(p-q)n}, \\ &\quad \text{for } q, p \in \mathbb{Z}^+ \text{ s.t. } k_l \leq \tau + qn < \tau + pn \leq k_{l+1}.\end{aligned} \tag{5.57}$$

If we solve (5.52) iteratively and use the above bound, we obtain

$$\begin{aligned}\|\mathcal{Y}(\tau + pn)\| \leq \bar{\gamma}_2\lambda^{(p-q)n}\|\mathcal{Y}(\tau + qn)\| + \sum_{j=q}^{p-1} c_8\bar{\gamma}_2(\lambda^n)^{p-j-1}\tilde{w}(\tau + jn), \\ \text{for } q, p \in \mathbb{Z}^+, \text{ s.t. } k_l \leq \tau + qn < \tau + pn \leq k_{l+1}.\end{aligned} \tag{5.58}$$

We can now use Proposition 5.2 (for no more than n steps at a time):

- to provide a bound on $\|\mathcal{Y}(t)\|$ between consecutive $(\tau + jn)$'s ;
- to provide a bound on $\|\mathcal{Y}(t)\|$ on the beginning part of the interval $[k_l, k_{l+1})$, until we get to the first admissible $\tau + jn$;
- to provide a bound on $\|\mathcal{Y}(t)\|$ on the last part of the interval $[k_l, k_{l+1})$, after the last admissible $\tau + jn$.

After simplification, we conclude that there exists a constant $\bar{\gamma}_3 \geq \bar{\gamma}_1$ so that

$$\|\mathcal{Y}(t)\| \leq \bar{\gamma}_3 \lambda^{t-k_l} \|\mathcal{Y}(k_l)\| + \bar{\gamma}_3 \sum_{j=k_l}^{t-1} \lambda^{t-j-1} \left(|y^*(j+1)| + |\bar{w}(j)| \right), \quad t \in [k_l, k_{l+1}], \tag{5.59}$$

as desired.

Step 4: Combine Case 1 and Case 2 of Step 3 into a general bound on $\mathcal{Y}$.

Now we will combine Case 1 and Case 2 into a general bound on $\mathcal{Y}$. Using an argument similar to that used in the end of proof of the main result in Chapter 4, we glue the bounds of Case 1 and Case 2 together. To this end, first define

$$\bar{\gamma} := \max \left\{ \bar{\gamma}_3, \left(c_6 + \tfrac{1}{\sqrt{\nu}} \right), \bar{\gamma}_3 \left(c_6 + \tfrac{1}{\sqrt{\nu}} \right) \right\}.$$

Claim 5.3. *The following bound holds:*

$$\|\mathcal{Y}(t)\| \leq \bar{\gamma} \lambda^{t-\tau} \|\mathcal{Y}(\tau)\| + \sum_{j=\tau}^{t-1} \bar{\gamma} \lambda^{t-j-1} (|y^*(j+1)| + |\bar{w}(j)|), \quad t \geq \tau. \tag{5.60}$$

Proof of the Claim 5.3. If $[k_0, k_1) = [\tau, k_1) \subset S_{good}$, then (5.60) is true for $t \in [k_0, k_1]$ by (5.45) and (5.59). If $[k_0, k_1) \subset S_{bad}$, then by (5.44) we have

$$\|\mathcal{Y}(j)\| \leq \begin{cases} \|\mathcal{Y}(\tau)\| & j = k_0 = \tau \\ \frac{1}{\sqrt{\nu}} |\bar{w}(j)| & j = k_0 + 1, k_0 + 2, \ldots, k_1 - 1 \\ \left(c_6 + \frac{1}{\sqrt{\nu}} \right) |\bar{w}(j-1)| & j = k_1, \end{cases}$$

which means that (5.60) holds on $[k_0, k_1]$ for this case as well.

We now use induction: suppose that (5.60) is true for $t \in [k_0, k_l]$; we need to prove that it holds for $t \in (k_l, k_{l+1}]$ as well. If $[k_l, k_{l+1}) \subset S_{bad}$, then from (5.44) we see that

$$\|\mathcal{Y}(j)\| \leq \begin{cases} \frac{1}{\sqrt{\nu}}|\bar{w}(j)| & j = k_l, k_l+1, \ldots, k_{l+1}-1 \\ \left(c_6 + \frac{1}{\sqrt{\nu}}\right)|\bar{w}(j-1)| & j = k_{l+1}, \end{cases}$$

which means (5.60) holds on $(k_l, k_{l+1}]$. On the other hand, suppose that $[k_l, k_{l+1}) \subset S_{good}$. Using (5.45) and (5.59) to analyze the behavior on $[k_l, k_{l+1}]$, we have

$$\|\mathcal{Y}(t)\| \leq \bar{\gamma}_3 \lambda^{t-k_l} \|\mathcal{Y}(k_l)\| + \bar{\gamma}_3 \sum_{j=k_l}^{t-1} \lambda^{t-j-1} \left(|y^*(j+1)| + |\bar{w}(j)|\right), \quad t \in [k_l, k_{l+1}].$$

But $k_l - 1 \in S_{bad}$, so by (5.44) we have

$$\|\mathcal{Y}(k_l)\| \leq \left(c_6 + \frac{1}{\sqrt{\nu}}\right)|\bar{w}(k_l - 1)|;$$

combining these, we obtain

$$\begin{aligned} \|\mathcal{Y}(t)\| &\leq \bar{\gamma}_3 \lambda^{t-k_l} \left(c_6 + \frac{1}{\sqrt{\nu}}\right)|\bar{w}(k_l-1)| + \bar{\gamma}_3 \sum_{j=k_l}^{t-1} \lambda^{t-j-1} \left(|y^*(j+1)| + |\bar{w}(j)|\right) \\ &\leq \bar{\gamma} \sum_{j=k_l-1}^{t-1} \lambda^{t-j-1} \left(|y^*(j+1)| + |\bar{w}(j)|\right), \quad t \in [k_l, k_{l+1}], \end{aligned}$$

which means that (5.60) holds on $(k_l, k_{l+1}]$ in this case as well. This concludes the proof of the claim. □

Step 5: Obtain a bound on $u(t)$.

Now we derive a bound on $u(t)$. Substituting (5.60) into (5.30) yields

$$\begin{aligned} |u(t)| \leq\, & c_\theta c_\varphi \|\mathcal{S}\| \bar{\gamma} \lambda^{t-\tau} \|\mathcal{Y}(\tau)\| + \\ & c_\theta c_\varphi \|\mathcal{S}\| \bar{\gamma} \sum_{j=\tau}^{t-1} \lambda^{t-j-1}(|y^*(j+1)| + |\bar{w}(j)|) + c_\theta |y^*(t+1)|, \quad t \geq \tau. \end{aligned} \tag{5.61}$$

Now we combine the above with (5.60): there clearly exists a constant γ_1 so that

$$\left\| \begin{bmatrix} \mathcal{Y}(t) \\ u(t) \end{bmatrix} \right\| \le \gamma_1 \lambda^{t-\tau} \|\mathcal{Y}(\tau)\| + \sum_{j=\tau}^{t-1} \gamma_1 \lambda^{t-j-1} (|y^*(j+1)| + |\bar{w}(j)|) + \gamma_1 |y^*(t+1)|, \qquad t \ge \tau. \tag{5.62}$$

Since $\tau \ge t_0$ is arbitrary, we conclude the proof. ■

5.3.4 Robustness Results

Now we show that the exponential stability property and the linear-like convolution bounds proven in Theorem 5.1 will guarantee robustness to a degree of time-variations and unmodelled dynamics. In this way, the approach has a lot in common with LTI systems, which also enjoys this feature. To this end, we consider a time-varying version of the plant (5.5) along with the term $d_\Delta(t) \in \mathbb{R}$ added to represent the unmodelled dynamics:

$$\begin{aligned} y(j+1) &= \theta^*(t)^\top \begin{bmatrix} \varphi\big(\mathcal{X}(j)\big) \\ u(j) \end{bmatrix} + w(j) + v(j+1) + d_\Delta(t) \\ &= \theta^*(t)^\top \phi(j) + w(j) + v(j+1) + \underbrace{\theta^*(t)^\top \begin{bmatrix} \varphi\big(\mathcal{X}(j)\big) - \varphi\big(\mathcal{Y}(j)\big) \\ 0 \end{bmatrix}}_{=:\tilde{\Delta}(j)} + d_\Delta(t). \end{aligned} \tag{5.63}$$

As discussed in Chapter 2, with $g : \mathbb{R}^{n+1} \mapsto \mathbb{R}$ having a bounded gain, we assume that d_Δ satisfies

$$m(t+1) = \beta m(t) + \beta \left| g\left(\begin{bmatrix} \mathcal{Y}(t) \\ u(t) \end{bmatrix} \right) \right|, \qquad m(t_0) = m_0 \tag{5.64a}$$

$$|d_\Delta(t)| \le \mu m(t) + \mu \left| g\left(\begin{bmatrix} \mathcal{Y}(t) \\ u(t) \end{bmatrix} \right) \right|, \qquad t \ge t_0. \tag{5.64b}$$

Theorem 5.2. *Suppose that the adaptive controller (5.16) and (5.21) is applied to the time-varying nonlinear plant (5.63) with d_Δ satisfying (5.64). Then for every $\delta \in (0, \infty]$, $\beta \in (0,1)$ and $\bar{c}_0 \geq 0$, there exist $\bar{\epsilon} > 0$, $\mu > 0$, $\tilde{\lambda} \in (\beta, 1)$ and $\tilde{\gamma} > 0$ such that for every $t_0 \in \mathbb{Z}$, $\mathcal{X}(t_0) \in \mathbb{R}^n$, $\theta^* \in \mathbb{S}(\mathcal{S}^*, \bar{c}_0, \bar{\epsilon})$, $\hat{\theta}(t_0) \in \mathcal{S}$, and $w, v, y^* \in \boldsymbol{\ell}_\infty$, the following holds:*

$$\left\| \begin{bmatrix} \mathcal{Y}(t) \\ u(t) \\ m(t) \end{bmatrix} \right\| \leq \tilde{\gamma} \tilde{\lambda}^{t-t_0} \left\| \begin{bmatrix} \mathcal{Y}(t_0) \\ u(t_0) \\ m_0 \end{bmatrix} \right\| + \sum_{j=t_0}^{t-1} \tilde{\gamma} \tilde{\lambda}^{t-j-1} (|\bar{w}(j)| + |y^*(j+1)|) + \tilde{\gamma} |y^*(t+1)|, \qquad t \geq t_0.$$

Proof of Theorem 5.2. We observe here that the plant (5.63) and the controller (5.16) and (5.21) fit into the paradigm of Chapter 2: we set

$$\begin{aligned}
\vartheta(t) &= \begin{bmatrix} \mathcal{Y}(t) \\ u(t) \end{bmatrix}, \\
f(\vartheta(\cdot)) = \phi(\cdot) &= \begin{bmatrix} \varphi\big(\mathcal{Y}(\cdot)\big) \\ u(\cdot) \end{bmatrix}, \\
z(t) &= \varnothing, \\
\hat{\theta}(t) &= \hat{\theta}(t), \\
r(t) &= y^*(t+1), \\
w(t) &= w(t) + v(t+1) + \tilde{\Delta}(t), \\
\Omega &= \mathcal{S}.
\end{aligned}$$

In Theorem 5.1 it is proven the controller (5.16) and (5.21) provides a convolution bound for (5.5), which is equivalent to (5.10). Observe that the term $\tilde{\Delta}(t)$ has the same upper bound as its counterpart in (5.10) because $\theta^*(t) \in \mathcal{S}$ for all $t \in \mathbb{Z}$; indeed, we utilize the upper bound given in (5.11). By the definition of $\bar{w}$ given in (5.12), it means that there exists a constant c so that

$$|w(t) + v(t+1) + \tilde{\Delta}(t)| \leq c|\bar{w}(t)|.$$

We apply Theorems 2.2, 2.2 and 2.3 to immediately see that the linear-like convolution bound holds in the presence of parameter time-variation and/or unmodelled dynamics. ∎

5.3.5 Tracking Results

We now move from the stability problem to the tracking problem. We first analyze the case when the disturbance is absent: we start with the original case of constant parameters, and then we move to the case in which the parameters are slowly time-varying. Second of all, we consider the original case with a disturbance.

Constant Parameters - No Disturbance

In the literature it is typically proven that the tracking error is square summable, e.g. see [20]. Here we can prove an explicit bound on the 2−norm of the error signal in terms of the plant initial condition and the size of the reference signal, which is a significant improvement.

Theorem 5.3. *Suppose that the adaptive controller* (5.16) *and* (5.21) *is applied to the nonlinear plant* (5.5) *in the presence of a zero disturbance* $w = v = 0$*. Then for every* $\delta \in (0, \infty]$ *and* $\lambda \in (0, 1)$*, there exists a constant* $c > 0$ *so that, for every* $t_0 \in \mathbb{Z}$*,* $\theta^* \in \mathcal{S}^*$*,* $y^* \in \boldsymbol{\ell}_\infty$*,* $\hat{\theta}(t_0) \in \mathcal{S}$*, and* $\mathcal{X}(t_0) \in \mathbb{R}^n$*, the following bound holds:*

$$\sum_{t=t_0+1}^{\infty} \varepsilon(t)^2 \leq c\left(\|\mathcal{X}(t_0)\|^2 + \|y^*\|_\infty^2\right).$$

Proof. See Appendix C. ■

Slowly Time-Varying Parameters - No Disturbance

Now we turn to the case in which the plant parameter is slowly time-varying (with no jumps in the parameters) and the disturbance is zero. We should not expect to get exact tracking; we will be able to prove, roughly speaking, that the average tracking error is small on average if the time-variation is small. To proceed, we consider the time-varying plant (5.63) without the unmodelled dynamics and with zero noise ($v = w = 0$):

$$y(t+1) = x(t+1) = \theta^*(t)^\top \begin{bmatrix} \varphi\big(\mathcal{X}(t)\big) \\ u(t) \end{bmatrix} = \theta^*(t)^\top \phi(t). \tag{5.65}$$

Theorem 5.4. *For every $\delta \in (0, \infty]$, there exist constants $\bar{\epsilon} > 0$ and $\gamma > 0$ so that for every $t_0 \in \mathbb{Z}$, $\epsilon \in (0, \bar{\epsilon})$, $\mathcal{X}(t_0) \in \mathbb{R}^n$, $\theta^* \in \mathbb{S}(\mathcal{S}^*, 0, \epsilon)$, $\hat{\theta}(t_0) \in \mathcal{S}$, and $y^* \in \boldsymbol{\ell}_\infty$, when the adaptive controller (5.16) and (5.21) is applied to the time-varying nonlinear plant (5.65), the following holds:*

$$\limsup_{T \to \infty} \frac{1}{T} \sum_{j=t_0}^{t_0+T-1} |\varepsilon(j)|^2 \quad \leq \quad \gamma \epsilon^{2/3} \|y^*\|_\infty^2.$$

Proof. See Appendix C. ∎

Remark 5.6. *The proof of Theorem 5.4 is a modified version of a proof in the paper Miller and Shahab [48] which deals with the LTI one-step-ahead adaptive control paradigm.*

Tracking in the Presence of a Disturbance

Now we turn to the much harder problem of tracking in the presence of a disturbance; throughout this sub-section we assume that plant parameters are constant. Our goal is to show that if the noise is small, then the tracking error is small; this is a stringent requirement, since in adaptive control it is usually only proven that if the noise is bounded, then the error is bounded. We can, of course, measure signal sizes in a variety of ways, with the 2−norm and the ∞-norm the most common; given that a large disturbance can lead the estimator astray and cause “temporary instability”, the 2−norm seems to be the most appropriate here.

If the closed-loop system were LTI, then by Parseval's Theorem we could conclude that the average power of the tracking error is bounded by the average power of the disturbance, i.e. there exists a constant c so that

$$\limsup_{T \to \infty} \frac{1}{T} \sum_{j=t_0}^{t_0+T-1} [\varepsilon(t)]^2 \leq c \times \limsup_{T \to \infty} \frac{1}{T} \sum_{j=t_0}^{t_0+T-1} [w(t)]^2; \tag{5.66}$$

unfortunately, while the closed-loop system has some desirable linear-like closed-loop properties, the closed-loop system is clearly not LTI.

We will prove something weaker than (5.66), but with much the same flavor; it is, however, stronger than the standard result in the literature.

Theorem 5.5. *Suppose that the adaptive controller* (5.16) *and* (5.21) *is applied to the nonlinear plant* (5.5)*. Then for every* $\delta \in (0, \infty]$*, there exists a constant* $\gamma > 0$ *so that, for every* $t_0 \in \mathbb{Z}$*,* $\theta^* \in \mathcal{S}^*$*,* $y^*, w, v \in \boldsymbol{\ell}_\infty$*,* $\hat{\theta}(t_0) \in \mathcal{S}$*, and* $\mathcal{X}(t_0) \in \mathbb{R}^n$*, the following holds: if* $\liminf_{t\to\infty} |y^*(t)| > 0$*, then*

$$\limsup_{T\to\infty} \frac{1}{T} \sum_{j=t_0}^{t_0+T-1} |\varepsilon(j)|^2 \leq \gamma \times \limsup_{T\to\infty} \frac{1}{T} \sum_{j=t_0}^{t_0+T-1} (|w(j)|^2 + |v(j)|^2) \times \frac{\limsup_{t\to\infty} |y^*(t)|^2 + \limsup_{t\to\infty}(|w(t)|^2 + |v(t)|^2)}{\liminf_{t\to\infty} |y^*(t)|^2}. \tag{5.67}$$

Proof. See Appendix C. ■

Remark 5.7. *So we see that the bound proven here is similar to that of* (5.66) *which holds in the LTI case, although we have an extra term multiplied on the RHS:*

$$\frac{\limsup_{t\to\infty} |y^*(t)|^2 + \limsup_{t\to\infty}(|w(t)|^2 + |v(t)|^2)}{\liminf_{t\to\infty} |y^*(t)|^2}.$$

If the reference signal is larger than the noise, which is what one would normally expect, then this would be bounded by

$$2\frac{\limsup_{t\to\infty} |y^*(t)|^2}{\liminf_{t\to\infty} |y^*(t)|^2};$$

if $|y^*(t)| \in \{-1, 1\}$ *then this is exactly two. It is curious that the quantity gets large if* $y^*(t)$ *gets close to zero; we suspect that this is an artifact of the proof, since all simulations indicate that the LTI-like bound* (5.66) *holds.*

Remark 5.8. *The proof of Theorem 5.5 is a modified version of a proof in the paper Miller and Shahab [48] which dealt with the LTI one-step-ahead adaptive control paradigm.*

5.4 The Case of the Sign of θ_p being Unknown

As shown earlier, the convexity requirement on the set of uncertainty plays a crucial role in proving nice closed-loop properties. In this section, we only impose Assumptions 5.1 and 5.2; the set of admissible parameters may not be convex. In this case, the standard trick in adaptive control is to replace the set with its closed convex hull; however, that set may contain "uncontrollable models", i.e. it may be that $\theta_p = 0$. If this is the case then instead, we "cover" 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 two convex sets.

Proposition 5.3. *For every compact set $\mathcal{S}^* \subset \{\theta^* \in \mathbb{R}^p : \theta_p \neq 0\}$, there exist compact and convex sets $\mathcal{S}_1$ and $\mathcal{S}_2$ which also lie in $\{\theta^* \in \mathbb{R}^p : \theta_p \neq 0\}$ such that $\mathcal{S}^* \subset \mathcal{S}_1 \cup \mathcal{S}_2$.*

Proof of Proposition 5.3. For a given $\mathcal{S}^*$, define

$$\mathcal{S}_1 := \text{convex hull of } \{\theta^* \in \mathcal{S}^* : \theta_p > 0\},$$
$$\mathcal{S}_2 := \text{convex hull of } \{\theta^* \in \mathcal{S}^* : \theta_p < 0\}.$$

The result follows immediately. ■

At this point we assume that we have at hand sets $\mathcal{S}_1$ and $\mathcal{S}_2$ of the form discussed in Proposition 5.3. Similar to Chapter 4, now define the index set

$$\mathcal{I}^* := \{1, 2\}.$$

For each $\theta^* \in \mathcal{S}_i$, $i = 1, 2$, we define

$$i^*(\theta^*) = \min \{i \in \mathcal{I}^* : \theta^* \in \mathcal{S}_i\};$$

when there is no ambiguity, we will drop the argument and simply write i^*. Before proceeding, define

$$\bar{\mathbf{s}} := \max_i \|\mathcal{S}_i\|.$$

5.4.1 Parameter Estimation

For $i = 1$ and 2, given an estimate $\hat{\theta}_i(t)$ of θ^* at time t, we define the prediction error by

$$e_i(t+1) := y(t+1) - \hat{\theta}_i(t)^\top \phi(t); \tag{5.68}$$

as usual, this is a measure of the error in $\hat{\theta}_i(t)$, since it is zero if $\hat{\theta}_i(t) = \theta^*$ and $w = v = 0$. For each $i = 1, 2$, we will be using the same projection algorithm as in Section 5.3.1; however here we project updates onto the corresponding $\mathcal{S}_i$. To this end, with $\delta \in (0,\infty]$ and

$$\tilde{\phi}(t) = \begin{bmatrix} \phi(t) \\ \mathcal{Y}(t) \end{bmatrix},$$

we define $\rho_i : \mathbb{Z} \mapsto \{0,1\}$ by

$$\rho_i(t) := \begin{cases} 1 & \text{if } |e_i(t+1)| < (2\bar{\mathbf{s}}+\delta)\|\tilde{\phi}(t)\| \\ 0 & \text{otherwise;} \end{cases} \tag{5.69}$$

here ρ_i is defined in this manner as i^* is unknown and $\|\mathcal{S}_{i^*}\| \leq \bar{\mathbf{s}}$. With an initial condition of $\hat{\theta}_i(t_0) \in \mathcal{S}_i$, for $t \geq t_0$ the i^{th} estimator is updated via

$$\check{\theta}_i(t+1) = \hat{\theta}_i(t) + \rho_i(t)\frac{\phi(t)}{\|\tilde{\phi}(t)\|^2}e_i(t+1) \tag{5.70a}$$

$$\hat{\theta}_i(t+1) = \operatorname*{Proj}_{\mathcal{S}_i}\left\{\check{\theta}_i(t+1)\right\}. \tag{5.70b}$$

Define the parameter estimation error

$$\tilde{\theta}_i(t) := \hat{\theta}_i(t) - \theta^*.$$

The following estimator property follows directly from Proposition 5.1. Of course, the difference here is that we do not know the value of i^*.

Proposition 5.4. *There exists a constant $\bar{c} > 0$ such that for every $t_0 \in \mathbb{Z}$, $\mathcal{X}(t_0) \in \mathbb{R}^n$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ $(i = 1,2)$, $\theta^* \in \mathcal{S}^*$, and $v, w \in \boldsymbol{\ell}_\infty$, when the estimator (5.70) is applied to the plant (5.5), the following holds:*

$$\left\|\tilde{\theta}_{i^*}(t)\right\|^2 \leq \left\|\tilde{\theta}_{i^*}(\tau)\right\|^2 + \sum_{j=\tau}^{t-1}\rho_{i^*}(j)\left[-\frac{1}{2}\frac{e_{i^*}(j+1)^2}{\|\tilde{\phi}(j)\|^2} + \bar{c}\frac{\bar{w}(j)^2}{\|\tilde{\phi}(j)\|^2}\right], \quad t > \tau \geq t_0. \tag{5.71}$$

5.4.2 The Switching Control Law

For each i, we partition $\hat{\theta}_i(t)$ in a natural way as

$$\hat{\theta}_i(t) =: \begin{bmatrix} \hat{\theta}_{i,1}(t) \\ \hat{\theta}_{i,2}(t) \\ \vdots \\ \hat{\theta}_{i,p}(t) \end{bmatrix},$$

and define $\hat{\hat{\theta}}_i(t)$ to be the first $p-1$ elements:

$$\hat{\hat{\theta}}_i(t) := \begin{bmatrix} \hat{\theta}_{i,1}(t) \\ \hat{\theta}_{i,2}(t) \\ \vdots \\ \hat{\theta}_{i,p-1}(t) \end{bmatrix}.$$

With $y^*(\cdot)$ as the reference signal to be tracked, the natural choice for the one-step-ahead adaptive control law associated with the i^{th} estimator is given by

$$u(t) = \frac{1}{\hat{\theta}_{i,p}(t)} \left[y^*(t+1) - \hat{\hat{\theta}}_i(t)^\top \varphi\Big(\mathcal{Y}(t)\Big) \right];$$

note that if θ^* were known (i.e. $\hat{\theta}_i(t) = \theta^*$) and no noise is entering the system, this control law will ensure that $y(t+1) = y^*(t+1)$.

We define a switching signal $\sigma : \mathbb{Z} \mapsto \mathcal{I}^*$ that decides which controller to use at any given point in time; in the next sub-section, we will show how to choose $\sigma(t)$. With that in mind, we set the control law to be

$$u(t) = \frac{1}{\hat{\theta}_{\sigma(t),p}(t)} \left[y^*(t+1) - \hat{\hat{\theta}}_{\sigma(t)}(t)^\top \varphi\Big(\mathcal{Y}(t)\Big) \right]. \tag{5.72}$$

With the tracking error defined as in the previous section, namely $\varepsilon(t) := y(t) - y^*(t)$, from (5.72) we see that

$$y^*(t+1) = \hat{\theta}_{\sigma(t)}(t)^\top \phi(t);$$

combining this with (5.68) yields

$$e_{\sigma(t)}(t+1) = \varepsilon(t+1), \qquad t \geq t_0. \tag{5.73}$$

5.4.3 The Switching Algorithm

Unlike in the first-order LTI case of Chapter 4, to facilitate analysis, we update $\sigma(t)$ only every n steps; however, we keep the estimators running at all times. The effect of this will become clear in the proof of the main result of this section. To this end, we define a sequence of switching times as follows: we initialize $\hat{t}_0 := t_0$ and then define

$$\hat{t}_\ell := \hat{t}_0 + \ell n, \qquad \ell \in \mathbb{N};$$

the switching signal is piecewise constant of the form

$$\sigma(t) = \sigma(\hat{t}_\ell), \qquad t \in [\hat{t}_\ell, \hat{t}_{\ell+1}), \ \ \ell \in \mathbb{Z}^+. \tag{5.74}$$

For each $i \in \mathcal{I}^*$ we define a performance signal $J_i : \left\{\hat{t}_0, \hat{t}_1, \hat{t}_2, \ldots\right\} \to \mathbb{R}^+$ by

$$J_i(\hat{t}_\ell) := \sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+1}-1} \rho_i(j) \frac{|e_i(j+1)|}{\|\tilde{\phi}(j)\|}; \tag{5.75}$$

in the absence of noise, this is a crude measure of the size of $\|\hat{\theta}_i(\hat{t}_{\ell+1}) - \hat{\theta}_i(\hat{t}_\ell)\|$. With $\sigma(\hat{t}_0) = \sigma_0$, we choose the following switching rule:

$$\sigma(\hat{t}_{\ell+1}) = \operatorname*{argmin}_{i \in \mathcal{I}^*} J_i(\hat{t}_\ell), \qquad \ell \in \mathbb{Z}^+. \tag{5.76}$$

For the case when $J_1(\hat{t}_\ell) = J_2(\hat{t}_\ell)$, we (somewhat arbitrarily) select $\sigma(\hat{t}_{\ell+1})$ to be 1.

Lemma 5.1. *Suppose that the adaptive controller (5.70), (5.72), (5.74), (5.75) and (5.76) is applied to the plant (5.5). Then for every $t_0 \in \mathbb{Z}$, $\mathcal{X}(t_0) \in \mathbb{R}^n$, $\sigma_0 \in \mathcal{I}^*$, $\theta^* \in \mathcal{S}^*$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ $(i = 1, 2)$ and $y^*, w, v \in \boldsymbol{\ell}_\infty$, we have that for every $\ell \in \mathbb{Z}^+$,*

(a) $J_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell) \leq J_{i^*}(\hat{t}_\ell)$ *or*

(b) $J_{\sigma(\hat{t}_{\ell+1})}(\hat{t}_{\ell+1}) \leq J_{i^*}(\hat{t}_{\ell+1})$.

Proof of Lemma 5.1. Fix $t_0 \in \mathbb{Z}$, $\mathcal{X}(t_0) \in \mathbb{R}^n$, $\sigma_0 \in \mathcal{I}^*$, $\theta^* \in \mathcal{S}^*$, and $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ $(i = 1, 2)$ and $y^*, w, v \in \boldsymbol{\ell}_\infty$, and let $\ell \in \mathbb{Z}^+$ be arbitrary.

Assume that (a) does not hold, i.e. $J_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell) > J_{i^*}(\hat{t}_\ell)$; then according to (5.76), this means that $\sigma(\hat{t}_{\ell+1}) = i^*$, i.e. (b) will hold. ∎

In the above we do not make any claim that $\theta^* \in \mathcal{S}_{\sigma(t)}$ at any time; it only makes a indirect statement about the size of the prediction error. It turns out that this is enough to ensure that the desirable closed-loop properties are attained.

5.4.4 The Main Result

We now prove that the proposed switching adaptive controller has the same desirable linear-like properties as in the case when sign of θ_p is known.

Theorem 5.6. *Suppose that the adaptive controller (5.70), (5.72), (5.74), (5.75) and (5.76) is applied to the nonlinear plant (5.5). Then for every $\delta \in (0, \infty]$ and $\lambda \in (0, 1)$, there exists a constant $\gamma > 0$ so that, for every $t_0 \in \mathbb{Z}$, $\mathcal{X}(t_0) \in \mathbb{R}^n$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$, $\theta^* \in \mathcal{S}^*$, and $v, w, y^* \in \boldsymbol{\ell}_\infty$, the following holds:*

$$\left\| \begin{bmatrix} \mathcal{Y}(t) \\ u(t) \end{bmatrix} \right\| \leq \gamma \lambda^{t-\tau} \|\mathcal{Y}(\tau)\| + \sum_{j=\tau}^{t-1} \gamma \lambda^{t-\tau-1} \left(|y^*(j+1)| + |\bar{w}(j)| \right) + \gamma |y^*(t+1)|,$$
$$t > \tau \geq t_0. \tag{5.77}$$

Remark 5.9. *The above result shows that the closed-loop system experiences linear-like behavior. There is a uniform exponential decay bound on the effect of the initial condition, and there is a convolution sum bound on the effect of the exogenous signals. As in the previous section, this implies that the system has a bounded gain (from w, v and y^* to y) in every p-norm; in particular for $p = \infty$, it follows from (5.77) that there exists $c > 0$ such that*

$$\left\| \begin{bmatrix} \mathcal{Y}(t) \\ u(t) \end{bmatrix} \right\| \leq \frac{c\gamma}{1-\lambda} \left(\lambda^{t-t_0} \|\mathcal{Y}(t_0)\| + \|y^*\|_\infty + \|w\|_\infty + \|v\|_\infty \right), \qquad t \geq t_0.$$

Hence, if $w, v, y^* \in \boldsymbol{\ell}_\infty$, then $y, x, u \in \boldsymbol{\ell}_\infty$, so ε, e_i $(i \in \mathcal{I}^*)$ lie in $\boldsymbol{\ell}_\infty$ as well; all signals in the closed-loop system are uniformly bounded.

Similar to Proposition 5.2 that provided a crude bound on the closed-loop behavior in the case of one parameter estimator and a non-switching control law, the following also provide a similar crude bound on the closed-loop behavior for the case considered here. Due to the boundedness of $\mathcal{S}^*$, $\mathcal{S}_1$ and $\mathcal{S}_2$, it is easy to prove the following result in the same way as that of the proof of Proposition 5.2.

Proposition 5.5. *Suppose that the adaptive controller (5.70), (5.72), (5.74), (5.75) and (5.76) is applied to (5.5). Then for every $p \geq 0$, there exists a constant $c_1 > 0$ so that, for every $t_0 \in \mathbb{Z}$, $\mathcal{X}_0 \in \mathbb{R}$, $\sigma_0 \in \mathcal{I}^*$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ $(i = 1, 2)$, $\theta^* \in \mathcal{S}^*$, and $v, w, r \in \boldsymbol{\ell}_\infty$, the following holds:*

$$\|\mathcal{Y}(t+p)\| \leq c_1 \|\mathcal{Y}(t)\| + c_1 \sum_{j=0}^{p-1} \left(|\bar{w}(t+j)| + |y^*(t+j+1)|\right). \tag{5.78}$$

Before proceeding, define

$$\hat{c}_\theta := \max\left\{\frac{1}{|\theta_p|} : \theta^* \in \mathcal{S}_i,\ i = 1, 2\right\}.$$

We now proceed to prove the Theorem.

Proof of Theorem 5.6:

Fix $\delta \in (0, \infty]$ and $\lambda \in (0, 1)$. Let $t_0 \in \mathbb{Z}$, $\mathcal{X}(t_0) \in \mathbb{R}^n$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$, $(i = 1, 2)$, $\sigma_0 \in \mathcal{I}^*$, $\theta^* \in \mathcal{S}^*$, and $v, w, r \in \boldsymbol{\ell}_\infty$ be arbitrary. In many aspects, this proof will be similar to the proof of Theorem 5.1; the key difference is in Step 3 where we analyze the switching behavior in the closed-loop system.

Step 1: Construct a useful difference equation.

From the tracking error we have

$$y(t+1) = \varepsilon(t+1) + y^*(t+1);$$

then from (5.73), we have

$$y(t+1) = e_{\sigma(t)}(t+1) + y^*(t+1). \tag{5.79}$$

Now define:

$$\alpha_i(t) := \rho_i(t) \frac{e_i(t+1)}{\|\tilde{\phi}(t)\|^2} \tilde{\phi}(t)^\top, \qquad i \in \mathcal{I}^*; \tag{5.80}$$

it is clear that

$$\alpha_i(t)\tilde{\phi}(t) = \rho_i(t) e_i(t+1). \tag{5.81}$$

Also, observe that

$$e_i(t+1) = \rho_i(t) e_i(t+1) + \underbrace{[1 - \rho_i(t)] e_i(t+1)}_{=:\eta_i(t)}, \qquad i \in \mathcal{I}^*;$$

if we combine this with (5.81) and substitute the result into (5.79), then we obtain

$$y(t+1) = \alpha_{\sigma(t)}(t)\tilde{\phi}(t) + \eta_{\sigma(t)}(t) + y^*(t+1). \tag{5.82}$$

Step 2: Analyze the difference equation (5.82).

We will analyze (5.82) in somewhat the same way as inside the proof of Theorem 5.1. Before proceeding, let us analyze the term $|\eta_{\sigma(t)}(t)|$.

Claim 5.4. *There exists a constant c_1 such that*

$$|\eta_i(t)| \le c_1 |\bar{w}(t)|, \qquad t \ge t_0, \;\; i \in \mathcal{I}^*. \tag{5.83}$$

Proof of Claim 5.4. If $\rho_i(t) = 1$, then $\eta_i(t) = 0$. If $\rho_i(t) = 0$, then $\eta_i(t) = e_i(t+1)$ and from the estimator definition we obtain

$$|e_i(t+1)| \ge (2\bar{\mathbf{s}} + \delta)\|\tilde{\phi}(t)\|;$$

similar to the proof of Claim 5.1, we can also obtain

$$|e_i(t+1)| \le 2\bar{\mathbf{s}}\|\tilde{\phi}(t)\| + (1 + c_\varphi \|\mathcal{S}^*\|)(n+2)|\bar{w}(t)|. \tag{5.84}$$

Combining the above two statements yields

$$\begin{aligned} 2\bar{\mathbf{s}}\|\tilde{\phi}(t)\| + (1 + c_\varphi \|\mathcal{S}^*\|)(n+2)|\bar{w}(t)| &\ge (2\bar{\mathbf{s}} + \delta)\|\tilde{\phi}(t)\| \\ \Rightarrow \|\tilde{\phi}(t)\| &\le \frac{(1 + c_\varphi \|\mathcal{S}^*\|)(n+2)}{\delta}|\bar{w}(t)|; \end{aligned}$$

if we combine this with (5.84), we end up with the desired bound:

$$|\eta_i(t)| = |e_i(t+1)| \le \underbrace{(1 + c_\varphi \|\mathcal{S}^*\|)(n+2)\left(\frac{2\bar{\mathbf{s}}}{\delta} + 1\right)}_{=:c_1} |\bar{w}(t)|, \qquad t \ge t_0, \;\; i \in \mathcal{I}^*.$$

$\square$

Using the definition of $\mathcal{Y}(\cdot)$, the formula for $y(t+1)$ in (5.82), Claim 5.4, and (5.29)–(5.37) inside the proof of Theorem 5.1, we conclude that there exists a constant c_2 such that the following holds:

$$\|\mathcal{Y}(t+n)\| \le \sum_{j=0}^{n-1} \Bigg[c_2 \|\alpha_{\sigma(t+j)}(t+j)\| \|\mathcal{Y}(t+j)\| + \\ \Big(1 + \hat{c}_\theta \|\alpha_{\sigma(t+j)}(t+j)\|\Big) |y^*(t+j+1)| + c_1 |\bar{w}(t+j)| \Bigg], \quad t \ge t_0. \tag{5.85}$$

Now define the quantity

$$\tilde{\alpha}_n(t) := \max_{j=0,1,\dots,n-1} \|\alpha_{\sigma(t+j)}(t+j)\|, \qquad t \ge t_0. \tag{5.86}$$

This means that we can rewrite (5.85) as

$$\|\mathcal{Y}(t+n)\| \le c_2 \tilde{\alpha}_n(t) \sum_{j=0}^{n-1} \|\mathcal{Y}(t+j)\| + \\ \sum_{j=0}^{n-1} \Bigg[\Big(1 + \hat{c}_\theta \|\alpha_{\sigma(t+j)}(t+j)\|\Big) |y^*(t+j+1)| + c_1 |\bar{w}(t+j)| \Bigg]. \tag{5.87}$$

It follows from Proposition 5.5 (applied for $p = 1, 2, \dots, n-1$) that there exists a constant c_3 so that the following holds:

$$\sum_{j=0}^{n-1} \|\mathcal{Y}(t+j)\| \le c_3 \|\mathcal{Y}(t)\| + c_3 \sum_{j=0}^{n-2} (|y^*(t+j+1)| + |\bar{w}(t+j)|);$$

after substituting this into (5.87), it follows that there exists a constant c_4 such that

$$\|\mathcal{Y}(t+n)\| \leq c_4 \tilde{\alpha}_n(t) \|\mathcal{Y}(t)\| + c_4 \left(1 + \tilde{\alpha}_n(t)\right) \sum_{j=0}^{n-1} \left(|y^*(t+j+1)| + |\bar{w}(t+j)|\right). \tag{5.88}$$

Step 3: Analyze the switching behavior, and obtain a bound on $\mathcal{Y}(\cdot)$ which depends solely on $J_{i^*}(\cdot)$.

As $\hat{t}_{\ell+1} - \hat{t}_\ell = n$, by changing indexes we can rewrite (5.88) to describe the behavior between two switching times:

$$\|\mathcal{Y}(\hat{t}_{\ell+1})\| \leq c_4 \tilde{\alpha}_n(\hat{t}_\ell) \|\mathcal{Y}(\hat{t}_\ell)\| + c_4 \left(1 + \tilde{\alpha}_n(\hat{t}_\ell)\right) \sum_{j=0}^{n-1} \left(|y^*(\hat{t}_\ell+j+1)| + |\bar{w}(\hat{t}_\ell+j)|\right), \quad \ell \in \mathbb{Z}^+. \tag{5.89}$$

Now we will analyze the $\tilde{\alpha}_n(j)$ term carefully. First, from the definition of α_i given in (5.80), we have

$$\|\alpha_i(t)\| = \rho_i(t) \frac{|e_i(t+1)|}{\|\tilde{\phi}(t)\|}; \tag{5.90}$$

so from the definition of the performance signal in (5.75) and the fact that $\sigma(j)$ is constant for $j \in [\hat{t}_\ell, \hat{t}_{\ell+1})$, we see that

$$\begin{aligned}
\tilde{\alpha}_n(\hat{t}_\ell) &= \max_{q=0,1,\ldots,n-1} \|\alpha_{\sigma(\hat{t}_\ell+q)}(\hat{t}_\ell+q)\| \\
&\leq \sum_{q=0}^{n-1} \|\alpha_{\sigma(\hat{t}_\ell+q)}(\hat{t}_\ell+q)\| \\
&= \sum_{q=0}^{n-1} \rho_{\sigma(\hat{t}_\ell+q)}(\hat{t}_\ell+q) \frac{|e_{\sigma(\hat{t}_\ell+q)}(\hat{t}_\ell+q+1)|}{\|\tilde{\phi}(\hat{t}_\ell+q)\|} \\
&= J_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell), \qquad \ell \in \mathbb{Z}^+.
\end{aligned} \tag{5.91}$$

Incorporating the bound (5.91) into (5.89) yields

$$\|\mathcal{Y}(\hat{t}_{\ell+1})\| \leq c_4 J_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell) \|\mathcal{Y}(\hat{t}_\ell)\| + c_4 \left(1 + J_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\right) \sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+1}-1} \left(|y^*(j+1)| + |\bar{w}(j)|\right), \quad \ell \in \mathbb{Z}^+. \tag{5.92}$$

In the remainder of this step, we will analyze the closed-loop behavior on intervals of length $2n$; we want to obtain a bound on $\mathcal{Y}(\hat{t}_{\ell+2})$ in terms of $\mathcal{Y}(\hat{t}_\ell)$. To proceed, the first goal is to replace the $J_{\sigma(\cdot)}$ term by J_{i^*}: for every $\ell \in \mathbb{Z}^+$, from Lemma 5.1 we have either

$$J_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell) \leq J_{i^*}(\hat{t}_\ell) \tag{5.93}$$

or

$$J_{\sigma(\hat{t}_{\ell+1})}(\hat{t}_{\ell+1}) \leq J_{i^*}(\hat{t}_{\ell+1}). \tag{5.94}$$

So, if (5.93) is true, then we can substitute it into (5.92) to obtain a bound on $\|\mathcal{Y}(\hat{t}_{\ell+1})\|$ in terms of $\|\mathcal{Y}(\hat{t}_\ell)\|$ and then use Proposition 5.5 for n steps to get a bound on $\|\mathcal{Y}(\hat{t}_{\ell+2})\|$ in terms of $\|\mathcal{Y}(\hat{t}_{\ell+1})\|$; it follows that there exists a constant c_5 so that

$$\begin{aligned} \|\mathcal{Y}(\hat{t}_{\ell+2})\| \leq c_5 c_4 J_{i^*}(\hat{t}_\ell)\|\mathcal{Y}(\hat{t}_\ell)\| + c_5 c_4 \left(1 + J_{i^*}(\hat{t}_\ell)\right) \sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+1}-1} \left(|y^*(j+1)| + |\bar{w}(j)|\right) + \\ c_5 \sum_{j=\hat{t}_{\ell+1}}^{\hat{t}_{\ell+2}-1} \left(|y^*(j+1)| + |\bar{w}(j)|\right). \end{aligned} \tag{5.95}$$

On the the other hand, if (5.94) is true, then we can use Proposition 5.5 for n steps to get a bound on $\|\mathcal{Y}(\hat{t}_{\ell+1})\|$ in terms of $\|\mathcal{Y}(\hat{t}_\ell)\|$, and then we can substitute (5.94) into (5.92) to obtain a bound on $\|\mathcal{Y}(\hat{t}_{\ell+2})\|$ in terms of $\|\mathcal{Y}(\hat{t}_{\ell+1})\|$; it follows that there exists a constant c_5 so that

$$\begin{aligned} \|\mathcal{Y}(\hat{t}_{\ell+2})\| \leq c_5 c_4 J_{i^*}(\hat{t}_{\ell+1})\|\mathcal{Y}(\hat{t}_\ell)\| + c_5 c_4 J_{i^*}(\hat{t}_{\ell+1}) \sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+1}-1} \left(|y^*(j+1)| + |\bar{w}(j)|\right) + \\ c_4 \left(1 + J_{i^*}(\hat{t}_{\ell+1})\right) \sum_{j=\hat{t}_{\ell+1}}^{\hat{t}_{\ell+2}-1} \left(|y^*(j+1)| + |\bar{w}(j)|\right). \end{aligned} \tag{5.96}$$

Define

$$\beta_\ell := \max\left\{J_{i^*}(\hat{t}_\ell), J_{i^*}(\hat{t}_{\ell+1})\right\}, \qquad \ell \in \mathbb{Z}^+; \tag{5.97}$$

then there exists a constant c_6 so that (5.95) and (5.96) can be combined to yield

$$\|\mathcal{Y}(\hat{t}_{\ell+2})\| \leq c_6\beta_\ell\|\mathcal{Y}(\hat{t}_\ell)\| + c_6\left(1+\beta_\ell\right)\sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+2}-1}\left(|y^*(j+1)|+|\bar{w}(j)|\right), \qquad \ell \in \mathbb{Z}^+. \tag{5.98}$$

Before proceeding, for ease of analysis we further rewrite the above as follows:

$$\|\mathcal{Y}(\hat{t}_{2j+2})\| \leq c_6\beta_{2j}\|\mathcal{Y}(\hat{t}_{2j})\| + c_6\left(1+\beta_{2j}\right)\underbrace{\sum_{k=\hat{t}_{2j}}^{\hat{t}_{2j+2}-1}\left(|y^*(k+1)|+|\bar{w}(k)|\right)}_{=:\tilde{w}(j)}, \quad j \in \mathbb{Z}^+. \tag{5.99}$$

Step 4: Solve the key difference inequality (5.99).

Now we proceed to analyze the closed-loop behavior on the whole timeline. We define

$$\lambda_1 = \frac{\lambda^{2n}}{\max\{1, c_6\}} \in (0,1)$$

and

$$\nu := \frac{\lambda_1^2}{16n^2\bar{c}} \quad \in (0,1). \tag{5.100}$$

To proceed, let $\tau \geq t_0$ be arbitrary. We now partition the timeline into two parts: one in which $\bar{w}(\cdot)$ is small versus $\tilde{\phi}(\cdot)$ and one where it is not. With ν defined above, we define

$$S_{good} = \left\{ j \geq \tau : \tilde{\phi}(j) \neq 0 \text{ and } \frac{|\bar{w}(j)|^2}{\|\tilde{\phi}(j)\|^2} < \nu \right\},$$

$$S_{bad} = \left\{ j \geq \tau : \tilde{\phi}(j) = 0 \text{ or } \frac{|\bar{w}(j)|^2}{\|\tilde{\phi}(j)\|^2} \geq \nu \right\};$$

clearly $\{j \in \mathbb{Z} : \; j \geq \tau\} = S_{good} \cup S_{bad}$. We can clearly define a (possibly infinite) sequence of intervals of the form $[k_l, k_{l+1})$ which satisfy:
(i) $k_0 = \tau$ serves as the initial instant of the first interval;
(ii) $[k_l, k_{l+1})$ either belongs to S_{good} or S_{bad}; and
(iii) if $k_{l+1} \neq \infty$ and $[k_l, k_{l+1})$ belongs to S_{good} then $[k_{l+1}, k_{l+2})$ belongs to S_{bad} and vice versa.

Now we analyze the behavior during each interval.

Case 1: $[k_l, k_{l+1})$ lies in S_{bad}.

In this case, the analysis is identical to Case 1 inside the proof of Theorem 5.1; we conclude that there exists a constant $c_7 := \frac{\|\mathcal{S}^*\|}{\sqrt{\nu}} + (1 + c_\varphi\|\mathcal{S}^*\|)\,(n+2)$ so that following holds:

$$\|\mathcal{Y}(j)\| \leq \begin{cases} c_7|\bar{w}(j)|, & j = k_l, k_l+1, \ldots, k_{l+1}-1 \\ \left(\frac{1}{\sqrt{\nu}} + c_7\right)|\bar{w}(j-1)|, & j = k_{l+1}. \end{cases} \tag{5.101}$$

Case 2: $[k_l, k_{l+1})$ lies in S_{good}.

First suppose that $k_{l+1} - k_l \leq 4n$; then by Proposition 5.5 it can be easily proven that there exists a constant γ_1 so that

$$\|\mathcal{Y}(t)\| \leq \gamma_1\lambda^{t-k_l}\|\mathcal{Y}(k_l)\| + \gamma_1\sum_{j=k_l}^{t-1}\lambda^{t-j-1}\left(|y^*(j+1)| + |\bar{w}(j)|\right), \quad t \in [k_l, k_{l+1}]. \tag{5.102}$$

Now suppose that $k_{l+1} - k_l > 4n$; this means that in the interval of interest, namely $[k_l, k_{l+1})$, there are at least two switching times: there exist $0 \leq q < p$ so that

$$k_l \leq \hat{t}_q < \hat{t}_p \leq k_{l+1};$$

in fact, there may be many choices of q and p. For $j \in [k_l, k_{l+1})$, we have $\tilde{\phi}(j) \neq 0$ and

$$\frac{\bar{w}(j)^2}{\|\tilde{\phi}(j)\|^2} < \nu. \tag{5.103}$$

By Proposition 5.4, it follows that for every $t_2 > t_1 \geq t_0$:

$$\sum_{j=t_1}^{t_2-1}\rho_{i^*}(j)\frac{|e_{i^*}(j+1)|^2}{\|\tilde{\phi}(j)\|^2} \leq 2\|\tilde{\theta}_{i^*}(t_1)\|^2 + 2\bar{c}\sum_{j=t_1}^{t_2-1}\rho_{i^*}(j)\frac{|\bar{w}(j)|^2}{\|\tilde{\phi}(j)\|^2}; \tag{5.104}$$

from this and using the bound in (5.103) which holds on $[k_l, k_{l+1})$, together with the fact that $\|\tilde{\theta}_{i^*}(t)\| \leq 2\|\mathcal{S}_{i^*}\| \leq 2\bar{\mathbf{s}}$, we see that

$$\sum_{j=t_1}^{t_2-1}\rho_{i^*}(j)\frac{|e_{i^*}(j+1)|^2}{\|\tilde{\phi}(j)\|^2} \leq 8\bar{\mathbf{s}}^2 + 2\bar{c}\nu\,(t_2 - t_1)\,, \qquad k_l \leq t_1 < t_2 \leq k_{l+1}. \tag{5.105}$$

We now analyze the square sum of β_{2j}; from the definition of β_{2j} and the Cauchy-Schwarz property, we have

$$\begin{aligned}
\sum_{j=q}^{p-1}\beta_{2j}^2 &= \sum_{j=q}^{p-1}\max\left\{J_{i^*}(\hat{t}_{2j}), J_{i^*}(\hat{t}_{2j+1})\right\}^2 \\
&\le \sum_{j=q}^{p-1}\left(J_{i^*}(\hat{t}_{2j}) + J_{i^*}(\hat{t}_{2j+1})\right)^2 \\
&\le 2\sum_{j=q}^{p-1}\left[J_{i^*}(\hat{t}_{2j})^2 + J_{i^*}(\hat{t}_{2j+1})^2\right] \\
&= 2\sum_{j=q}^{p-1}\left[\left(\sum_{k=\hat{t}_{2j}}^{\hat{t}_{2j+1}-1}\rho_{i^*}(k)\frac{|e_{i^*}(k+1)|}{\|\tilde{\phi}(k)\|}\right)^2 + \left(\sum_{k=\hat{t}_{2j+1}}^{\hat{t}_{2j+2}-1}\rho_{i^*}(k)\frac{|e_{i^*}(k+1)|}{\|\tilde{\phi}(k)\|}\right)^2\right] \\
&\le 2n\sum_{j=q}^{p-1}\left(\sum_{k=\hat{t}_{2j}}^{\hat{t}_{2j+2}-1}\rho_{i^*}(k)\frac{|e_{i^*}(k+1)|^2}{\|\tilde{\phi}(k)\|^2}\right) \\
&= 2n\sum_{j=\hat{t}_{2q}}^{\hat{t}_{2p}-1}\rho_{i^*}(j)\frac{|e_{i^*}(j+1)|^2}{\|\tilde{\phi}(j)\|^2}, \qquad 0 \le q < p. \qquad (5.106)
\end{aligned}$$

Applying (5.105) into the above yields

$$\begin{aligned}
\sum_{j=q}^{p-1}\beta_{2j}^2 &\le 2n\sum_{j=\hat{t}_{2q}}^{\hat{t}_{2p}-1}\rho_{i^*}(j)\frac{|e_{i^*}(j+1)|^2}{\|\phi(j)\|^2} \\
&\le 16n\bar{\mathbf{s}}^2 + 4\bar{c}\nu n\left(\hat{t}_{2p} - \hat{t}_{2q}\right), \\
&= 16n\bar{\mathbf{s}}^2 + 8\bar{c}\nu n^2\left(p - q\right), \\
&\quad \text{for all } q, p \in \mathbb{Z}^+ \text{ s.t. } k_l \le \hat{t}_{2q} < \hat{t}_{2p} \le k_{l+1}. \qquad (5.107)
\end{aligned}$$

Before proceeding, we would like to obtain an upper bound on the β_{2j} term that appears in the second term on the RHS of (5.99). From the above we see that

$$\beta_{2j} \le \left(16n\bar{\mathbf{s}}^2 + 8\bar{c}\nu n^2\right)^{\frac{1}{2}} =: c_8;$$

so we can rewrite (5.99) as

$$\|\mathcal{Y}(\hat{t}_{2j+2})\| \leq c_6\beta_{2j}\|\mathcal{Y}(\hat{t}_{2j})\| + \underbrace{c_6\,(1+c_8)}_{=:c_9}\tilde{w}(j). \tag{5.108}$$

We now utilize the inequality of arithmetic and geometric means.

Claim 5.5. *There exists a constant $\gamma_2 > 1$ such that*

$$\prod_{j=q}^{p-1}\beta_{2j} \leq \gamma_2\lambda_1^{p-q}, \qquad \textit{for all } q,p \in \mathbb{Z}^+ \textit{ s.t. } k_l \leq \hat{t}_{2q} < \hat{t}_{2p} \leq k_{l+1}. \tag{5.109}$$

Proof of Claim 5.5. Let $q,p \in \mathbb{Z}^+$ be arbitrary such that $k_l \leq \hat{t}_{2q} < \hat{t}_{2p} \leq k_{l+1}$. By the fact that $\beta_{2j} \geq 0$, we obtain

$$\prod_{j=q}^{p-1}\beta_{2j} \leq \left[\frac{1}{p-q}\sum_{j=q}^{p-1}\beta_{2j}^2\right]^{\frac{p-q}{2}}. \tag{5.110}$$

Substituting (5.107) and the definition of ν in (5.100) into (5.110) yields

$$\begin{aligned}\prod_{j=q}^{p-1}\beta_{2j} &\leq \left[\frac{16n\bar{\mathbf{s}}^2}{p-q} + 8\bar{c}\nu n^2\right]^{\frac{p-q}{2}} \\ &\leq \left[\frac{16n\bar{\mathbf{s}}^2}{p-q} + \frac{\lambda_1^2}{2}\right]^{\frac{p-q}{2}}.\end{aligned} \tag{5.111}$$

So it is enough to prove that there exists a constant γ_2 so that

$$\left(\underbrace{\left[\frac{16n\bar{\mathbf{s}}^2}{k} + \frac{\lambda_1^2}{2}\right]^{\frac{1}{2}}}_{=:g(k)}\right)^k \leq \gamma_2\lambda_1^k, \qquad k > 0.$$

We can easily show that with $\bar{k} := 32n \times \left\lceil\left(\frac{\bar{\mathbf{s}}}{\lambda_1}\right)^2\right\rceil$, we have

$$\frac{16n\bar{\mathbf{s}}^2}{\bar{k}} \leq \frac{\lambda_1^2}{2},$$

which means that

$$g(k)^k \leq \lambda_1^k, \qquad k \geq \bar{k}.$$

Since $g(k)$ decreases as $k \geq 1$ increases, we conclude that if we define $\gamma_2 := \max\left\{1, \left(\frac{g(1)}{\lambda_1}\right)^{\bar{k}}\right\}$, then

$$g(k)^k \leq \gamma_2 \lambda_1^k, \quad k = 1, 2, \ldots, \bar{k},$$

as well, so the claim holds. □

We now solve (5.108). First of all, using the bound in (5.109) and the definition of λ_1 we obtain

$$\begin{aligned}\prod_{j=q}^{p-1} [c_6 \beta_{2j}] &\leq \gamma_2 \lambda_1^{p-q} c_6^{p-q} \\ &\leq \gamma_2 \lambda^{2n(p-q)}, \qquad \text{for all } q, p \in \mathbb{Z}^+ \text{ s.t. } k_l \leq \hat{t}_{2q} < \hat{t}_{2p} \leq k_{l+1}.\end{aligned} \tag{5.112}$$

If we solve (5.108) iteratively and apply the above inequality, we obtain

$$\begin{aligned}\|\mathcal{Y}(\hat{t}_{2p})\| \leq \gamma_2 \lambda^{2n(p-q)} \|\mathcal{Y}(\hat{t}_{2q})\| + \sum_{j=q}^{p-1} c_9 \gamma_2 \left(\lambda^{2n}\right)^{p-j-1} \tilde{w}(j), \\ \text{for all } q, p \in \mathbb{Z}^+ \text{ s.t. } k_l \leq \hat{t}_{2q} < \hat{t}_{2p} \leq k_{l+1}.\end{aligned} \tag{5.113}$$

We can now use Proposition 5.5 (for no more than $2n$ steps at a time):

- to provide a bound on $\|\mathcal{Y}(t)\|$ between the switching times $\hat{t}_{2j}$'s;
- to provide a bound on $\|\mathcal{Y}(t)\|$ on the beginning part of the interval $[k_l, k_{l+1})$, until we get to the first admissible switching time;
- to provide a bound on $\|\mathcal{Y}(t)\|$ on the last part of the interval $[k_l, k_{l+1})$, after the last admissible switching time.

After simplification, we conclude that there exists a constant $\gamma_3 \geq \gamma_1$ so that

$$\|\mathcal{Y}(t)\| \leq \gamma_3 \lambda^{t-k_l} \|\mathcal{Y}(k_l)\| + \gamma_3 \sum_{j=k_l}^{t-1} \lambda^{t-j-1} \left(|y^*(j+1)| + |\bar{w}(j)|\right), \quad t \in [k_l, k_{l+1}], \tag{5.114}$$

as desired.

Now we combine Case 1 and Case 2 into a general bound on $\mathcal{Y}$. We use the same argument that of Claim 5.3 inside the proof of Theorem 5.1 to glue the bounds of Case 1 and Case 2 together: we conclude that there exists a constant $\bar{\gamma}$ so that

$$\|\mathcal{Y}(t)\| \leq \bar{\gamma}\lambda^{t-\tau}\|\mathcal{Y}(\tau)\| + \sum_{j=\tau}^{t-1} \bar{\gamma}\lambda^{t-j-1}(|y^*(j+1)| + |\bar{w}(j)|), \qquad t > \tau. \tag{5.115}$$

Step 5: Obtain a bound on $u(t)$.

Now we derive a bound on $u(t)$ in a similar manner to that of Step 4 inside the proof of Theorem 5.1; we see that

$$|u(t)| \leq \hat{c}_\theta c_\varphi \bar{\mathbf{s}}\bar{\gamma}\lambda^{t-\tau}\|\mathcal{Y}(\tau)\| + \hat{c}_\theta c_\varphi \bar{\mathbf{s}}\bar{\gamma} \sum_{j=\tau}^{t-1} \lambda^{t-j-1}(|y^*(j+1)| + |\bar{w}(j)|) + \hat{c}_\theta |y^*(t+1)|, \;\; t \geq \tau. \tag{5.116}$$

If we combine this with (5.115) we conclude that there exists a constant $\bar{\gamma}_2$ so that

$$\left\| \begin{bmatrix} \mathcal{Y}(t) \\ u(t) \end{bmatrix} \right\| \leq \bar{\gamma}_2 \lambda^{t-\tau}\|\mathcal{Y}(\tau)\| + \\ \sum_{j=\tau}^{t-1} \bar{\gamma}_2 \lambda^{t-j-1}(|y^*(j+1)| + |\bar{w}(j)|) + \bar{\gamma}_2 |y^*(t+1)|, \;\; t \geq \tau. \tag{5.117}$$

Since $\tau \geq t_0$ is arbitrary, we conclude the proof. ■

Remark 5.10. *Observe that the result in Theorem 5.6 subsumes the result in Theorem 4.1 of Chapter 4 dealing with one-step-ahead adaptive control of first-order linear plants.*

5.4.5 Robustness Results

Similar to the case of a known plant's input gain, the linear-like convolution bounds proven in Theorem 5.6 will guarantee robustness to a degree of time-variations and unmodelled dynamics in this case of an unknown plant's input gain and using switching control law. We have in (5.63) a time-varying version of the plant (5.5) along with the term $d_\Delta(t) \in \mathbb{R}$

added to represent the unmodelled dynamics; for clarity we present it here again:

$$\begin{aligned} y(t+1) &= \theta^*(t)^\top \begin{bmatrix} \varphi\big(\mathcal{X}(j)\big) \\ u(t) \end{bmatrix} + w(t) + v(t+1) + d_\Delta(t) \\ &= \theta^*(t)^\top \phi(t) + w(t) + v(t+1) + \underbrace{\theta^*(t)^\top \begin{bmatrix} \varphi\big(\mathcal{X}(t)\big) - \varphi\big(\mathcal{Y}(t)\big) \\ 0 \end{bmatrix}}_{=:\tilde{\Delta}(t)} + d_\Delta(t). \end{aligned} \tag{5.118}$$

Similar to the previous section, as discussed in Chapter 2, with $g : \mathbb{R}^{n+1} \mapsto \mathbb{R}$ having a bounded gain, we assume that d_Δ satisfies

$$m(t+1) = \beta m(t) + \beta \left| g\left(\begin{bmatrix} \mathcal{Y}(t) \\ u(t) \end{bmatrix} \right) \right|, \qquad m(t_0) = m_0 \tag{5.119a}$$

$$|d_\Delta(t)| \le \mu m(t) + \mu \left| g\left(\begin{bmatrix} \mathcal{Y}(t) \\ u(t) \end{bmatrix} \right) \right|, \qquad t \ge t_0. \tag{5.119b}$$

Theorem 5.7. *Suppose that the adaptive controller (5.70), (5.72), (5.74), (5.75) and (5.76) is applied to the time-varying nonlinear plant (5.118) with d_Δ satisfying (5.119). Then for every $\delta \in (0, \infty]$, $\beta \in (0,1)$ and $\bar{c}_0 \ge 0$, there exist $\bar{\epsilon} > 0$, $\mu > 0$, $\tilde{\lambda} \in (\beta, 1)$ and $\tilde{\gamma} > 0$ such that for every $t_0 \in \mathbb{Z}$, $\mathcal{X}(t_0) \in \mathbb{R}^n$, $\sigma_0 \in \mathcal{I}^*$, $\theta^* \in \mathbb{S}(\mathcal{S}^*, \bar{c}_0, \bar{\epsilon})$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ $(i = 1, 2)$, and $w, v, y^* \in \boldsymbol{\ell}_\infty$, the following holds:*

$$\left\| \begin{bmatrix} \mathcal{Y}(t) \\ u(t) \\ m(t) \end{bmatrix} \right\| \le \tilde{\gamma} \tilde{\lambda}^{t-t_0} \left\| \begin{bmatrix} \mathcal{Y}(t_0) \\ u(t_0) \\ m_0 \end{bmatrix} \right\| + \sum_{j=t_0}^{t-1} \tilde{\gamma} \tilde{\lambda}^{t-j-1} (|\bar{w}(j)| + |y^*(j+1)|) + \tilde{\gamma} |y^*(t+1)|, \qquad t \ge t_0.$$

Proof of Theorem 5.7. We observe here that the plant (5.118) and the controller consisting of (5.70), (5.72), (5.74), (5.75) and (5.76) fit into the paradigm of Chapter 2: we set

$$\vartheta(t) = \begin{bmatrix} \mathcal{Y}(t) \\ u(t) \end{bmatrix},$$

$$
\begin{aligned}
f(\vartheta(\cdot)) = \phi(\cdot) &= \begin{bmatrix} \varphi\big(\mathcal{Y}(\cdot)\big) \\ u(\cdot) \end{bmatrix}, \\
z(t) &= \varnothing, \\
\hat{\theta}(t) &= \begin{bmatrix} \hat{\theta}_1(t) \\ \hat{\theta}_2(t) \end{bmatrix}, \\
r(t) &= y^*(t+1), \\
w(t) &= w(t) + v(t+1) + \tilde{\Delta}(t), \\
\Omega &= \mathcal{S}_1 \times \mathcal{S}_2.
\end{aligned}
$$

In Theorem 5.6 it is proven the controller (5.70), (5.72), (5.74), (5.75) and (5.76) provides a convolution bound for (5.5). Observe that the term $\tilde{\Delta}(t)$ has the same upper bound of as its counterpart in (5.10) because $\theta^*(t) \in \mathcal{S}$ for all $t \in \mathbb{Z}$; indeed, we utilize the upper bound given in (5.11). By the definition of $\bar{w}$ given in (5.12), it means that there exists a constant c so that

$$
|w(t) + v(t+1) + \tilde{\Delta}(t)| \leq c|\bar{w}(t)|.
$$

We apply Theorems 2.2, 2.2 and 2.3 to immediately see that the linear-like convolution bound holds in the presence of parameter time-variation and/or unmodelled dynamics. ■

Remark 5.11. *Just as in the first-order LTI case of Chapter 4, in the absence of noise, we are unable to prove that we obtain asymptotic tracking; however, if switching eventually stops then this will be the case. We may be able to alleviate this problem if we apply the switching algorithm using a performance signal which has some memory.*

5.5 Simulation Examples

5.5.1 The One-link Manipulator

Here we apply the proposed approach to a one-link manipulator system discussed in Example 5.1. There we provide a continuous-time model (5.3), together with a discretized approximation (5.4); with a sampling period of $h = 0.1$ seconds, the latter is given by

$$
y(t+1) = \theta_1 y(t) + \theta_2 \sin\big(y(t)\big) + \theta_3 y(t-1) + \theta_4 u(t), \tag{5.120}
$$

with

$$\begin{bmatrix}\theta_1\\\theta_2\\\theta_3\\\theta_4\end{bmatrix} = \begin{bmatrix}2\\-\frac{[m+2M]gLh^2}{J+4ML^2}\\-1\\\frac{h^2}{J+4ML^2}\end{bmatrix} \qquad \text{and} \qquad \varphi\Big(y(t), y(t-1)\Big) = \begin{bmatrix}y(t)\\\sin\Big(y(t)\Big)\\y(t-1)\end{bmatrix}.$$

Here we allow the time-varying mass load M to take values between 0 and 5; we set other parameters to be: $m = J = 0.5$ and $L = 1$. This means that the uncertainty set of the system is as follows:

$$\mathcal{S} = \mathcal{S}^* = \left\{\begin{bmatrix}\theta_1\\\theta_2\\\theta_3\\\theta_4\end{bmatrix} \in \mathbb{R}^4 : \theta_1 = 2, \theta_2 \in [-1, -0.0502], \theta_3 = -1, \theta_4 \in [0.0005, 0.02]\right\};$$

notice that the estimation of the parameters θ_1 and θ_3 is trivial.

We illustrate the result in Theorem 5.4 on tracking in the presence of time-variation; we show that the average tracking error is proportional to the speed of the parameter variation. We apply the proposed controller (5.16) and (5.21) (with $\delta = \infty$) to the plant when the mass load is varying between 0 and 2 as follows:

$$M(t) = 1 + \sin(\omega_0 t)$$

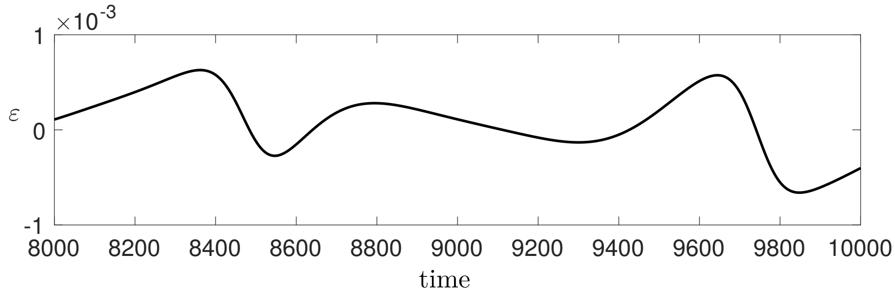
with

$$\omega_0 \in \{0.005, 0.01, 0.05, 0.1\}.$$

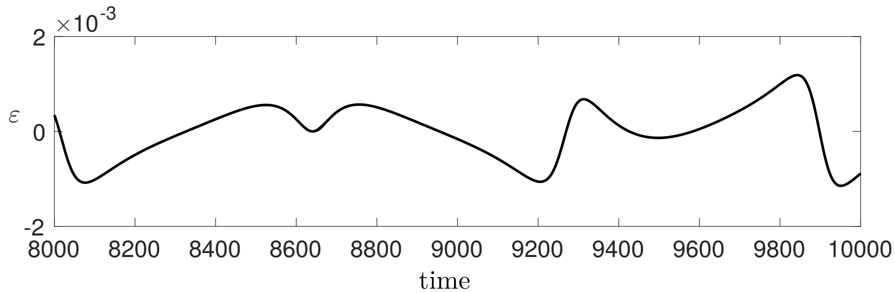
We choose plant initial conditions of $y(-1) = y(0) = \pi$, and the initial parameter estimate corresponding to a mass load of 2; this corresponds to values of $\hat{\theta}_2(0) = -0.0678$ and $\hat{\theta}_4(0) = 0.0077$. With a zero disturbance and

$$y^*(t) = \frac{\pi}{2}\sin\left(\tfrac{1}{250}t\right),$$

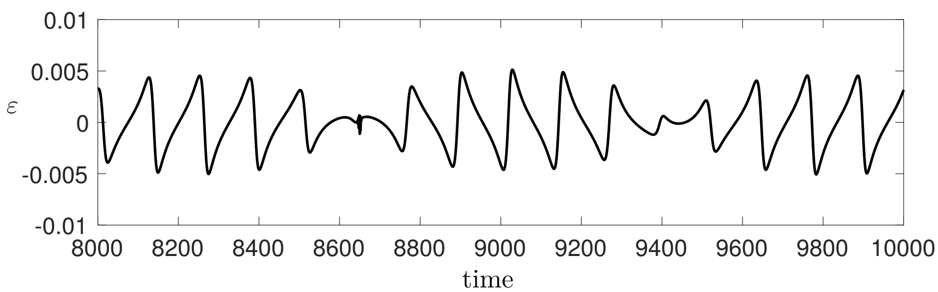
we simulate the closed-loop system for $T = 10000$ steps, i.e. 1000 seconds; we plot the tracking error for the last 2000 steps in Figure 5.2, i.e. after the transient effect has been eliminated. We see that, consistent with Theorem 5.4, average tracking error increases with the speed of the plant parameter variation. In Figure 5.3 and Figure 5.4 we show the details of the closed-loop behavior for the case of $\omega_0 = 0.005$; although the system exhibits a large transient control action, stability is retained.



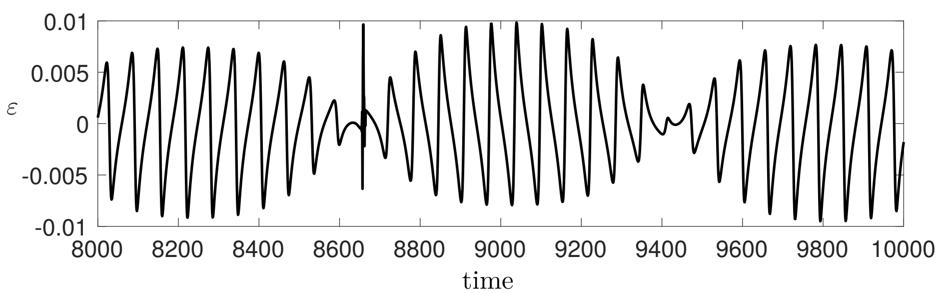
$(\omega_0 = 0.005)$



$(\omega_0 = 0.01)$



$(\omega_0 = 0.05)$



$(\omega_0 = 0.1)$

Figure 5.2: The plots show the tracking error for simulation of a one-link manipulator for various choices of the time-varying parameter.

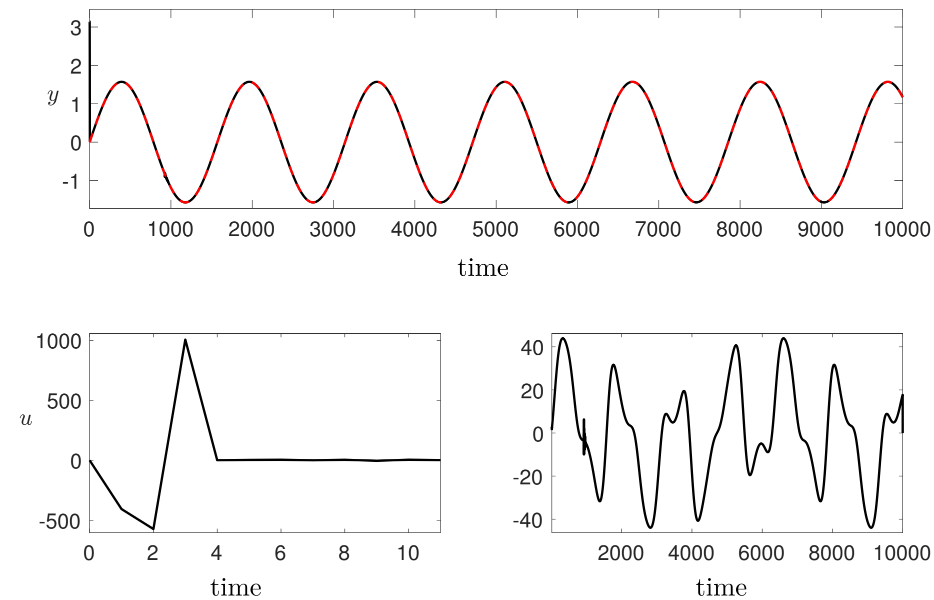


Figure 5.3: The top plot shows both the output $y(t)$ (solid black) and the reference $y^*(t)$ (dashed red); the bottom left plot shows the transient control input $u(t)$ of the first 10 steps, and the bottom right plot shows the control input $u(t)$ for $t \geq 10$.

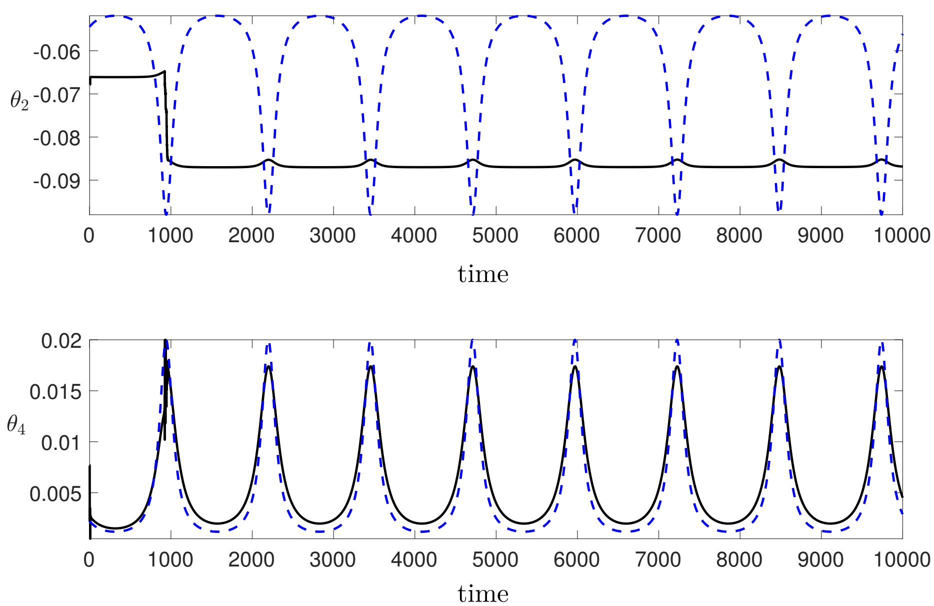


Figure 5.4: The two plots shows both the parameter estimate $\hat{\theta}_2(t)$ and $\hat{\theta}_4(t)$ (solid black) and the actual parameters $\theta_2(t)$ and $\theta_4(t)$ (dashed blue), respectively.

5.5.2 An Example with an Unknown θ_p Sign

Here we consider an example to illustrate the switching controller of Section 5.4. To this end, consider the following second-order nonlinear plant:

$$\begin{aligned} x(t+1) &= \theta_1(t)x(t) + \theta_2(t)\sin\Big(10x(t)\Big) + \theta_3(t)|x(t-1)| + \theta_4(t)u(t) + w(t) \\ y(t) &= x(t) + v(t), \end{aligned}$$

with the unknown parameters belonging to the compact set

$$\mathcal{S}^* = \left\{ \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \end{bmatrix} \in \mathbb{R}^4 : \theta_1 \in [1,3], \theta_2 \in [-2,2], \theta_3 \in [-4,-2], \theta_4 \in [1,3] \cup [-3,-1] \right\};$$

observe that this set satisfies Assumptions 5.1. Here we have

$$\varphi\Big(x(t), x(t-1)\Big) = \begin{bmatrix} x(t) \\ \sin\Big(10x(t-1)\Big) \\ |x(t-1)| \end{bmatrix}$$

which clearly satisfies Assumption 5.2. To apply our approach, we will need compact and convex sets $\mathcal{S}_1$ and $\mathcal{S}_2$ so that their union contains $\mathcal{S}^*$; there is a natural choice: define

$$\mathcal{S}_1 = \left\{ \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \end{bmatrix} \in \mathbb{R}^4 : \theta_1 \in [1,3], \theta_2 \in [-2,2], \theta_3 \in [-4,-2], \theta_4 \in [1,3] \right\},$$

$$\mathcal{S}_2 = \left\{ \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \end{bmatrix} \in \mathbb{R}^4 : \theta_1 \in [1,3], \theta_2 \in [-2,2], \theta_3 \in [-4,-2], \theta_4 \in [-3,-1] \right\}.$$

Now we will carry out a simulation with time-varying parameters:

$$\begin{aligned} \theta_1(t) &= 2 + \cos\left(\tfrac{1}{300}t\right), & \theta_2(t) &= \cos\left(\tfrac{1}{250}t\right), \\ \theta_3(t) &= -2, & \theta_4(t) &= \begin{cases} -1 & 2000 < t \leq 3500 \\ 1 & \text{otherwise;} \end{cases} \end{aligned}$$

we set the noise terms to

$$w(t) = \begin{cases} \frac{1}{2}\sin(10t), & 1000 < t \leq 2500 \\ 0 & \text{otherwise,} \end{cases}$$

$$v(t) = \begin{cases} \frac{1}{5}\cos(20t), & 1000 < t \leq 2500 \\ 0 & \text{otherwise.} \end{cases}$$

We apply the switching adaptive controller (5.70), (5.72), (5.74), (5.75) and (5.76) (with $\delta = 2\bar{\mathbf{s}} \approx 6$) with initial parameter estimates set to $\theta_1(0) = \begin{bmatrix} 2 & 0 & -3 & 3 \end{bmatrix}^\top$, and $\theta_2(0) = \begin{bmatrix} 2 & 0 & -3 & -3 \end{bmatrix}^\top$ and $\sigma_0 = 2$. We set the reference signal to $y^*(t) = 20\cos\left(\frac{2\pi}{750}t\right)$ and set initial condition to $x(0) = x(-1) = 11$. We plot the results in Figures 5.5 and 5.6. We see that the closed-loop system provides a good job of tracking despite inaccurate parameter estimation. While performance degrades temporarily when measurement noise is turned on or when the sign of θ_4 changes, it improves quickly thereafter. Furthermore, the estimate choice is mostly correct.

Remark 5.12. *We can also compare the performance here with that which arises when we use the classical estimator* (3.4) *as part of the adaptive controller; we often end up with the same sort of result as in Example 2 of the simulation section of Chapter 4, namely a degradation in performance.*

5.6 Conclusion

We have shown in this chapter that we are able to obtain desirable linear-like closed-loop behavior for a class of nonlinear systems. The set of parametric uncertainty is only assumed to be compact. The first half of the chapter discussed the approach when the sign of the control/input gain is known. The second half of the chapter deals with the case of sign of the control/input gain is unknown: multiple estimators, together with a switching control law, is used. Utilizing the resulting convolution bounds, we have shown that the adaptive controller tolerates a certain degree of time-variations and unmodelled dynamics. In the case of known sign of the control gain, some tracking results are also proven which are stronger than what is usually found in the literature. This work can be considered as a first step towards obtaining the desired linear-like properties when applying adaptive control to discrete-time nonlinear systems; further study of the adaptive control of more general discrete-time nonlinear systems is merited.

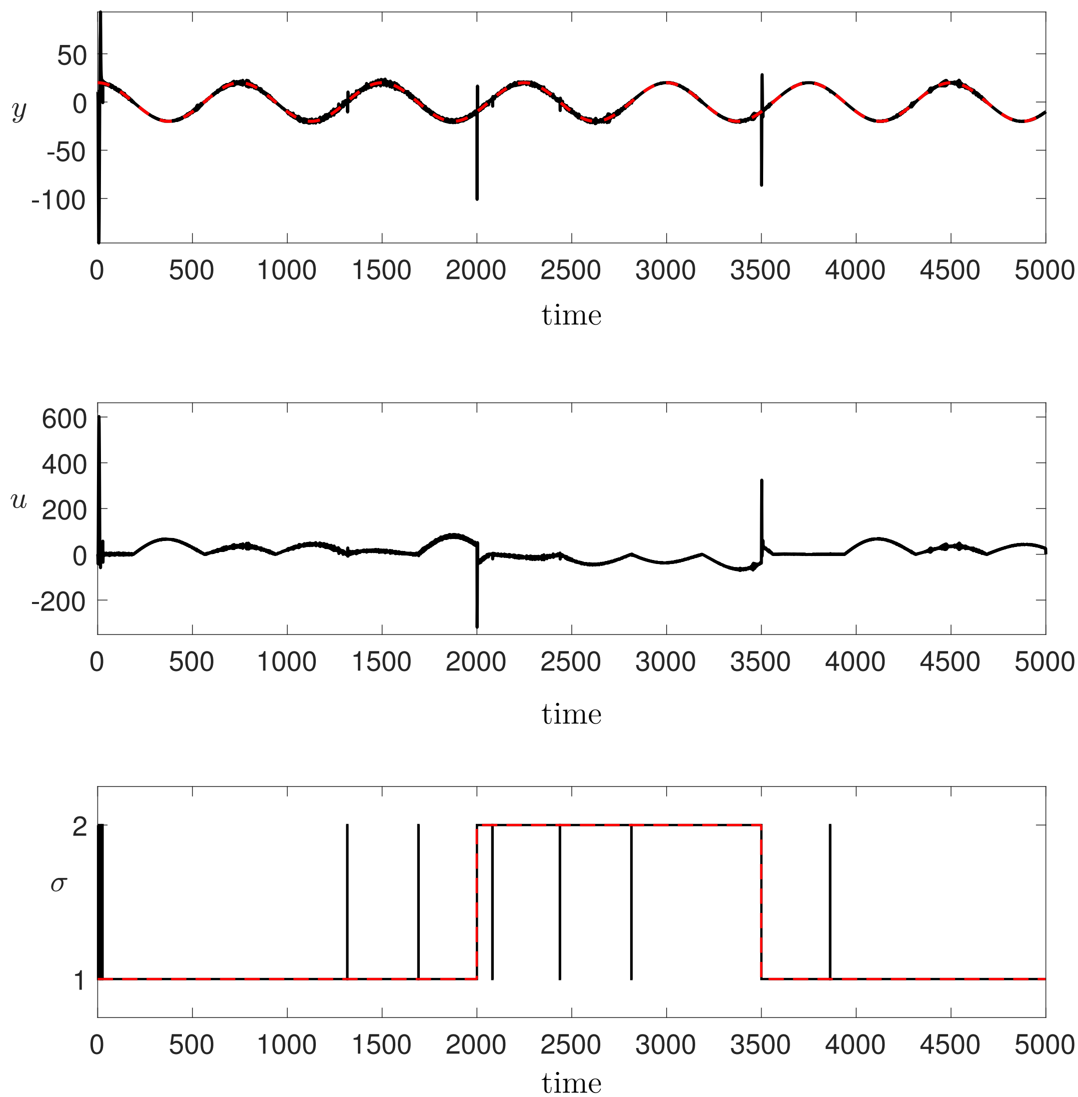


Figure 5.5: The top plot shows both the output $y(t)$ (solid black) and the reference $y^*(t)$ (dashed red); the 2nd plot shows the control input $u(t)$; the bottom plot shows both the switching signal $\sigma(t)$ (solid black) and correct index $i^*(t)$ (dashed red).

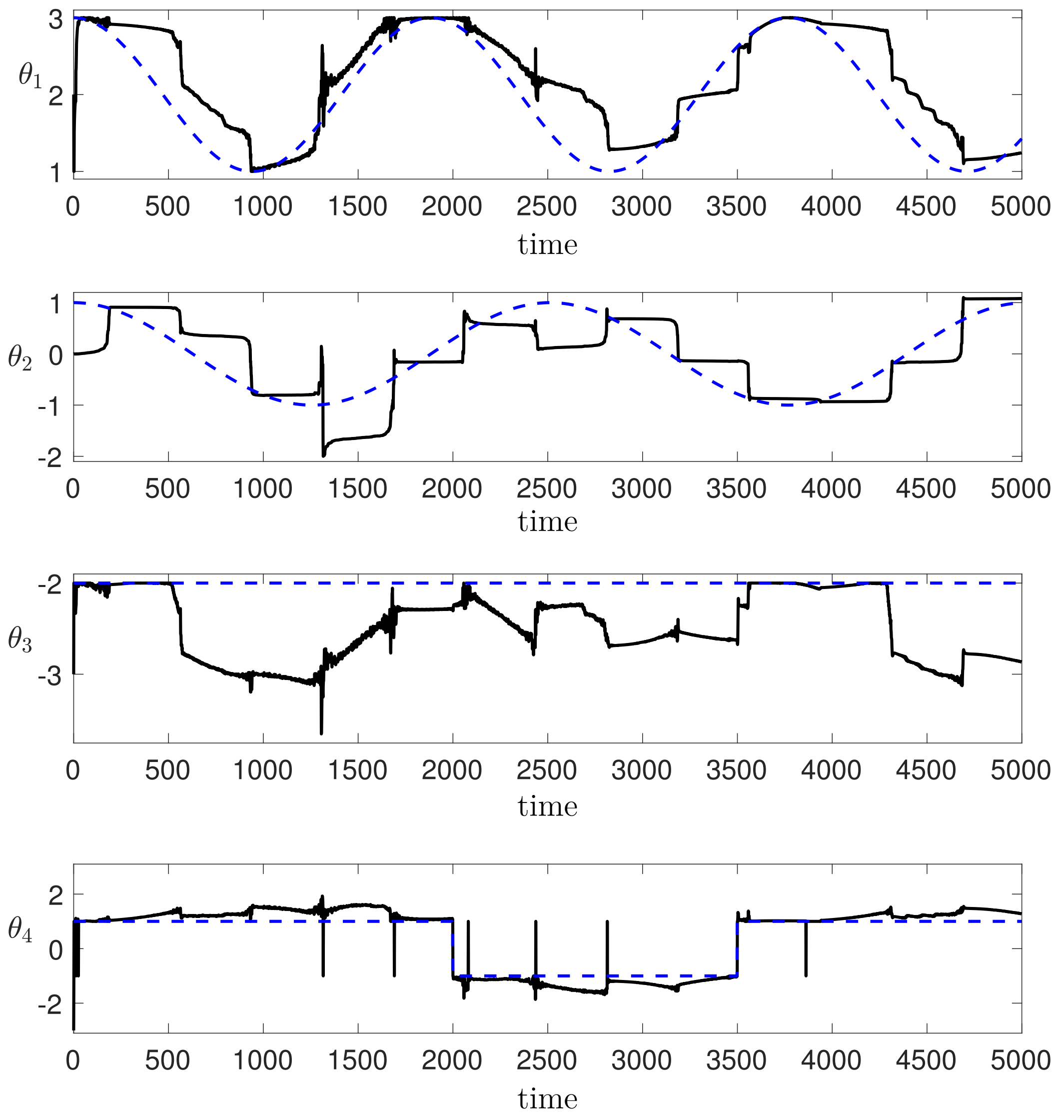


Figure 5.6: The plots show the parameter estimates $\hat{\theta}_{\sigma(t)}$ (solid) as well as the actual parameters $\theta^*(t)$ (dashed).

Chapter 6

Adaptive Stabilization of Possibly Non-Minimum Phase Plants Using Two Estimators

6.1 Introduction

We have shown in Chapter 4 in the first-order LTI case that a switching adaptive controller utilizing projection-algorithm estimators stabilizes the system and provide a linear-like closed-loop behavior; the same is proven in Chapter 5 for a special class of nonlinear systems with stable zero dynamics. In this chapter, we extend the approach to higher order LTI plants that may be non-minimum phase. We adopt a pole-placement based adaptive controller to achieve the desirable linear-like properties.

6.2 The Setup

Here, we consider an n^{th}-order linear time-invariant discrete-time plant given by

$$y(t+1) = \sum_{j=1}^{n} a_j y(t-j+1) + \sum_{j=1}^{n} b_j u(t-j+1) + w(t)$$

$$= \underbrace{\begin{bmatrix} y(t) \\ y(t-1) \\ \vdots \\ y(t-n+1) \\ u(t) \\ u(t-1) \\ \vdots \\ u(t-n+1) \end{bmatrix}^{\top}}_{=:\phi(t)^{\top}} \underbrace{\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \\ b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}}_{=:\theta^*} + w(t), \tag{6.1}$$

with $y(t) \in \mathbb{R}$ the measured output, $u(t) \in \mathbb{R}$ the control input, $\phi(t) \in \mathbb{R}^{2n}$ a vector of input-output data, and $w(t) \in \mathbb{R}$ the noise (or disturbance) input. We assume that θ^* is unknown but belongs to a set $\mathcal{S}^* \subset \mathbb{R}^{2n}$. Associated with this plant model are the polynomials

$$\mathbf{A}(z^{-1}) := 1 - a_1 z^{-1} - a_2 z^{-2} - \cdots - a_n z^{-n}$$

and

$$\mathbf{B}(z^{-1}) := b_1 z^{-1} + b_2 z^{-2} + \cdots + b_n z^{-n}$$

and the transfer function $\frac{\mathbf{B}(z^{-1})}{\mathbf{A}(z^{-1})}$.

Remark 6.1. *It is straight-forward to verify that if the system has a disturbance at both the input and output, then it can be converted to a system of the above form. To see this, suppose that we start with the model*

$$\begin{aligned} x(t+1) &= \sum_{j=1}^{n} a_j x(t-j+1) + \sum_{j=1}^{n} b_j u(t-j+1) \\ y(t) &= x(t) + w_1(t) \\ u(t) &= u_c(t) + w_2(t); \end{aligned}$$

here $y(t)$ is the measured output, $u_c(t)$ is the input generated by the controller, $w_1(t)$ is an output disturbance/noise and $w_2(t)$ is an input disturbance/noise. This can be rewritten as

$$y(t+1) = \sum_{j=1}^{n} a_j y(t-j+1) + \sum_{j=1}^{n} b_j u_c(t-j+1) +$$

$$\underbrace{w_1(t+1) - \sum_{j=1}^{n} a_j w_1(t-j+1) + \sum_{j=1}^{n} b_j w_2(t-j+1)}_{=:w(t)},$$

which is exactly of the form (6.1)*. Nonlinear plants generally do not enjoy this property.*

It has been shown in the main result of Miller and Shahab [46] on adaptive pole-placement control that convexity and coprimeness assumptions on the set of admissible plant parameters play a crucial role in obtaining closed-loop stability and other desirable closed-loop properties. Here we will show that it is possible to weaken the convexity requirement with the objective of stabilizing the system. The proposed approach is modelled on the first-order one-step-ahead control set-up in Chapter 4 which is deadbeat in nature; of course, here the plant is larger than a first-order one, which increases the complexity. While we are able to remove the convexity requirement completely (see the next chapter), in this chapter we are only weakening it to illustrate the approach. To this end, we impose an assumption on the set of parameter uncertainty.

Assumption 6.1. $\mathcal{S}^* \subset \mathcal{S}_1 \cup \mathcal{S}_2$ with $\mathcal{S}_1$ and $\mathcal{S}_2$ compact and convex, and for each $\theta^* \in \mathcal{S}_1 \cup \mathcal{S}_2$, the corresponding pair of polynomials $z^n \mathbf{A}(z^{-1})$ and $z^n \mathbf{B}(z^{-1})$ are coprime.

As has been discussed earlier, the boundedness assumption is quite reasonable in practical situations; it is used here to ensure that we can prove uniform bounds and decay rates on the closed-loop behavior. The coprimeness assumption arises since our goal is to place the closed-loop poles at desirable locations.

The main goal here is to prove a form of stability. To proceed, we use a parameter estimator together with an adaptive pole placement control law. The idea is to use an estimator for each of $\mathcal{S}_1$ and $\mathcal{S}_2$, and at each point in time we choose which one to use in the control law. In a similar manner as in previous chapters, for each $\theta^* \in \mathcal{S}_i$, $i = 1, 2$, we define

$$i^*(\theta^*) = \min \{ i \in \{1, 2\} \, : \, \theta^* \in \mathcal{S}_i \} \, ;$$

when there is no ambiguity, we will drop the argument and simply write i^*. Also define

$$\bar{\mathbf{s}} := \max \{ \|\mathcal{S}_1\|, \|\mathcal{S}_2\| \} \, .$$

6.2.1 Parameter Estimation

For each $\mathcal{S}_i$ and $\hat{\theta}_i(t_0) \in \mathcal{S}_i$, we construct an estimator which generates an estimate $\hat{\theta}_i(t) \in \mathcal{S}_i$ at each $t > t_0$. The associated prediction error is defined as

$$e_i(t+1) = y(t+1) - \phi(t)^\top \hat{\theta}_i(t). \tag{6.2}$$

Here we apply the simplest form[1] of the estimator (3.9) of Chapter 3, where we set $\phi_m = \phi$ and $\delta = \infty$; observe that in this case

$$\rho_\delta(\phi(t), e_i(t+1)) = 1 \quad \Leftrightarrow \quad \phi(t) \neq 0,$$

and the parameter update law is given by

$$\check{\theta}_i(t+1) = \begin{cases} \hat{\theta}_i(t) + \dfrac{\phi(t)}{\|\phi(t)\|^2} e_i(t+1) & \text{if } \phi(t) \neq 0 \\ \hat{\theta}_i(t) & \text{if } \phi(t) = 0, \end{cases} \tag{6.3a}$$

$$\hat{\theta}_i(t+1) = \operatorname*{Proj}_{\mathcal{S}_i} \left\{ \check{\theta}_i(t+1) \right\}. \tag{6.3b}$$

Associated with this estimator is the parameter estimation error $\tilde{\theta}_i(t) := \hat{\theta}_i(t) - \theta^*$. We now list some properties of the above algorithm; the following result follows directly from combining the results in Propositions 3.1 and 3.2 of Chapter 3, where $\phi_m = \phi$ and $\delta = \infty$. The index i^* is not known, of course.

[1] For ease of exposition, we do not use the more general version of the estimator.

Proposition 6.1. *For every $t_0 \in \mathbb{Z}$, $t_2 > t_1 \geq t_0$, $\phi(t_0) \in \mathbb{R}^{2n}$, $\theta^* \in \mathcal{S}^*$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ ($i = 1, 2$) and $w \in \boldsymbol{\ell}_\infty$, when the estimation algorithm in (6.3) is applied to the plant (6.1), the following holds:*

1. *for* **every estimator** $i = 1, 2$,

$$\|\hat{\theta}_i(t_2) - \hat{\theta}_i(t_1)\| \leq \sum_{j=t_1,\, \phi(j)\neq 0}^{t_2-1} \frac{|e_i(j+1)|}{\|\phi(j)\|}. \tag{6.4}$$

2. *for the* **correct estimator** i^*,

$$\|\tilde{\theta}_{i^*}(t_2)\|^2 \leq \|\tilde{\theta}_{i^*}(t_1)\|^2 + \sum_{j=t_1,\, \phi(j)\neq 0}^{t_2-1} \left[-\frac{1}{2} \frac{e_{i^*}(j+1)^2}{\|\phi(j)\|^2} + 2 \frac{w(j)^2}{\|\phi(j)\|^2} \right]. \tag{6.5}$$

6.2.2 The Switching Control Law

The elements of $\hat{\theta}_i(t)$ are partitioned in a natural way as

$$\hat{\theta}_i(t) =: \begin{bmatrix} \hat{a}_{i,1}(t) \\ \hat{a}_{i,2}(t) \\ \vdots \\ \hat{a}_{i,n}(t) \\ \hat{b}_{i,1}(t) \\ \hat{b}_{i,2}(t) \\ \vdots \\ \hat{b}_{i,n}(t) \end{bmatrix};$$

associated with these estimates are the polynomials

$$\hat{\mathbf{A}}_i(t, z^{-1}) = 1 - \hat{a}_{i,1}(t) z^{-1} - \hat{a}_{i,2}(t) z^{-2} \cdots - \hat{a}_{i,n}(t) z^{-n},$$

and

$$\hat{\mathbf{B}}_i(t, z^{-1}) = \hat{b}_{i,1}(t) z^{-1} + \hat{b}_{i,2}(t) z^{-2} \cdots + \hat{b}_{i,n}(t) z^{-n}.$$

Next, we design a n^{th}-order *strictly* proper controller; we choose the following polynomials

$$\hat{\mathbf{L}}_i(t, z^{-1}) = 1 + \hat{l}_{i,1}(t)z^{-1} + \hat{l}_{i,2}(t)z^{-2} + \cdots + \hat{l}_{i,n}(t)z^{-n},$$

$$\hat{\mathbf{P}}_i(t, z^{-1}) = \hat{p}_{i,1}(t)z^{-1} + \hat{p}_{i,2}(t)z^{-2} + \cdots + \hat{p}_{i,n}(t)z^{-n}$$

to place all closed-loop poles at zero, so we need

$$\hat{\mathbf{A}}_i(t, z^{-1})\hat{\mathbf{L}}_i(t, z^{-1}) + \hat{\mathbf{B}}_i(t, z^{-1})\hat{\mathbf{P}}_i(t, z^{-1}) = 1. \tag{6.6}$$

Given the assumption that the $z^n\hat{\mathbf{A}}_i(t, z^{-1})$ and $z^n\hat{\mathbf{B}}_i(t, z^{-1})$ are coprime, we know that there exist unique $\hat{\mathbf{L}}_i(t, z^{-1})$ and $\hat{\mathbf{P}}_i(t, z^{-1})$ which satisfy this equation; this entails solving a linear equation—see Appendix A for more details. It is also easy to prove that the coefficients of $\hat{\mathbf{L}}_i(t, z^{-1})$ and $\hat{\mathbf{P}}_i(t, z^{-1})$ are analytic functions of $\hat{\theta}_i(t) \in \mathcal{S}_i$.

We can now discuss the candidate control law to be used. Define the control parameters

$$\hat{K}_i(t) := \begin{bmatrix} -\hat{p}_{i,1}(t) & -\hat{p}_{i,2}(t) & \cdots & -\hat{p}_{i,n}(t) & -\hat{l}_{i,1}(t) & -\hat{l}_{i,2}(t) & \cdots & -\hat{l}_{i,n}(t) \end{bmatrix} \tag{6.7}$$

and a switching signal $\sigma : \mathbb{Z} \mapsto \{1, 2\}$ that decides which controller to use at any given point in time. A natural choice for a control law is

$$u(t) = \hat{K}_{\sigma(t-1)}(t-1)\phi(t-1). \tag{6.8}$$

Remark 6.2. *In order to describe the closed-loop system, it is illustrative to first consider the case in which $\mathcal{S}^*$ is a convex set and use only one estimator, as in the first part of Miller and Shahab [46]. In this case, we can exactly obtain an update equation for $\phi(t)$: we use (6.8) to provide the update for $u(t+1)$ and the prediction error equation (6.2) to provide an update for $y(t+1)$, and end up with the equation*

$$\phi(t+1) = \underbrace{\begin{bmatrix} \hat{a}_1(t) & \hat{a}_2(t) & \cdots & \hat{a}_n(t) & \hat{b}_1(t) & \cdots & \cdots & \hat{b}_n(t) \\ 1 & 0 & \cdots & 0 & 0 & \cdots & \cdots & 0 \\ & \ddots & & \vdots & \vdots & \cdots & \cdots & \vdots \\ & & 1 & 0 & 0 & \cdots & \cdots & 0 \\ -\hat{p}_1(t) & -\hat{p}_2(t) & \cdots & -\hat{p}_n(t) & -\hat{l}_1(t) & -\hat{l}_2(t) & \cdots & -\hat{l}_n(t) \\ 0 & \cdots & \cdots & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \cdots & \cdots & \vdots & & \ddots & & \vdots \\ 0 & \cdots & \cdots & 0 & & & 1 & 0 \end{bmatrix}}_{=:\bar{A}(t)} \phi(t) + \mathbf{e}_1 e(t+1); \tag{6.9}$$

here we have omitted the estimator index since it is no longer needed.

As we will soon see, our closed-loop system behavior will in large part be determined by the natural extension of $\bar{A}(t)$ (defined above) to the case of switching—we label this as $\bar{A}_{\sigma(t)}(t) \in \mathbb{R}^{2n}$ (see (6.20) for a precise definition); the closed-loop behavior is captured by

$$\phi(t+1) = \bar{A}_{\sigma(t)}(t)\phi(t) + \mathbf{e}_1 e_{\sigma(t)}(t+1). \tag{6.10}$$

As mentioned earlier, our approach is based on exploiting the deadbeat nature of the problem. While $\bar{A}_{\sigma(t)}(t)$ is a deadbeat matrix (i.e. all of its eigenvalues is at zero) for every t, the product

$$\bar{A}_{\sigma(t)}(t) \times \bar{A}_{\sigma(t-1)}(t-1) \times \cdots \times \bar{A}_{\sigma(t_0)}(t_0), \qquad t \geq t_0$$

will not usually have all eigenvalues at zero. A natural attempt to deal with the problem is to hold $\sigma(t)$ constant in (6.8) for $2n$ steps at a time as well as updating the estimators every $2n$ steps as well; the problem here is that we end up with no information about $e_i(t+1)$ between the updates, so the closed-loop system is not amenable to analysis. So our solution procedure will be different: we are going to change $\sigma(t)$ every $N \geq 2n$ steps; we keep the estimators running, but adjust the control parameters every $N \geq 2n$ steps as well. The effect of this will become clear in the proof of the result. To this end, we define a sequence of switching times as follows: we initialize $\hat{t}_0 := t_0$ and then define

$$\hat{t}_\ell := t_0 + \ell N, \qquad \ell \in \mathbb{N}.$$

So now define the associated control parameters by

$$K_i(t) := \begin{bmatrix} -\hat{p}_{i,1}(\hat{t}_\ell) & -\hat{p}_{i,2}(\hat{t}_\ell) & \cdots & -\hat{p}_{i,n}(\hat{t}_\ell) & -\hat{l}_{i,1}(\hat{t}_\ell) & -\hat{l}_{i,2}(\hat{t}_\ell) & \cdots & -\hat{l}_{i,n}(\hat{t}_\ell) \end{bmatrix},$$
$$t \in [\hat{t}_\ell, \hat{t}_{\ell+1}), \quad \ell \in \mathbb{Z}^+, \tag{6.11}$$

and the switching signal by

$$\sigma(t) = \sigma(\hat{t}_\ell), \qquad t \in [\hat{t}_\ell, \hat{t}_{\ell+1}), \quad \ell \in \mathbb{Z}^+. \tag{6.12}$$

We subsequently define the control law as

$$u(t) = K_{\sigma(t-1)}(t-1)\phi(t-1), \qquad t > t_0. \tag{6.13}$$

What remains to be defined is the choice of switching signal $\sigma(\hat{t}_\ell)$, which we do in the next subsection.

6.2.3 The Switching Algorithm

To proceed, we define a performance signal $J_i : \{\hat{t}_0, \hat{t}_1, \ldots\} \mapsto \mathbb{R}^+$ for estimator i, which produces a measure of "accuracy" of estimation; for $\ell \in \mathbb{Z}^+$, we define

$$J_i(\hat{t}_\ell) := \begin{cases} 0 & \text{if } \phi(j) = 0 \text{ for all } j \in [\hat{t}_\ell, \hat{t}_{\ell+1}), \\ \max\limits_{j \in [\hat{t}_\ell, \hat{t}_{\ell+1}), \phi(j) \neq 0} \dfrac{|e_i(j+1)|}{\|\phi(j)\|} & \text{otherwise.} \end{cases} \tag{6.14}$$

With $\sigma(\hat{t}_0) = \sigma_0$, we use the following switching rule:

$$\sigma(\hat{t}_{\ell+1}) = \underset{i \in \{1,2\}}{\operatorname{argmin}}\, J_i(\hat{t}_\ell), \qquad \ell \in \mathbb{Z}^+. \tag{6.15}$$

For the case when $J_1(\hat{t}_\ell) = J_2(\hat{t}_\ell)$, we (somewhat arbitrarily) select $\sigma(\hat{t}_{\ell+1})$ to be 1. Before presenting the main result of this chapter, we first show that the logic in (6.15) yields a desirable closed-loop property; this result is equivalent to the one in Lemma 5.1.

Lemma 6.1. *Consider the plant (6.1) and suppose that the adaptive controller consisting of the parameter estimator (6.3), the control law (6.13), the performance signal (6.14) and the switching rule (6.15) is applied. Then for every $t_0 \in \mathbb{Z}$, $\phi(t_0) \in \mathbb{R}^{2n}$, $\sigma_0 \in \{1, 2\}$, $N \geq 1$, $\theta^* \in \mathcal{S}^*$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ $(i = 1, 2)$ and $w \in \boldsymbol{\ell}_\infty$, we have that, for every $\ell \in \mathbb{Z}^+$, either*

(a) $J_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell) \leq J_{i^*}(\hat{t}_\ell)$, *or*

(b) $J_{\sigma(\hat{t}_{\ell+1})}(\hat{t}_{\ell+1}) \leq J_{i^*}(\hat{t}_{\ell+1})$.

Proof of Lemma 6.1. Fix $t_0 \in \mathbb{Z}$, $\phi(t_0) \in \mathbb{R}^{2n}$, $\sigma_0 \in \{1, 2\}$, $N \geq 1$, $\theta^* \in \mathcal{S}^*$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ $(i = 1, 2)$, and $w \in \boldsymbol{\ell}_\infty$. Let $\ell \in \mathbb{Z}^+$ be arbitrary. Assume that (a) does not hold, i.e. $J_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell) > J_{i^*}(\hat{t}_\ell)$; then according to (6.15), this means that $\sigma(\hat{t}_{\ell+1}) = i^*$, i.e. (b) will hold. ■

Similar to previous results in this thesis, observe here that we do not make any claim that $\theta^* \in \mathcal{S}_{\sigma(t)}$ at any time; it only makes an indirect statement about the size of the prediction error. It turns out that this is enough to ensure that closed-loop stability is attained.

6.3 The Main Result

Now we present the main result of this chapter.

Theorem 6.1. *Consider the n^{th}-order plant (6.1) and suppose that the adaptive controller consisting of the parameter estimator (6.3), the control law (6.13), the performance signal (6.14) and the switching rule (6.15) is applied to the plant (6.1). For every $\lambda \in (0, 1)$ and $N \geq 2n$, there exists a constant $\gamma > 0$ such that for every $t_0 \in \mathbb{Z}$, $\phi(t_0) \in \mathbb{R}^{2n}$, $\sigma_0 \in \{1, 2\}$, $\theta^* \in \mathcal{S}^*$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ $(i = 1, 2)$, and $w \in \boldsymbol{\ell}_\infty$, the following bound holds:*

$$\|\phi(t)\| \leq \gamma \lambda^{t-\tau} \|\phi(\tau)\| + \gamma \sum_{j=\tau}^{t-1} \lambda^{t-1-j} |w(j)|, \qquad t > \tau \geq t_0. \tag{6.16}$$

Remark 6.3. *The above result shows linear-like closed-loop behavior. There is a* uniform *exponential decay bound on the effect of the initial condition, and a convolution bound on* *the effect of the exogenous signals. This implies that the system has a bounded gain (from w and y^* to y) in every p-norm; in particular, for $p = \infty$ we see from the above bound that*

$$\|\phi(t)\| \leq \gamma\lambda^{t-t_0}\|\phi(t_0)\| + \frac{\gamma}{1-\lambda}\sup_{j\in[t_0,t)}|w(j)| \leq \frac{\gamma}{1-\lambda}\Big(\lambda^{t-t_0}\|\phi(t_0)\| + \|w\|_\infty\Big), \quad t \geq t_0.$$

Hence, if $w \in \boldsymbol{\ell}_\infty$, then $y, u \in \boldsymbol{\ell}_\infty$.

It is natural to ask if the proposed approach would work if $\mathcal{S} \subset \bigcup_{i=1}^{m} \mathcal{S}_i$ where $m > 2$ with each $\mathcal{S}_i$ compact and convex sets for which the corresponding pair of polynomials $\mathbf{A}(z^{-1})$ and $\mathbf{B}(z^{-1})$ are coprime. While the proposed controller (6.3), (6.13), (6.14) and (6.15) is well defined in this case, we have been unable to prove that it will work; as mentioned in earlier chapters a potential problem is that the switching algorithm would oscillate between two bad choices, and never (or rarely) choose the correct one. In the next chapter we provide a more complicated approach which deals with this problem and more.

In order to prove Theorem 6.1, we need the following result that produces a crude bound on the closed-loop behavior.

Proposition 6.2. *Consider the plant (6.1) and suppose that the adaptive controller consisting of the parameter estimator (6.3), the control law (6.13), the performance signal (6.14) and the switching rule (6.15) is applied. Then for every $p \geq 0$, there exists a constant $\bar{c} \geq 1$ such that for every $t_0 \in \mathbb{Z}$, $t \geq t_0$, $\phi(t_0) \in \mathbb{R}^{2n}$, $\sigma_0 \in \{1, 2\}$, $N \geq 1$, $\theta^* \in \mathcal{S}^*$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ $(i = 1, 2)$ and $w \in \boldsymbol{\ell}_\infty$, the following holds:*

$$\|\phi(t+p)\| \leq \bar{c}\|\phi(t)\| + \bar{c}\sum_{j=0}^{p-1}|w(t+j)|. \tag{6.17}$$

Proof. See Appendix B. ∎

Proof of Theorem 6.1:

Fix $\lambda \in (0, 1)$ and $N \geq 2n$. Let $t_0 \in \mathbb{Z}$, $\phi(t_0) \in \mathbb{R}^{2n}$, $\sigma_0 \in \{1, 2\}$, $\theta^* \in \mathcal{S}^*$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ $(i = 1, 2)$, and $w \in \boldsymbol{\ell}_\infty$ be arbitrary.

Step 1: Obtain a state-space model describing $\phi(t)$ for $t \in [\hat{t}_\ell, \hat{t}_{\ell+1})$.

By definition of the prediction error (6.2) and by the property of switching signal (6.12) being constant on $[\hat{t}_\ell, \hat{t}_{\ell+1})$ we have

$$\begin{aligned}
y(t+1) &= \phi(t)^\top \hat{\theta}_{\sigma(t)}(t) + e_{\sigma(t)}(t+1) \\
&= \phi(t)^\top \hat{\theta}_{\sigma(\hat{t}_\ell)}(t) + e_{\sigma(\hat{t}_\ell)}(t+1) + \phi(t)^\top \hat{\theta}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell) - \phi(t)^\top \hat{\theta}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell) \\
&= \hat{\theta}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)^\top \phi(t) + \left[\hat{\theta}_{\sigma(\hat{t}_\ell)}(t) - \hat{\theta}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\right]^\top \phi(t) + e_{\sigma(\hat{t}_\ell)}(t+1), \qquad t \in [\hat{t}_\ell, \hat{t}_{\ell+1}).
\end{aligned} \tag{6.18}$$

From the control law (6.13) and the control gains (6.11) we have

$$\begin{aligned}
u(t+1) &= K_{\sigma(t)}(t)\phi(t) \\
&= K_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\phi(t), \qquad t \in [\hat{t}_\ell, \hat{t}_{\ell+1}).
\end{aligned} \tag{6.19}$$

We now derive a state-space equation for $\phi(t)$; we first define the following $2n \times 2n$ matrix

$$\bar{A}_{\sigma(t)}(t) := \begin{bmatrix}
\hat{a}_{\sigma(t),1}(t) & \hat{a}_{\sigma(t),2}(t) & \cdots & \hat{a}_{\sigma(t),n}(t) & \hat{b}_{\sigma(t),1}(t) & \cdots & \cdots & \hat{b}_{\sigma(t),n}(t) \\
1 & 0 & \cdots & 0 & 0 & \cdots & \cdots & 0 \\
 & \ddots & & \vdots & \vdots & \cdots & \cdots & \vdots \\
 & & 1 & 0 & 0 & \cdots & \cdots & 0 \\
-\hat{p}_{\sigma(t),1}(t) & -\hat{p}_{\sigma(t),2}(t) & \cdots & -\hat{p}_{\sigma(t),n}(t) & -\hat{l}_{\sigma(t),1}(t) & -\hat{l}_{\sigma(t),2}(t) & \cdots & -\hat{l}_{\sigma(t),n}(t) \\
0 & \cdots & \cdots & 0 & 1 & 0 & \cdots & 0 \\
\vdots & \cdots & \cdots & \vdots & & \ddots & & \vdots \\
0 & \cdots & \cdots & 0 & & & 1 & 0
\end{bmatrix}; \tag{6.20}$$

then, in light of (6.18) and (6.19), the following holds

$$\phi(t+1) = \bar{A}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\phi(t) + \mathbf{e}_1 \left(\left[\hat{\theta}_{\sigma(\hat{t}_\ell)}(t) - \hat{\theta}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\right]^\top \phi(t) + e_{\sigma(\hat{t}_\ell)}(t+1) \right), \quad t \in [\hat{t}_\ell, \hat{t}_{\ell+1}), \;\; \ell \in \mathbb{Z}^+. \tag{6.21}$$

Step 2: Obtain a bound on $\|\phi(\hat{t}_{\ell+1})\|$ in terms of $\|\phi(\hat{t}_\ell)\|$.

In (6.21) we have $\bar{A}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell) \in \mathbb{R}^{2n\times 2n}$ to be a constant matrix with all eigenvalues equal to zero; since $N \geq 2n$, clearly

$$\left[\bar{A}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\right]^{\hat{t}_{\ell+1}-\hat{t}_\ell} = \left[\bar{A}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\right]^N = \mathbf{0}.$$

So, solving (6.21) for $\phi(\hat{t}_{\ell+1})$ yields

$$\phi(\hat{t}_{\ell+1}) = \sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+1}-1} \left[\bar{A}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\right]^{\hat{t}_{\ell+1}-j-1} \left[\mathbf{e}_1\left(\left[\hat{\theta}_{\sigma(\hat{t}_\ell)}(j) - \hat{\theta}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\right]^\top \phi(j) + e_{\sigma(\hat{t}_\ell)}(j+1)\right)\right]. \tag{6.22}$$

It follows from the compactness of the $\mathcal{S}_i$'s that $\left\|\left[\bar{A}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\right]^j\right\|$, $j = 0, 1 \ldots, N-1$, is bounded above by a constant which we label c_1. Using this fact together with Part (1) of Proposition 6.1 which provides a bound on the difference between parameter estimates at two different points in time, we obtain

$$\begin{aligned}
\|\phi(\hat{t}_{\ell+1})\| &\leq c_1 \sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+1}-1} \left(\left\|\hat{\theta}_{\sigma(\hat{t}_\ell)}(j) - \hat{\theta}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\right\| \|\phi(j)\| + |e_{\sigma(\hat{t}_\ell)}(j+1)|\right) \\
&\leq c_1 \sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+1}-1} \left(\left[\sum_{q=\hat{t}_\ell, \phi(q)\neq 0}^{j-1} \frac{|e_{\sigma(\hat{t}_\ell)}(q+1)|}{\|\phi(q)\|}\right] \|\phi(j)\| + |e_{\sigma(\hat{t}_\ell)}(j+1)|\right).
\end{aligned}$$

By definition of the prediction error, if $\phi(j) = 0$ then

$$|e_i(j+1)| = |w(j)|,$$

and if $\phi(j) \neq 0$, then

$$|e_i(j+1)| = \frac{|e_i(j+1)|}{\|\phi(j)\|}\|\phi(j)\|.$$

Incorporating this into the above equation yields

$$\begin{aligned}
\Rightarrow \|\phi(\hat{t}_{\ell+1})\| &\leq c_1 \sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+1}-1} \left(\left[\sum_{q=\hat{t}_\ell, \phi(q)\neq 0}^{j} \frac{|e_{\sigma(\hat{t}_\ell)}(q+1)|}{\|\phi(q)\|}\right] \|\phi(j)\| + |w(j)|\right) \\
&\leq c_1 \sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+1}-1} \left(\left[\sum_{q=\hat{t}_\ell, \phi(q)\neq 0}^{\hat{t}_{\ell+1}-1} \frac{|e_{\sigma(\hat{t}_\ell)}(q+1)|}{\|\phi(q)\|}\right] \|\phi(j)\| + |w(j)|\right) \\
&= c_1 \left[\sum_{q=\hat{t}_\ell, \phi(q)\neq 0}^{\hat{t}_{\ell+1}-1} \frac{|e_{\sigma(\hat{t}_\ell)}(q+1)|}{\|\phi(q)\|}\right] \sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+1}-1} \|\phi(j)\| + c_1 \sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+1}-1} |w(j)|
\end{aligned}$$

$$\leq \ c_1(\hat{t}_{\ell+1} - \hat{t}_\ell)\left[\max_{j\in[\hat{t}_\ell,\hat{t}_{\ell+1}),\phi(j)\neq 0} \frac{|e_{\sigma(\hat{t}_\ell)}(j+1)|}{\|\phi(j)\|}\right]\sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+1}-1}\|\phi(j)\| + c_1\sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+1}-1}|w(j)|. \tag{6.23}$$

Since $\hat{t}_{\ell+1} - \hat{t}_\ell = N$, it follows from Proposition 6.2 that there exists a constant c_2 so that the following holds:

$$\sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+1}-1}\|\phi(j)\| \ \leq \ c_2\|\phi(\hat{t}_\ell)\| + c_2\sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+1}-2}|w(j)|; \tag{6.24}$$

so, substituting (6.24) into (6.23) and using the definition of the performance signal $J_{\sigma(\hat{t}_\ell)}(\cdot)$ given in (6.14) it follows that there exists a constant c_3 so that

$$\begin{aligned}\|\phi(\hat{t}_{\ell+1})\| \ &\leq \ c_1 N J_{\sigma(\hat{t}_\ell)}\left(c_2\|\phi(\hat{t}_\ell)\| + c_2\sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+1}-2}|w(j)|\right) + c_1\sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+1}-1}|w(j)| \\ &\leq \ c_3 J_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\|\phi(\hat{t}_\ell)\| + c_3\left(1 + J_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\right)\sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+1}-1}|w(j)|. \end{aligned} \tag{6.25}$$

Step 3: Apply Lemma 6.1 and Proposition 6.2 to obtain a bound on $\|\phi(\hat{t}_{\ell+2})\|$ in terms of $\|\phi(\hat{t}_\ell)\|$.

For an arbitrary $\ell \in \mathbb{Z}^+$, from Lemma 6.1 either

$$J_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell) \leq J_{i^*}(\hat{t}_\ell) \tag{6.26}$$

or

$$J_{\sigma(\hat{t}_{\ell+1})}(\hat{t}_{\ell+1}) \leq J_{i^*}(\hat{t}_{\ell+1}). \tag{6.27}$$

If (6.26) is true, then we can substitute this into (6.25) to obtain a bound on $\|\phi(\hat{t}_{\ell+1})\|$ in terms of $J_{i^*}(\hat{t}_\ell)$, and then apply Proposition 6.2 to get a bound on $\|\phi(\hat{t}_{\ell+2})\|$ in terms of $\|\phi(\hat{t}_{\ell+1})\|$ and the exogenous inputs; it follows that there exists a constant c_4 so that

$$\|\phi(\hat{t}_{\ell+2})\| \leq c_3 c_4 J_{i^*}(\hat{t}_\ell)\|\phi(\hat{t}_\ell)\| +$$

$$c_3c_4\left(1+J_{i^*}(\hat{t}_\ell)\right)\sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+1}-1}|w(j)|+c_4\sum_{j=\hat{t}_{\ell+1}}^{\hat{t}_{\ell+2}-1}|w(j)|. \tag{6.28}$$

On the other hand, if (6.27) is true, we can use (6.25) to get a bound on $\|\phi(\hat{t}_{\ell+2})\|$ in terms of $J_{i^*}(\hat{t}_{\ell+1})$, and then apply Proposition 6.2 to get a bound on $\|\phi(\hat{t}_{\ell+1})\|$ in terms of $\|\phi(\hat{t}_\ell)\|$; it follows that there exists a constant c_5 so that

$$\begin{aligned}\|\phi(\hat{t}_{\ell+2})\| \le c_3c_5J_{i^*}(\hat{t}_{\ell+1})\|\phi(\hat{t}_\ell)\|+c_3c_5J_{i^*}(\hat{t}_{\ell+1})\sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+1}-1}|w(j)|\\ +c_3\left(1+J_{i^*}(\hat{t}_{\ell+1})\right)\sum_{j=\hat{t}_{\ell+1}}^{\hat{t}_{\ell+2}-1}|w(j)|.\end{aligned} \tag{6.29}$$

If we define

$$\alpha(\hat{t}_\ell):=\max\left\{J_{i^*}(\hat{t}_\ell),J_{i^*}(\hat{t}_{\ell+1})\right\},$$

then there exist a constant c_6 so that (6.28) and (6.29) can be combined to yield

$$\|\phi(\hat{t}_{\ell+2})\|\le c_6\alpha(\hat{t}_\ell)\|\phi(\hat{t}_\ell)\|+c_6\left(1+\alpha(\hat{t}_\ell)\right)\sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+2}-1}|w(j)|,\qquad \ell\in\mathbb{Z}^+. \tag{6.30}$$

Step 4: Analyze the first-order difference inequality (6.30).

First we change notation in (6.30) to facilitate analysis:

$$\|\phi(\hat{t}_{2j+2})\|\le c_6\alpha(\hat{t}_{2j})\|\phi(\hat{t}_{2j})\|+c_6\left(1+\alpha(\hat{t}_{2j})\right)\sum_{q=\hat{t}_{2j}}^{\hat{t}_{2j+2}-1}|w(q)|,\qquad j\in\mathbb{Z}^+. \tag{6.31}$$

Next, we will analyze (6.31) to obtain a bound on the closed-loop behavior. Define

$$\lambda_1:=\frac{\lambda^{2N}}{\max\{1,c_6\}}$$

and

$$\nu:=\left(\frac{\lambda_1}{4N}\right)^2.$$

Let $\tau\ge t_0$ be arbitrary. We now partition the timeline into two parts: one in which the noise is small versus ϕ and one where it is not. With ν defined above, we define

$$S_{good} = \left\{ j \geq \tau : \phi(j) \neq 0 \text{ and } \frac{|w(j)|^2}{\|\phi(j)\|^2} < \nu \right\},$$

$$S_{bad} = \left\{ j \geq \tau : \phi(j) = 0 \text{ or } \frac{|w(j)|^2}{\|\phi(j)\|^2} \geq \nu \right\};$$

clearly $\{j \in \mathbb{Z} : \ \ j \geq \tau\} = S_{good} \cup S_{bad}$. We can clearly define a (possibly infinite) sequence of intervals of the form $[k_l, k_{l+1})$ which satisfy:
(i) $k_0 = \tau$ serves as the initial instant of the first interval;
(ii) $[k_l, k_{l+1})$ either belongs to S_{good} or S_{bad}; and
(iii) if $k_{l+1} \neq \infty$ and $[k_l, k_{l+1})$ belongs to S_{good} then $[k_{l+1}, k_{l+2})$ belongs to S_{bad} and vice versa.

Now we analyze the behavior during each interval.

Step 4.1: $[k_l, k_{l+1})$ lies in S_{bad}.

Let $j \in [k_l, k_{l+1})$ be arbitrary. In this case

$$\frac{|w(j)|^2}{\|\phi(j)\|^2} \geq \nu \qquad \text{or} \qquad \|\phi(j)\| = 0;$$

in either case

$$\|\phi(j)\| \leq \underbrace{\frac{1}{\sqrt{\nu}}}_{=:c_7} |w(j)|.$$

Also, applying Proposition 6.2 for one step, there exists constant c_8 so that

$$\|\phi(j)\| \leq c_8 |w(j-1)|.$$

Then for $j \in [k_l, k_{l+1})$, we have

$$\|\phi(j)\| \leq \begin{cases} c_7 |w(j)| & j = k_l \\ c_8 |w(j-1)| & j = k_l + 1, k_l + 2, \ldots, k_{l+1}. \end{cases} \tag{6.32}$$

Step 4.2: $[k_l, k_{l+1})$ lies in S_{good}.

First suppose that $k_{l+1} - k_l \leq 4N$. From Proposition 6.2 it can be easily proven that there exists a constant c_9 so that

$$\|\phi(t)\| \leq c_9 \lambda^{t-k_l} \|\phi(k_l)\| + c_9 \sum_{j=k_l}^{t-1} \lambda^{t-j-1} |w(j)|, \qquad t \in [k_l, k_{l+1}]. \tag{6.33}$$

Now suppose that $k_{l+1} - k_l > 4N$. This means, in particular, that there exist $j_1 < j_2$ so that

$$k_l \leq \hat{t}_{2j_1} < \hat{t}_{2j_2} \leq k_{l+1}.$$

Let $j \in [k_l, k_{l+1})$ be arbitrary. To proceed, observe that $\|\phi(j)\| \neq 0$ and

$$\frac{|w(j)|^2}{\|\phi(j)\|^2} < \nu. \tag{6.34}$$

It follows from Part (2) of Proposition 6.1 and the definition of $\alpha(\cdot)$ that

$$\begin{aligned} \sum_{q=j_1}^{j_2-1} \alpha(\hat{t}_{2q})^2 &\leq N \sum_{p=\tau+2j_1N, \phi(p)\neq 0}^{\tau+2j_2N-1} \frac{|e_{i^*}(p+1)|^2}{\|\phi(p)\|^2} \\ &\leq 2N\|\tilde{\theta}_{i^*}(\hat{t}_{2j_1})\|^2 + 4N \sum_{p=\hat{t}_{2j_1}, \phi(p)\neq 0}^{\hat{t}_{2j_2}-1} \frac{|w(p)|^2}{\|\phi(p)\|^2}, \qquad 0 \leq j_1 < j_2; \end{aligned} \tag{6.35}$$

using the bound given in (6.34) which holds on $[k_l, k_{l+1})$, this becomes

$$\sum_{q=j_1}^{j_2-1} \alpha(\hat{t}_{2q})^2 \leq 8N\bar{\mathbf{s}}^2 + 8N^2(j_2 - j_1)\nu, \quad \text{for all } j_1, j_2 \in \mathbb{Z}^+ \text{ s.t. } k_l \leq \hat{t}_{2j_1} < \hat{t}_{2j_2} \leq k_{l+1}. \tag{6.36}$$

We want now to utilize the inequality of arithmetic and geometric means.

Claim 6.1. *There exists a constant $\gamma_1 > 1$ such that*

$$\prod_{j=j_2}^{j_1-1} \alpha(\hat{t}_{2j}) \leq \gamma_1 \lambda_1^{j_2-j_1}, \qquad \textit{for } j_1, j_2 \in \mathbb{Z}^+ \textit{ s.t. } k_l \leq \hat{t}_{2j_1} < \hat{t}_{2j_2} \leq k_{l+1}. \tag{6.37}$$

Proof of Claim 6.1. Let $j_1, j_2 \in \mathbb{Z}^+$ be arbitrary such that $k_l \leq \hat{t}_{2j_1} < \hat{t}_{2j_2} \leq k_{l+1}$. By the fact that $\alpha(\cdot) \geq 0$, we obtain

$$\prod_{j=j_1}^{j_2-1} \alpha(\hat{t}_{2j}) \leq \left[\frac{1}{j_2 - j_1} \sum_{j=j_1}^{j_2-1} \alpha(\hat{t}_{2j})^2\right]^{\frac{j_2-j_1}{2}}. \tag{6.38}$$

Substituting (6.36) into the above yields

$$\prod_{j=j_1}^{j_2-1} \alpha(\hat{t}_{2j}) \leq \left[\frac{8N\bar{\mathbf{s}}^2}{j_2 - j_1} + 8N^2\nu\right]^{\frac{j_2-j_1}{2}}. \tag{6.39}$$

So it is enough to prove that there exists a constant γ_1 so that

$$\Bigg(\underbrace{\left[\frac{8N\bar{\mathbf{s}}^2}{k} + 8N^2\nu\right]^{\frac{1}{2}}}_{=:g(k)}\Bigg)^k \leq \gamma_1 \lambda_1^k, \qquad k > 0.$$

We can easily show that with $\bar{k} := 16N \times \left\lceil \left(\frac{\bar{\mathbf{s}}}{\lambda_1}\right)^2 \right\rceil$, we have

$$\frac{8N\bar{\mathbf{s}}^2}{\bar{k}} \leq \frac{\lambda_1^2}{2},$$

which means that by the choice of ν, we see that

$$g(k)^k \leq \lambda_1^k, \qquad k \geq \bar{k}.$$

So if we define $\gamma_1 := \max\left\{1, \left(\frac{g(1)}{\lambda_1}\right)^{\bar{k}}\right\}$, we conclude the proof of the claim. $\square$

Using the bound in (6.37) and the definition of λ_1 we obtain

$$\begin{aligned}\prod_{j=j_1}^{j_2-1} [c_6 \alpha(\hat{t}_{2j})] &\leq \gamma_1 \lambda_1^{j_2-j_1} c_6^{j_2-j_1}, \\ &\leq \gamma_1 \lambda^{2N(j_2-j_1)}, \quad \text{for } j_1, j_2 \in \mathbb{Z}^+ \text{ s.t. } k_l \leq \hat{t}_{2j_1} < \hat{t}_{2j_2} \leq k_{l+1}.\end{aligned} \tag{6.40}$$

Now we can proceed to solve (6.31). Before proceeding, we use (6.36) to have an upper bound on $\alpha(\cdot)$:

$$\alpha(\hat{t}_{2j}) \leq \sqrt{8N(\bar{\mathbf{s}}^2 + N\nu)} =: c_{10}, \quad \text{for all } j \in \mathbb{Z}^+ \text{ s.t. } k_l \leq \hat{t}_{2j} < \hat{t}_{2(j+1)} \leq k_{l+1}.$$

The first step is to bound the second occurrence of $\alpha(\hat{t}_{2j})$ in (6.31), yielding

$$\|\phi(\hat{t}_{2j+2})\| \leq c_6 \alpha(\hat{t}_{2j}) \|\phi(\hat{t}_{2j})\| + \underbrace{c_6(1 + c_{10})}_{=:c_{11}} \underbrace{\sum_{q=\hat{t}_{2j}}^{\hat{t}_{2j+2}-1} |w(q)|}_{=:\bar{w}(j)},$$
$$\text{for all } j \in \mathbb{Z}^+ \text{ s.t. } k_l \leq \hat{t}_{2j} < \hat{t}_{2j+2} \leq k_{l+1}. \tag{6.41}$$

If we solve this iteratively and use the bounds in (6.40), we see that

$$\|\phi(\hat{t}_{2j_2})\| \leq \gamma_1 \lambda^{2N(j_2-j_1)} \|\phi(\hat{t}_{2j_1})\| + \sum_{q=j_1}^{j_2-1} \gamma_1 c_{11} \left(\lambda^{2N}\right)^{j_2-1-q} \bar{w}(q),$$
$$\text{for all } j_1, j_2 \in \mathbb{Z}^+ \text{ s.t. } k_l \leq \hat{t}_{2j_1} < \hat{t}_{2j_2} \leq k_{l+1}. \tag{6.42}$$

We can now use Proposition 6.2 for no more than $2N$ steps at a time:

- to provide a bound on $\|\phi(t)\|$ between consecutive $\hat{t}_{2j}$'s;
- to provide a bound on $\|\phi(t)\|$ on the beginning part of the interval $[k_l, k_{l+1})$ (until we get to the first admissible $\hat{t}_{2j}$);
- to provide a bound on $\|\phi(t)\|$ on the last part of the interval $[k_l, k_{l+1})$ (after the last admissible $\hat{t}_{2j}$).

We conclude that there exist a constant $\gamma_2 \geq c_9$ so that

$$\|\phi(t)\| \leq \gamma_2 \lambda^{t-k_l} \|\phi(k_l)\| + \gamma_2 \sum_{j=k_l}^{t-1} \lambda^{t-j-1} |w(j)|, \qquad t \in [k_l, k_{l+1}]. \tag{6.43}$$

Now we combine Step 4.1 and Step 4.2 into a general bound on ϕ. Define

$$\gamma := \max\{\gamma_2, c_8, c_8\gamma_2\}.$$

Claim 6.2. *The following bound holds:*

$$\|\phi(t)\| \leq \gamma\lambda^{t-\tau}\|\phi(\tau)\| + \sum_{j=\tau}^{t-1} \gamma\lambda^{t-j-1}|w(j)|, \qquad t \geq \tau. \tag{6.44}$$

Proof of the Claim 6.2. If $[k_0, k_1) = [\tau, k_1) \subset S_{good}$, then (6.44) is true for $t \in [k_0, k_1]$ by (6.43). If $[k_0, k_1) \subset S_{bad}$, then from (6.32) we obtain

$$\|\phi(j)\| \leq \begin{cases} \|\phi(k_0)\| = \|\phi(\tau)\| & j = k_0 = \tau \\ c_8|w(j-1)| & j = k_0+1, k_0+2, \ldots, k_1. \end{cases}$$

which means that (6.44) holds on $[k_0, k_1]$ for this case as well.

We now use induction: suppose that (6.44) is true for $t \in [k_0, k_l]$; we need to prove it holds for $t \in (k_l, k_{l+1}]$ as well. If $k \in [k_l, k_{l+1}) \subset S_{bad}$, then from (6.32) we see that

$$\|\phi(j)\| \leq c_8|w(j-1)|, \quad j = k_l+1, k_l+2, \ldots, k_{l+1},$$

which means (6.44) holds on $(k_l, k_{l+1}]$. On the other hand, if $[k_l, k_{l+1}) \subset S_{good}$, then $k_l - 1 \in S_{bad}$; from (6.32) we have that

$$\|\phi(k_l)\| \leq c_8|w(k_l - 1)|.$$

Using (6.43) to analyze the behavior on $[k_l, k_{l+1}]$, we have

$$\begin{aligned} \|\phi(t)\| &\leq \gamma_2\lambda^{t-k_l}[c_8|w(k_l-1)|] + \sum_{j=k_l}^{t-1} \gamma_2\lambda^{t-j-1}|w(j)|, \\ &\leq \gamma \sum_{j=k_l-1}^{t-1} \lambda^{t-j-1}|w(j)|, \qquad t \in [k_l, k_{l+1}], \end{aligned} \tag{6.45}$$

which implies that (6.44) holds. □

As $\tau \geq t_0$ is arbitrary, this concludes the proof. ■

6.4 Robustness Results

Here we show that we can leverage the fact that a convolution bound holds in the case of a fixed plant parameter to prove that a convolution bound also holds if we allow time-variation and/or unmodelled dynamics. To proceed, we consider a time-varying version of the plant (6.1) along with the term $d_\Delta(t) \in \mathbb{R}$ added to represent the unmodelled dynamics:

$$y(t+1) = \theta^*(t)^\top \phi(t) + w(t) + d_\Delta(t), \quad t \in \mathbb{Z}; \tag{6.46}$$

as discussed in Chapter 2, we assume that d_Δ satisfies

$$m(t+1) = \beta m(t) + \beta\|\phi(t)\|, \qquad m(t_0) = m_0 \tag{6.47a}$$
$$|d_\Delta(t)| \leq \mu m(t) + \mu\|\phi(t)\|, \qquad t \geq t_0. \tag{6.47b}$$

Theorem 6.2. *Suppose that the adaptive controller (6.3), (6.13), (6.14) and (6.15) is applied to the time-varying plant (6.46) with d_Δ satisfying (6.47). Then for every $\beta \in (0,1)$, $N \geq 2n$, and $\bar{c}_0 \geq 0$, there exist $\bar{\epsilon} > 0$, $\mu > 0$, $\tilde{\lambda} \in (\beta, 1)$ and $\tilde{\gamma} > 0$ such that for every $t_0 \in \mathbb{Z}$, $\phi(t_0) \in \mathbb{R}^{2n}$, $\sigma_0 \in \{1,2\}$, $\theta^* \in \mathbb{S}(\mathcal{S}^*, \bar{c}_0, \bar{\epsilon})$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ $(i = 1,2)$, and $w \in \boldsymbol{\ell}_\infty$, the following holds:*

$$\left\| \begin{bmatrix} \phi(t) \\ m(t) \end{bmatrix} \right\| \leq \tilde{\gamma}\tilde{\lambda}^{t-t_0} \left\| \begin{bmatrix} \phi(t_0) \\ m_0 \end{bmatrix} \right\| + \sum_{j=t_0}^{t-1} \tilde{\gamma}\tilde{\lambda}^{t-j-1}|w(j)|, \quad t \geq t_0.$$

Proof of Theorem 6.2. We observe here that the plant (6.46) and the controller (6.3), (6.13), (6.14), (6.15) fit into the paradigm of Chapter 2: we set

$$\begin{aligned}
\vartheta(t) &= \phi(t), \\
f(\vartheta(\cdot)) &= \phi(\cdot) \\
z(t) &= \varnothing, \\
\hat{\theta}(t) &= \begin{bmatrix} \hat{\theta}_1(t) \\ \hat{\theta}_2(t) \end{bmatrix}, \\
r(t) &= 0, \\
\Omega &= \mathcal{S}_1 \times \mathcal{S}_2.
\end{aligned}$$

In Theorem 6.1 it is proven the controller (6.3), (6.13), (6.14), (6.15) provides a convolution bound for (6.1). Then, by Theorems 2.2, 2.2 and 2.3 we immediately see that the same is true in the presence of time-variation and/or unmodelled dynamics. ■

6.5 A Simulation Example

We will consider the second order plant

$$y(t+1) = a_1(t)y(t) + a_2(t)y(t-1) + b_1(t)u(t) + b_2(t)u(t-1) + w(t) + d_\Delta(t)$$

with θ^* belonging to $\mathcal{S}_1 \cup \mathcal{S}_2$:

$$\mathcal{S}_1 := \left\{ \begin{bmatrix} a_1 \\ a_2 \\ b_1 \\ b_2 \end{bmatrix} \in \mathbb{R}^4 : a_1 \in [0,2], a_2 \in [1,3], b_1 \in [0,1], b_2 \in [-5,-2] \right\},$$

$$\mathcal{S}_2 := \left\{ \begin{bmatrix} a_1 \\ a_2 \\ b_1 \\ b_2 \end{bmatrix} \in \mathbb{R}^4 : a_1 \in [0,2], a_2 \in [1,3], b_1 \in [-1,0], b_2 \in [2,5] \right\};$$

observe that each of $\mathcal{S}_1$ and $\mathcal{S}_2$ is convex. Every admissible plant is unstable and non-minimum phase, which makes this a challenging plant to control; it has two complex unstable poles and a zero that lie in $[2,\infty)$. Note also that the convex hull of $\mathcal{S}_1 \cup \mathcal{S}_2$ includes the case when $b_1 = b_2 = 0$, that corresponds to a non-stabilizable plant, violating the coprimeness assumption; hence, the proposed approach is applied.

We want to illustrate the approach and its robustness; to this end, we will examine the case of the proposed controller when it is applied to the time-varying plant with unmodelled dynamics entering the system, a zero initial condition, and a non-zero noise. Specifically, we set the time-varying parameters to

$$a_1(t) = 1 + \sin\left(\tfrac{1}{1000}t\right),$$

$$a_2(t) = 2 + \cos\left(\tfrac{1}{1000}t\right),$$

$$b_1(t) = \begin{cases} -0.5 - 0.5\sin\left(\frac{1}{200}t\right) & 1500 \le t < 8000 \\ 0.5 + 0.5\sin\left(\frac{1}{200}t\right) & \text{otherwise,} \end{cases}$$

$$b_2(t) = \begin{cases} 3.5 - 1.5\sin\left(\frac{1}{200}t\right) & 1500 \le t < 8000 \\ -3.5 + 1.5\sin\left(\frac{1}{200}t\right) & \text{otherwise.} \end{cases}$$

We apply the proposed switching controller consisting of the estimator (6.3), the control law

(6.13), the performance signal (6.14) and the switching rule (6.15); we choose $N = 2n = 4$. Here we also set $y(0) = y(-1) = u(0) = u(-1) = 0$ and the noise to

$$w(t) = 0.01\sin(5t).$$

For the unmodelled dynamics, we choose a model of

$$\begin{aligned} m(t+1) &= 0.75m(t) + 0.75\|\phi(t)\|, \quad m(0) = 0, \\ d_\Delta(t) &= \begin{cases} 0 & t < 5000 \\ 0.025m(t) + 0.025\|\phi(t)\| & \text{otherwise.} \end{cases} \end{aligned}$$

Initial parameter estimates $\hat{\theta}_i(0)$ are set to the midpoints of each respective interval, and set $\sigma_0 = 2$.

The result for this case is plotted in Figure 6.1; we see that the controller does a reasonable job, even though the switching often chooses the wrong model. Larger transients may ensue, but on average the adaptive controller provides good performance. Furthermore, the estimator does a fairly good job of tracking the time-varying parameters. This illustrates that the approach handles time-variation and occasional jumps, as well as unmodelled dynamics.

Remark 6.4. *We can also compare the performance here with that which arises when we use the classical estimator (3.4) as part of the adaptive controller; we often end up with the same sort of result as in Example 2 of the simulation section of Chapter 4, namely a degradation in performance.*

6.6 Conclusion

In this chapter, we have shown that for possibly non-minimum phase plants, we are able to stabilize and obtain linear-like closed behavior subject to a standard coprimeness assumption on a compact, but possibly non-convex, set of plant uncertainty; that being said, we required the set of admissible plant parameters to lie in the union of two convex sets. Two parameter estimators along with a simple switching rule are used to choose which parameters are used in the pole-placement based control law. While we are able to extend that approach to more complex compact sets, we have decided to defer this to the next chapter, where we consider a more general situation which also incorporates a tracking requirement.

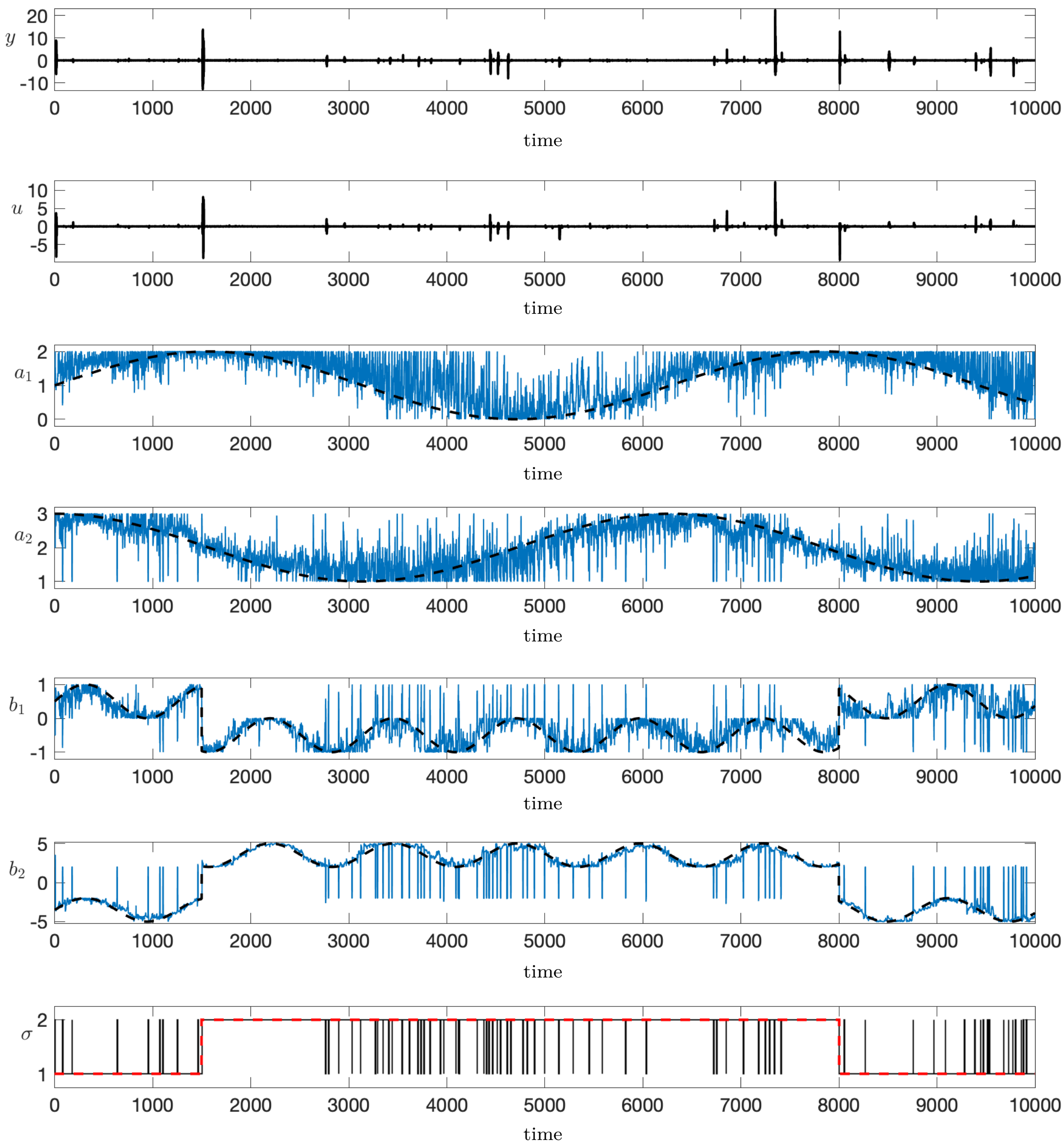


Figure 6.1: The upper two plots show system output and control input, respectively. The next four plots show the parameter estimates (solid) and actual plant parameters (dashed). The bottom plot shows the switching signal $\sigma(t)$ (solid) and the correct index (dashed).

Chapter 7

Multi-Model Adaptive Control and Tracking

7.1 Introduction

In Chapter 6, the convexity assumption is weakened slightly (without completely removing it) and stability is proven, but not tracking. In this chapter, we consider the problem of stabilization and tracking the sum of a finite number of sinusoids of known frequencies in the presence of plant uncertainty. We assume knowledge of an upper bound on the plant order, and for each admissible order we assume knowledge of a compact set in which the plant parameters lie; although we impose some natural technical assumptions on the sets, we do not assume that they are convex. To facilitate the tracking requirement, rather than directly estimating the plant parameters, we instead estimate the parameters of a suitably defined auxiliary plant model. We use the compactness of the parameter uncertainty set for each admissible order to prove that it is contained in a finite union of compact and convex sets; we construct a parameter estimator for each of these compact and convex sets, based on the original projection algorithm. A switching algorithm is used to determine which estimates are used in the controller at a given point in time; unlike many switching approaches in the literature, e.g. [41], [26] and [7], our approach does not assume that the switching stops; actually, the switching algorithms proposed in the previous chapters have this property as well. We prove that the desired linear-like convolution bounds are achieved, and if the reference and disturbance signals belong to the aforementioned class of sinusoids then the tracking error goes exponentially to zero. A preliminary version of this chapter appears in [69], and deals only with the problem of step tracking with a known plant order;

a complete version has been submitted [71].

We first discuss the unknown plant, the auxiliary plant to be used for estimation, and the uncertainty sets. Then we introduce the *multi-model adaptive controller*, followed by the main result, which shows that the closed-loop behavior satisfies the desired convolution bound, and asymptotic tracking for certain classes of reference and noise signals is provided.

7.2 The Setup

7.2.1 The Plant

As in the previous chapter, here we consider the n^{th}-order linear time-invariant discrete-time plant

$$y(t+1) = \sum_{j=1}^{n} a_j y(t-j+1) + \sum_{j=1}^{n} b_j u(t-j+1) + w(t), \qquad t \in \mathbb{Z}, \tag{7.1}$$

with $y(t), u(t), w(t) \in \mathbb{R}$ denoting the measured output, the control input, and the disturbance/noise input, respectively. Since our goal is more demanding here, we will proceed in a different way than in Chapter 6. A plant of the form (7.1) can be expressed in the (two-sided) z-transform form as

$$\mathbf{A}(z^{-1})Y(z) = \mathbf{B}(z^{-1})U(z) + z^{-1}W(z), \tag{7.2}$$

with the corresponding polynomials defined as

$$\mathbf{A}(z^{-1}) := 1 - a_1 z^{-1} - a_2 z^{-2} \cdots - a_n z^{-n},$$

and

$$\mathbf{B}(z^{-1}) := b_1 z^{-1} + b_2 z^{-2} \cdots + b_n z^{-n},$$

with $Y(z), U(z)$ and $W(z)$ denoting the z-transform of $y(t), u(t)$ and $w(t)$, respectively. The plant can be represented by the strictly proper transfer function $\frac{\mathbf{B}(z^{-1})}{\mathbf{A}(z^{-1})}$. Note that Remark 6.1 of Chapter 6 applies here too. We can represent the plant model by the vector of parameters

$$\theta = \begin{bmatrix} a_1 & a_2 & \cdots & a_n & b_1 & b_2 & \cdots & b_n \end{bmatrix}^\top \in \mathbb{R}^{2n}.$$

The objective is to control the system when θ is unknown but lies in a set of admissible parameters. Since our goal is to provide uniform bounds, we shall require that this set be compact; we will not insist on convexity. Since we will be using a pole-placement approach, we will require that $z^n\mathbf{A}(z^{-1})$ and $z^n\mathbf{B}(z^{-1})$ be coprime. Indeed, it turns out that our approach works if the plant order n is **not known** exactly, but rather we have an upper bound $\bar{n}$ on n. To this end, for each $n \in \{1, 2, \dots, \bar{n}\}$ we let

$$\mathbf{\Theta}_n \subset \mathbb{R}^{2n}$$

denote the set of admissible parameters, and impose

Assumption 7.1. For every $n \in \{1, 2, \dots, \bar{n}\}$:

(1) the set $\mathbf{\Theta}_n$ is compact[1]and

(2) for every $\theta \in \mathbf{\Theta}_n$, the corresponding polynomials $z^n\mathbf{A}(z^{-1})$ and $z^n\mathbf{B}(z^{-1})$ are coprime.

7.2.2 Control Objective

The objective is to prove an exponential form of stability, a bounded gain on the noise and on a general reference signal, and tracking (and disturbance rejection) of the sum of a finite number of sinusoids of known frequencies. To this end, consider a g^{th}-order polynomial of the form

$$\mathbf{Q}(z^{-1}) = 1 - q_1 z^{-1} - q_2 z^{-2} \cdots - q_g z^{-g} \tag{7.3}$$

with all of its roots belonging to the unit circle and with no multiplicities. Let $y^*(t) \in \mathbb{R}$ be the reference signal; the set of reference signals y^* to be tracked, and/or disturbance signals w to be rejected, includes those satisfying

$$\mathbf{Q}(z^{-1})Y^*(z) = 0, \quad \text{and/or} \quad \mathbf{Q}(z^{-1})W(z) = 0, \tag{7.4}$$

with $Y^*(z)$ and $W(z)$ being the z-transform of $y^*(t)$ and $w(t)$, respectively.

Remark 7.1. *If we wish to consider set-point tracking (and/or constant disturbance rejection), we should set*

$$\mathbf{Q}(z^{-1}) = 1 - z^{-1}.$$

[1]It could very well be that $\mathbf{\Theta}_n$ is empty for some n.

If we wish to track (and/or reject) a sinusoid of the form $A_0 \cos(\omega_0 t)$, we should set

$$\mathbf{Q}(z^{-1}) = 1 - 2\cos(\omega_0) z^{-1} + z^{-2}.$$

More generally, if we want to track a reference signal (and/or reject a disturbance signal) of the form

$$A_0 + \sum_{p=1}^{n_g} A_p \sin(\omega_p t + \phi_p)$$

with n_g distinct frequencies $\omega_1, \omega_2, \ldots, \omega_{n_g}$, then we should set

$$\begin{aligned}
\mathbf{Q}(z^{-1}) &= (1 - e^{j0} z^{-1}) \times \prod_{p=1}^{n_g} \Big((1 - e^{j\omega_p} z^{-1})(1 - e^{-j\omega_p} z^{-1}) \Big) \\
&= (1 - z^{-1}) \times \prod_{p=1}^{n_g} \Big(1 - 2\cos(\omega_p) z^{-1} + z^{-2} \Big).
\end{aligned}$$

As we know from the classical continuous-time control, if the plant has a zero at the origin then we cannot design an LTI stability controller which ensures the plant to track steps. To this end, we impose the following natural assumption:

Assumption 7.2. For each $n \in \{1, 2, \ldots, \bar{n}\}$, we assume that for every $\theta \in \mathbf{\Theta}_n$, the corresponding polynomial $z^n \mathbf{B}(z^{-1})$ and the polynomial $z^g \mathbf{Q}(z^{-1})$ are coprime.

Remark 7.2. *Observe that the plant may be non-minimum phase.*

7.2.3 The Auxiliary Plant

If n is known and the set of admissible parameters $\mathbf{\Theta}_n$ is convex, then the classical approach is to carry out system identification of the plant in the usual way, e.g. [21], and design the pole-placement based control law in such a way as to force an "internal model of $\mathbf{Q}(z^{-1})$" into the controller; this has been shown to be quite effective in classical results which prove asymptotic stability, e.g. see [21], as well as in Miller and Shahab [46] where exponential stability and step tracking is proven. If, however, the set of admissible parameters is not convex, which can be the case here, the standard trick is to replace it with its closed convex hull. Unfortunately, often that set will contain models that violate coprimeness, so we need another approach. The compactness of the set of admissible parameters can be utilized to easily prove that it is contained in a finite union of convex sets with desirable properties; we can then use an estimator for each convex set and from time to time switch between estimates for use in the control law. We have analyzed this approach, and while we have been able to prove an exponential type of stability, we have been unable to achieve our tracking objective. We will deal with this difficulty by doing system identification on a related auxiliary plant model rather than the original plant model.

Remark 7.3. *In many of the results on "switching adaptive control" in the literature, which typically considers the noise-free case, to prove asymptotic tracking they generally rely on the fact that the switching mechanism stops at some point, e.g. see [41], [26] and [7]. With unknown noise entering the system, as it is in our case, it is generally not possible to conclude that the switching eventually stops.*

With $y^*, w \in \boldsymbol{\ell}_\infty$, let us define the tracking error ε by

$$\varepsilon(t) := y(t) - y^*(t); \tag{7.5}$$

also define an auxiliary control input $v(t) \in \mathbb{R}$ and its z-transform counterpart by

$$V(z) := \mathbf{Q}(z^{-1})U(z) \tag{7.6a}$$

$$\Leftrightarrow v(t) = u(t) - q_1 u(t-1) - \cdots - q_g u(t-g). \tag{7.6b}$$

If we multiply both sides of the z-transformed counterpart of the plant model (7.2) by $\mathbf{Q}(z^{-1})$ and use the definition in (7.6), then we end up with

$$\mathbf{Q}(z^{-1})\mathbf{A}(z^{-1})Y(z) = \mathbf{B}(z^{-1})V(z) + z^{-1}\mathbf{Q}(z^{-1})W(z);$$

denoting the z-transform of $\varepsilon(t)$ by $\mathcal{E}(z)$, if we subtract $\mathbf{Q}(z^{-1})\mathbf{A}(z^{-1})Y^*(z)$ from both sides of the above equation then we obtain the auxiliary plant model

$$\underbrace{\mathbf{Q}(z^{-1})\mathbf{A}(z^{-1})}_{=:\bar{\mathbf{A}}(z^{-1})}\mathcal{E}(z) = \mathbf{B}(z^{-1})V(z) + z^{-1}\underbrace{\mathbf{Q}(z^{-1})\left[W(z) - z\mathbf{A}(z^{-1})Y^*(z)\right]}_{=:\bar{W}(z)}, \tag{7.7}$$

or in other words

$$\bar{\mathbf{A}}(z^{-1})\mathcal{E}(z) = \mathbf{B}(z^{-1})V(z) + z^{-1}\bar{W}(z). \tag{7.8}$$

We examine the polynomial $\bar{\mathbf{A}}(z^{-1})$ carefully; we have

$$\begin{aligned}\bar{\mathbf{A}}(z^{-1}) &=: 1 - \bar{a}_1 z^{-1} - \bar{a}_2 z^{-2} \cdots - \bar{a}_{n+g} z^{-(n+g)} \\ &= \mathbf{Q}(z^{-1})\mathbf{A}(z^{-1}) \\ &= \left(1 - \sum_{j=1}^{g} q_j z^{-j}\right)\left(1 - \sum_{j=1}^{n} a_j z^{-j}\right)\end{aligned}$$

which can be written as

$$\bar{\mathbf{A}}(z^{-1}) = 1 - \begin{bmatrix} z^{-1} & z^{-2} & \cdots & z^{-(n+g)} \end{bmatrix} \underbrace{\begin{bmatrix} q_1 & 1 & & & \\ q_2 & -q_1 & 1 & & \\ \vdots & -q_2 & -q_1 & \ddots & \\ \vdots & \vdots & -q_2 & \ddots & 1 \\ \vdots & \vdots & \vdots & \ddots & -q_1 \\ q_g & \vdots & \vdots & & -q_2 \\ & -q_g & \vdots & & \vdots \\ & & -q_g & & \vdots \\ & & & \ddots & \vdots \\ & & & & -q_g \end{bmatrix}}_{=:\mathcal{V}(\mathbf{Q},n)} \begin{bmatrix} 1 \\ a_1 \\ a_2 \\ \vdots \\ \vdots \\ a_n \end{bmatrix}. \tag{7.9}$$

We see that the parameters of $\bar{\mathbf{A}}(z^{-1})$ are determined in a simple way from those of $\mathbf{A}(z^{-1})$. Indeed, for each pair of n and $\mathbf{Q}(z^{-1})$, it is easy to obtain a matrix $\mathcal{V}(\mathbf{Q}, n) \in \mathbb{R}^{(n+g)\times(n+1)}$ defined in (7.9), from which we can form a matrix $\bar{\mathcal{V}}(\mathbf{Q}, n) \in \mathbb{R}^{(2n+g)\times(2n+1)}$ defined by

$$\bar{\mathcal{V}}(\mathbf{Q}, n) := \begin{bmatrix} \mathcal{V}(\mathbf{Q}, n) & \\ & I_n \end{bmatrix} \tag{7.10}$$

so that

$$\bar{\mathcal{V}}(\mathbf{Q}, n) \begin{bmatrix} 1 \\ a_1 \\ \vdots \\ a_n \\ b_1 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} \bar{a}_1 \\ \bar{a}_2 \\ \vdots \\ \bar{a}_{n+g} \\ b_1 \\ \vdots \\ b_n \end{bmatrix} =: \theta^*.$$

So we observe that the set of admissible parameters of (7.8) is given by

$$\tilde{\mathbf{\Theta}}_n := \left\{ \bar{\mathcal{V}}(\mathbf{Q}, n) \begin{bmatrix} 1 \\ \theta \end{bmatrix} : \ \theta \in \mathbf{\Theta}_n \right\} \subset \mathbb{R}^{2n+g}. \tag{7.11}$$

Using this notation and to facilitate analysis, the auxiliary plant (7.8) can now be put into regressor form:

$$\varepsilon(t+1) = \psi(t)^\top \theta^* + \bar{w}(t), \tag{7.12}$$

with $\bar{w}(t)$ as the inverse z-transform of $\bar{W}(z)$, $\psi(t) \in \mathbb{R}^{2n+g}$ defined as

$$\psi(t) := \begin{bmatrix} \varepsilon(t) \\ \varepsilon(t-1) \\ \vdots \\ \varepsilon(t-n-g+1) \\ v(t) \\ v(t-1) \\ \vdots \\ v(t-n+1) \end{bmatrix}$$

and $\theta^* \in \tilde{\mathbf{\Theta}}_n$. As in the case of the original plant (7.1), the order is not known, though it is known that $n \in \{1, 2, \ldots, \bar{n}\}$; hence, while the dimension of $\psi(t)$ clearly depends on n, to enhance readability this will not be made explicit.

Remark 7.4. *The new plant (7.12) is clearly overmodelled by g variables, which we consider to be a small price to pay to achieve our tracking objective.*

7.2.4 Uncertainty Sets

Since for every $n \in \{1, 2, \ldots, \bar{n}\}$, $\mathbf{\Theta}_n$ is compact, it follows that $\tilde{\mathbf{\Theta}}_n$ is as well; also, because of Assumptions 7.1 and 7.2 we see that for every $\bar{\theta} \in \tilde{\mathbf{\Theta}}_n$, the corresponding polynomials $z^{n+g}\bar{\mathbf{A}}(z^{-1})$ and $z^n\mathbf{B}(z^{-1})$ are coprime. Of course, if we were to replace $\tilde{\mathbf{\Theta}}_n$ by its convex hull, then those properties may fail to hold. This brings us to the following result. We show that for any $n \in \{1, 2, \ldots, \bar{n}\}$, $\tilde{\mathbf{\Theta}}_n$ can be approximated by a finite number of convex sets with desired properties.

Proposition 7.1. *For every $n \in \{1, 2, \ldots, \bar{n}\}$ and $\mu > 0$, there exist a finite number of convex, compact sets $\mathbf{\Theta}_n^i \subset \mathbb{R}^{2n+g}$ $(i = 1, 2, \ldots, m_n)$ that satisfy*

(i) $\tilde{\mathbf{\Theta}}_n \subset \bigcup_{i=1}^{m_n} \mathbf{\Theta}_n^i$,

(ii) for every $\theta^* \in \bigcup_{i=1}^{m_n} \mathbf{\Theta}_n^i$ there exists a $\bar{\theta}^* \in \tilde{\mathbf{\Theta}}_n$ that satisfy $\|\bar{\theta}^* - \theta^*\| \leq \mu$.

Furthermore, if $\mu > 0$ is sufficiently small, then we can choose the $\mathbf{\Theta}_n^i$'s to have additional property as well:

(iii) for every $\theta^* \in \bigcup_{i=1}^{m_n} \mathbf{\Theta}_n^i$, the corresponding pair of polynomials $z^{n+g}\bar{\mathbf{A}}(z^{-1})$ and $z^n\mathbf{B}(z^{-1})$ are coprime.

Proof of Proposition 7.1. Let $n \in \{1, 2, \ldots, \bar{n}\}$ be arbitrary. Fix $\mu > 0$. For every $x \in \tilde{\mathbf{\Theta}}_n$, let $\mathcal{O}_x \subset \mathbb{R}^{2n+g}$ denote the open ball of radius μ centered at x. Then

$$\left\{\mathcal{O}_x : x \in \tilde{\mathbf{\Theta}}_n\right\}$$

is an open cover of $\tilde{\mathbf{\Theta}}_n$, so by the Heine-Borel Theorem [66] there exist $x_1, x_2, \ldots, x_{m_n}$ so that

$$\tilde{\mathbf{\Theta}}_n \subset \bigcup_{i=1}^{m_n} \mathcal{O}_{x_i}.$$

If we set $\mathbf{\Theta}_n^i :=$ closure of $\mathcal{O}_{x_i}$, then (i) and (ii) of the required properties hold.

If $\bar{\mathbf{A}}(z^{-1})$ and $\mathbf{B}(z^{-1})$ are the corresponding polynomials associated with $x \in \mathbb{R}^{2n+g}$, then let $\mathsf{Syl}(x) \in \mathbb{R}^{(2n+g)\times(2n+g)}$ denote the Sylvester Matrix associated with the pair of polynomials (see [21, p. 482]). By the coprimeness requirement, we know that

$$\min_{\bar{\theta}^* \in \tilde{\mathbf{\Theta}}_n} \left|\det\left(\mathsf{Syl}(\bar{\theta}^*)\right)\right| > 0.$$

As $\det\Big(\mathsf{Syl}(x)\Big)$ is continuous in x, if a small enough $\mu > 0$ is used in the procedure (of the previous paragraph) to construct the $\mathbf{\Theta}_n^i$'s, we conclude that

$$\min_{\theta^* \in \mathbf{\Theta}_n^i} \left|\det\Big(\mathsf{Syl}(\theta^*)\Big)\right| > 0, \qquad i = 1, 2, \ldots, m_n. \qquad \blacksquare$$

In general, finding a set of m_n $\mathbf{\Theta}_n^i$'s which satisfy the desired properties of Proposition 7.1 for which m_n is small and $\mathbf{\Theta}_n^i$ has "nice[2] structure" is not easy. However, this is not the focus of this thesis. This *covering* problem is an open research problem—e.g. see [1], [15] and [33]. So at this point we assume that this process has been done for each $n \in \{1, 2, \ldots, \bar{n}\}$; we will show some examples on how to do this in Section 7.6.

To this end, the idea is to use a parameter estimator for each compact and convex set, and at each point in time we choose which one to use in constructing the control law. At this point, for every $n \in \{1, 2, \ldots, \bar{n}\}$ we have at hand m_n compact and convex parameter uncertainty sets (they can be disjoint or overlapping) that correspond to models of n^{th}-order plants; so for all possible plant orders $n \in \{1, 2, \ldots, \bar{n}\}$, we have compact and convex uncertainty sets:

$$\mathbf{\Theta}_1^1, \mathbf{\Theta}_1^2, \ldots, \mathbf{\Theta}_1^{m_1}, \mathbf{\Theta}_2^1, \mathbf{\Theta}_2^2, \ldots, \mathbf{\Theta}_2^{m_2}, \ldots \;\; \ldots, \mathbf{\Theta}_{\bar{n}}^1, \mathbf{\Theta}_{\bar{n}}^2 \ldots, \mathbf{\Theta}_{\bar{n}}^{m_{\bar{n}}},$$

yielding a total of

$$m := m_1 + m_2 + \cdots + m_{\bar{n}}$$

sets. For ease of notation, we re-label these sets as

$$\mathcal{S}_i \subset \mathbb{R}^{2n_i+g}, \quad i = 1, 2, \ldots m;$$

here $n_i \in \{1, 2, \ldots, \bar{n}\}$ represents the plant order of the associated model.

Now define the index set

$$\mathcal{I}^* := \{1, 2, \ldots, m\}.$$

[2] Nice in the sense that it is computationally easy to project onto it.

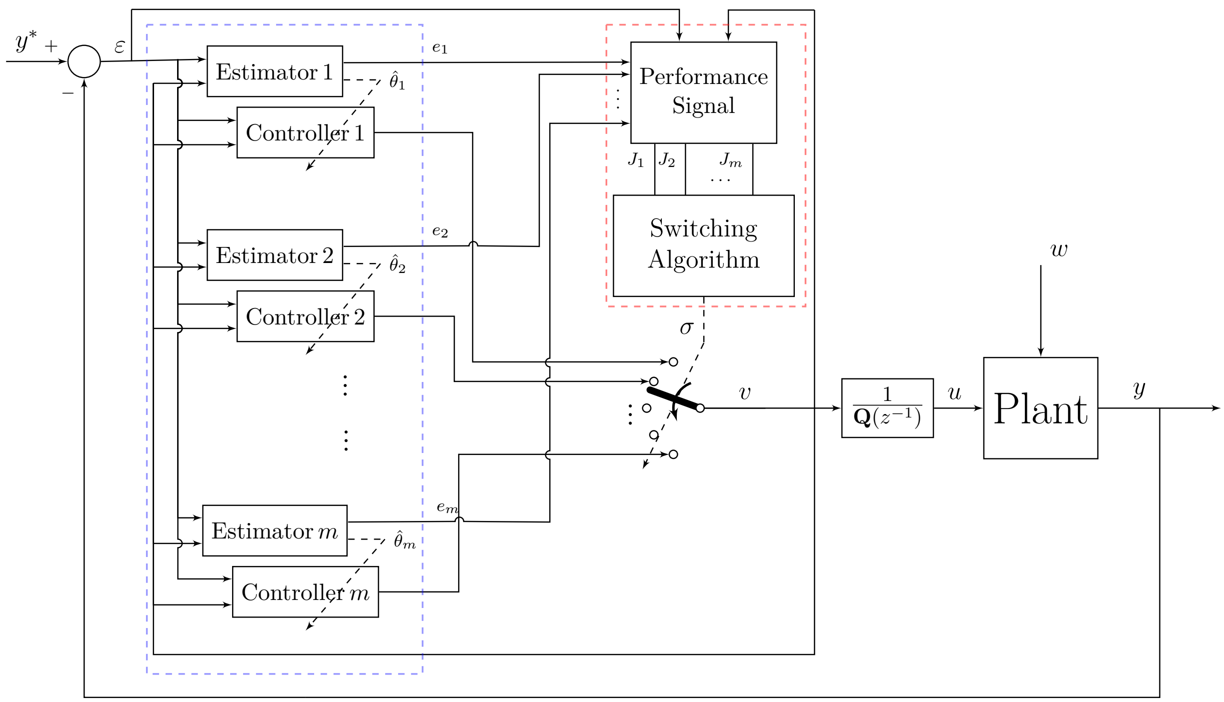


Figure 7.1: The Block diagram of the closed-loop system; enclosed inside the dashed boxes are the multiple estimators/controllers (blue), and the switching mechanism (red).

Similar to the previous chapters, for each $\theta^* \in \mathcal{S}_i$, $i = 1, 2, \ldots, m$, we define

$$i^*(\theta^*) = \min \{i \in \mathcal{I}^* : \theta^* \in \mathcal{S}_i\};$$

when there is no ambiguity, we will drop the argument and simply write i^*. Before proceeding, define

$$\bar{\mathbf{s}} := \max_{i \in \mathcal{I}^*} \|\mathcal{S}_i\|.$$

7.3 The Multi-Model Adaptive Controller

In this section we present the proposed adaptive controller; we discuss parameter multi-estimators, the associated switching control law, and the switching algorithm. The proposed controller is illustrated in the block diagram of the closed-loop system given in Figure 7.1.

7.3.1 Parameter Estimation

First, for each $i \in \mathcal{I}^*$, the corresponding regressor vector is defined by $\psi_i(t) \in \mathbb{R}^{2n_i+g}$:

$$\psi_i(t) := \begin{bmatrix} \varepsilon(t) \\ \varepsilon(t-1) \\ \vdots \\ \varepsilon(t-n_i-g+1) \\ v(t) \\ v(t-1) \\ \vdots \\ v(t-n_i+1) \end{bmatrix}.$$

So we know that the auxiliary plant (7.12) can be rewritten as

$$\varepsilon(t+1) = \psi_{i^*}(t)^\top \theta^* + \bar{w}(t).$$

Given an estimate $\hat{\theta}_{i^*}(t)$ at time t, we can now define the prediction error associated with this model by

$$e_{i^*}(t+1) := \varepsilon(t+1) - \psi_{i^*}(t)^\top \hat{\theta}_{i^*}(t). \tag{7.13}$$

Of course, we do not know the value of i^* but we define a prediction error for the i^{th} model in the natural way. Here we apply the same projection-algorithm based estimator (3.9) discussed in Chapter 3; to make our proof work we need to define the denominator in the estimation update law carefully. To proceed, we make the following observations:

1. We define the longest data vector by $\bar{\psi}(t) \in \mathbb{R}^{2\bar{n}+g}$:

$$\bar{\psi}(t) := \begin{bmatrix} \varepsilon(t) \\ \varepsilon(t-1) \\ \vdots \\ \varepsilon(t-\bar{n}-g+1) \\ v(t) \\ v(t-1) \\ \vdots \\ v(t-\bar{n}+1) \end{bmatrix}; \tag{7.14}$$

observe that $\|\psi_i(t)\| \leq \|\bar{\psi}(t)\|$ for all i.

2. By examining (7.13), we see that

$$e_{i^*}(t+1) = \psi_{i^*}(t)^\top \left[\theta^* - \hat{\theta}_{i^*}(t)\right] + \bar{w}(t), \tag{7.15}$$

which means that $|e_{i^*}(t+1)| \leq 2\bar{\mathbf{s}}\|\psi_{i^*}(t)\| + |\bar{w}(t)|$. So if $|e_{i^*}(t+1)| > 2\bar{\mathbf{s}}\|\psi_{i^*}(t)\|$, then the disturbance may be overwhelming the data, so we turn off the estimator.

To this end, define

$$e_i(t+1) := \varepsilon(t+1) - \psi_i(t)^\top \hat{\theta}_i(t), \qquad i \in \mathcal{I}^*, \tag{7.16}$$

and with $\delta \in (0, \infty]$, we define $\rho_i : \mathbb{Z} \mapsto \{0, 1\}$ by

$$\rho_i(t) := \begin{cases} 1 & \text{if } |e_i(t+1)| < (2\bar{\mathbf{s}} + \delta)\|\bar{\psi}(t)\| \\ 0 & \text{otherwise,} \end{cases} \tag{7.17}$$

which is used to determine when to turn off the algorithm. This leads to our proposed estimator: the estimator i updates are computed as follows:

$$\check{\theta}_i(t+1) = \hat{\theta}_i(t) + \rho_i(t)\frac{\psi_i(t)}{\|\bar{\psi}(t)\|^2}e_i(t+1) \tag{7.18a}$$

$$\hat{\theta}_i(t+1) = \operatorname*{Proj}_{\mathcal{S}_i}\left\{\check{\theta}_i(t+1)\right\}. \tag{7.18b}$$

Define the (correct) parameter estimation error $\tilde{\theta}_{i^*}(t) := \hat{\theta}_{i^*}(t) - \theta^*$. The following result lists properties of the estimation algorithm (7.18). These properties are direct result of applying Proposition 3.1 and 3.2 of Chapter 3, with the choices of $\phi(\cdot) = \psi_i(\cdot)$ and $\phi_m(\cdot) = \bar{\psi}(\cdot)$; note that the difference between the two vectors arises when the order is unknown, in which case $\psi_i(t)$ and $\bar{\psi}(t)$ may differ for some i.

Proposition 7.2. *For every $n \in \{1, 2, \ldots, \bar{n}\}$ and $\theta^* \in \tilde{\boldsymbol{\Theta}}_n$, $t_0 \in \mathbb{Z}$, $t_2 > t_1 \geq t_0$, $\bar{\psi}(t_0) \in \mathbb{R}^{2\bar{n}+g}$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ ($i \in \mathcal{I}^*$) and $w, y^* \in \boldsymbol{\ell}_\infty$, when the estimation algorithm in (7.18) is applied to the corresponding auxiliary plant (7.12), the following holds:*

1. *for* **every estimator** $i = 1, 2, \ldots, m$,
$$\|\hat{\theta}_i(t_2) - \hat{\theta}_i(t_1)\| \leq \sum_{j=t_1}^{t_2-1} \rho_i(j) \frac{|e_i(j+1)|}{\|\bar{\psi}(j)\|}. \tag{7.19}$$

2. *for the* **correct estimator** i^*,
$$\|\tilde{\theta}_{i^*}(t_2)\|^2 \leq \|\tilde{\theta}_{i^*}(t_1)\|^2 + \sum_{j=t_1}^{t_2-1} \rho_{i^*}(j) \left[-\frac{1}{2} \frac{e_{i^*}(j+1)^2}{\|\bar{\psi}(j)\|^2} + 2 \frac{\bar{w}(j)^2}{\|\bar{\psi}(j)\|^2} \right]. \tag{7.20}$$

7.3.2 The Switching Control Law

For each i, the parameter estimate $\hat{\theta}_i(t)$ is partitioned naturally as

$$\hat{\theta}_i(t) =: \begin{bmatrix} \hat{\bar{a}}_{i,1}(t) \\ \hat{\bar{a}}_{i,2}(t) \\ \vdots \\ \hat{\bar{a}}_{i,n_i+g}(t) \\ \hat{b}_{i,1}(t) \\ \hat{b}_{i,2}(t) \\ \vdots \\ \hat{b}_{i,n_i}(t) \end{bmatrix};$$

associated with these estimates are the polynomials

$$\begin{aligned} \hat{\bar{\mathbf{A}}}_i(t, z^{-1}) &= 1 - \hat{\bar{a}}_{i,1}(t) z^{-1} - \hat{\bar{a}}_{i,2}(t) z^{-2} \cdots - \hat{\bar{a}}_{i,n_i+g}(t) z^{-(n_i+g)}, \\ \hat{\mathbf{B}}_i(t, z^{-1}) &= \hat{b}_{i,1}(t) z^{-1} + \hat{b}_{i,2}(t) z^{-2} \cdots + \hat{b}_{i,n_i}(t) z^{-n_i}. \end{aligned}$$

Next we design a $(n_i + g)^{\text{th}}$-order *strictly* proper controller; we choose the following polynomials

$$\hat{\mathbf{L}}_i(t, z^{-1}) = 1 + \hat{l}_{i,1}(t) z^{-1} + \hat{l}_{i,2}(t) z^{-2} + \cdots + \hat{l}_{i,n_i}(t) z^{-n_i},$$

$$\hat{\mathbf{P}}_i(t, z^{-1}) = \hat{p}_{i,1}(t)z^{-1} + \hat{p}_{i,2}(t)z^{-2} + \cdots + \hat{p}_{i,n_i+g}(t)z^{-(n_i+g)}$$

so as to place all closed-loop poles at $z = 0$:

$$\hat{\bar{\mathbf{A}}}_i(t, z^{-1})\hat{\mathbf{L}}_i(t, z^{-1}) + \hat{\mathbf{B}}_i(t, z^{-1})\hat{\mathbf{P}}_i(t, z^{-1}) = 1. \tag{7.21}$$

Since $z^{n_i+g}\hat{\bar{\mathbf{A}}}_i(t, z^{-1})$ and $z^{n_i}\hat{\mathbf{B}}_i(t, z^{-1})$ are coprime by design, we know that there exist unique $\hat{\mathbf{L}}_i(t, z^{-1})$ and $\hat{\mathbf{P}}_i(t, z^{-1})$ which satisfy this equation; this entails solving a linear equation—see Appendix A. It is also easy to prove that the coefficients of $\hat{\mathbf{L}}_i(t, z^{-1})$ and $\hat{\mathbf{P}}_i(t, z^{-1})$ are analytic functions of $\hat{\theta}_i(t) \in \mathcal{S}_i$. For a suitable choice of $i \in \mathcal{I}^*$ at time t, we define the control input by

$$\hat{\mathbf{L}}_i(t-1, z^{-1})V(z) = -\hat{\mathbf{P}}_i(t-1, z^{-1})\mathcal{E}(z). \tag{7.22}$$

This can be written in terms of the data vector $\psi_i(t)$: to this end, we define the control gains $\hat{K}_i(t) \in \mathbb{R}^{2n_i+g}$ by

$$\hat{K}_i(t) := \begin{bmatrix} -\hat{p}_{i,1}(t) & -\hat{p}_{i,2}(t) & \cdots & -\hat{p}_{i,n_i+g}(t) & -\hat{l}_{i,1}(t) & -\hat{l}_{i,2}(t) & \cdots & -\hat{l}_{i,n_i}(t) \end{bmatrix} \tag{7.23}$$

so that (7.22) becomes

$$v(t) = \hat{K}_i(t-1)\psi_i(t-1).$$

We will use a switching signal $\sigma : \mathbb{Z} \to \mathcal{I}^*$ to denote the index i: $\sigma(t)$ denotes the index of the controller to use at time t.

In the previous chapter, we considered the problem of closed-loop stability (but not tracking) in the case of switching between two estimators of the same dimension. Unfortunately, the approach does not extend in a simple way to the case of $m > 2$ estimators, so we will need a different algorithm. However, our closed-loop system behavior will still, in large part, be determined by a time-varying matrix $\mathcal{A}_{\sigma(t)}(t) \in \mathbb{R}^{2\bar{n}+g}$ (see (7.39)); at all times this matrix will be deadbeat, i.e. all of its eigenvalues will be at zero. However, its product

$$\mathcal{A}_{\sigma(t)}(t) \times \mathcal{A}_{\sigma(t-1)}(t-1) \times \cdots \times \mathcal{A}_{\sigma(t_0)}(t_0), \qquad t \geq t_0$$

will not usually be deadbeat. A natural solution to this problem is to update the estimators every $2\bar{n} + g$ steps; the problem with this idea is that we end up with no information about $e_i(t+1)$ between the updates, so the closed-loop system is not amenable to analysis. So our solution procedure will need to be different: we update $\sigma(t)$ only every $N \geq 2\bar{n} + g$ steps; however, we keep the estimators running and the control gains updating; the aforementioned product of matrices is still not deadbeat, but it is close to being so, in a sense which will be apparent from the proof[3]. To this end, we define a sequence of switching times as follows: we initialize $\hat{t}_0 := t_0$ and then define

$$\hat{t}_\ell := \hat{t}_0 + \ell N, \qquad \ell \in \mathbb{N}.$$

So the switching signal is piecewise constant of the form

$$\sigma(t) = \sigma(\hat{t}_\ell), \qquad t \in [\hat{t}_\ell, \hat{t}_{\ell+1}), \quad \ell \in \mathbb{Z}^+; \tag{7.24}$$

the algorithm to compute $\sigma(\hat{t}_\ell)$ will be introduced shortly. We propose the choice of the control law

$$v(t) = \hat{K}_{\sigma(t-1)}(t-1)\psi_{\sigma(t-1)}(t-1), \qquad t > t_0, \tag{7.25}$$

which generates the auxiliary control input; this is combined with (7.6) to yield the plant control input

$$u(t) = v(t) + \sum_{j=1}^{g} q_j u(t-j), \qquad t > t_0. \tag{7.26}$$

What remains to be defined is the choice of the switching signal $\sigma(\hat{t}_\ell)$, which we will do in the next subsection.

7.3.3 The Switching Algorithm

With $N \in \mathbb{N}$, define the set of switching times by

$$\mathcal{T}_N := \left\{ \hat{t}_\ell \geq \hat{t}_0 \;:\; \hat{t}_\ell = \hat{t}_0 + \ell N,\ \ell \in \mathbb{Z}^+ \right\}. \tag{7.27}$$

To proceed, for each $i \in \mathcal{I}^*$ we define a performance signal $J_i : \mathcal{T}_N \to \mathbb{R}^+$ by

$$J_i(\hat{t}_\ell) := \sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+1}-1} \rho_i(j) \frac{|e_i(j+1)|}{\|\bar{\psi}(j)\|}, \qquad \ell \in \mathbb{Z}^+; \tag{7.28}$$

this quantity is an upper bound on the amount of change in $\hat{\theta}_i(t)$ on the interval $[\hat{t}_\ell, \hat{t}_{\ell+1})$. We may expect the estimator with the least amount of update to be the best one, which

[3] This is arguably a simplified version of the approach adopted in the previous chapter.

would lead to a switching signal of the form

$$\sigma(\hat{t}_{\ell+1}) = \underset{i \in \mathcal{I}^*}{\operatorname{argmin}}\ J_i(\hat{t}_\ell).$$

Although this rule works in every simulation that we have tried, a proof remains elusive; a potential problem is that the switching signal could oscillate between two bad choices, and never (or rarely) choose a "correct" one. Instead, we propose a different approach. At each switching time $\hat{t}_\ell$ we have an admissible set $\mathcal{I}(\hat{t}_\ell)$: we initialize $\mathcal{I}(\hat{t}_0) = \mathcal{I}^*$, and we obtain $\mathcal{I}(\hat{t}_{\ell+1})$ from $\mathcal{I}(\hat{t}_\ell)$ by removing all $j \in \mathcal{I}(\hat{t}_\ell)$ satisfying

$$J_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell) \le J_j(\hat{t}_\ell),$$

i.e. we keep all models in the admissible index set for which the performance signal is better (i.e. smaller) than the one we are currently using; clearly $j = \sigma(\hat{t}_\ell)$ satisfies this bound, but more j's may as well; if this results in $\mathcal{I}(\hat{t}_{\ell+1})$ being empty, then we **reset** $\mathcal{I}(\hat{t}_{\ell+1})$ to be $\mathcal{I}^*$. This *Switching Algorithm* is summarized as follows: with $\sigma(\hat{t}_0) = \sigma_0$ and $\mathcal{I}(\hat{t}_0) = \mathcal{I}^*$:

$$\hat{\mathcal{I}}(\hat{t}_\ell) = \left\{ i \in \mathcal{I}^* \ :\ J_i(\hat{t}_\ell) < J_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell) \right\}, \tag{7.29a}$$

$$\mathcal{I}(\hat{t}_{\ell+1}) = \begin{cases} \mathcal{I}^* & \text{if } \mathcal{I}(\hat{t}_\ell) \cap \hat{\mathcal{I}}(\hat{t}_\ell) = \varnothing \\ \mathcal{I}(\hat{t}_\ell) \cap \hat{\mathcal{I}}(\hat{t}_\ell) & \text{otherwise,} \end{cases} \tag{7.29b}$$

$$\sigma(\hat{t}_{\ell+1}) = \underset{i \in \mathcal{I}(\hat{t}_{\ell+1})}{\operatorname{argmin}}\ J_i(\hat{t}_\ell), \qquad \ell \in \mathbb{Z}^+. \tag{7.29c}$$

Remark 7.5. *In computing the* argmin *in the RHS of* (7.29c)*, it could very well that there are more values $i \in \mathcal{I}(\hat{t}_{\ell+1})$ which achieves the minimum. In such a case, we (somewhat arbitrarily) choose the smallest such index.*

Remark 7.6. *We define the* **index set reset times** *as those $\hat{t}_\ell$, $\ell \in \mathbb{Z}^+$, for which $\mathcal{I}(\hat{t}_\ell) = \mathcal{I}^*$.*

Remark 7.7. *The switching algorithm in* (7.29) *is an extended version of the one in* (4.33) *of Chapter 4 dealing with one-step-ahead adaptive control of a first-order plant, with $N = 1$ and $J_i(\hat{t}_\ell) = |e_i(t+1)|$. Note also that this algorithm can be applied to the approach of Chapter 5 dealing with nonlinear plants if more that two estimators were to be used (with $N = n$).*

In Figure 7.2 we provide an illustration of time steps, switching times, and index set reset times on the timeline. Now we present a desirable property of the switching algorithm (7.29).

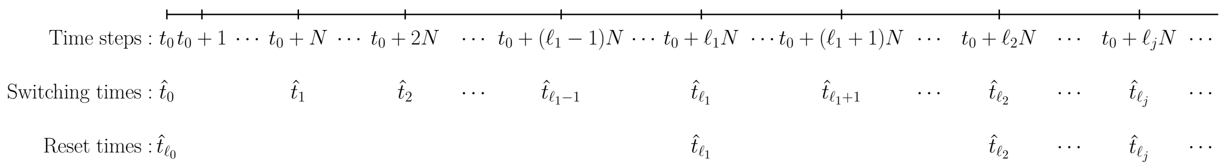


Figure 7.2: Illustration of the time instants, switching times, and the index set reset times.

Lemma 7.1. *Suppose that the adaptive controller (7.18), (7.23)–(7.25), and (7.27)–(7.29) is applied to the auxiliary plant (7.12). Then, for every $n \in \{1, 2, \ldots, \bar{n}\}$ and $\theta^* \in \tilde{\boldsymbol{\Theta}}_n$, $t_0 \in \mathbb{Z}$, $\sigma_0 \in \mathcal{I}^*$, $\bar{\psi}(t_0) \in \mathbb{R}^{2\bar{n}+g}$, $N \geq 1$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ ($i \in \mathcal{I}^*$) and $w, y^* \in \boldsymbol{\ell}_\infty$, if $\hat{t}_{\underline{\ell}}$ and $\hat{t}_{\bar{\ell}}$ are two consecutive index set reset times, then there exists a $\ell^* \in [\underline{\ell}, \bar{\ell})$ such that:*

$$J_{\sigma(\hat{t}_{\ell^*})}(\hat{t}_{\ell^*}) \leq J_{i^*}(\hat{t}_{\ell^*}). \tag{7.30}$$

Remark 7.8. *Lemma 7.1 says that, between every two index set resets, there is an interval of the form $[\hat{t}_{\ell^*}, \hat{t}_{\ell^*+1})$ for which the performance associated with the chosen index is equal to, or better than, that of the performance associated with the correct index.*

Proof of Lemma 7.1. Let $n \in \{1, 2, \ldots, \bar{n}\}$ and $\theta^* \in \tilde{\boldsymbol{\Theta}}_n$, $t_0 \in \mathbb{Z}$, $\sigma_0 \in \mathcal{I}^*$, $\bar{\psi}(t_0) \in \mathbb{R}^{2\bar{n}+g}$, $N \geq 1$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ ($i \in \mathcal{I}^*$), and $w, y^* \in \boldsymbol{\ell}_\infty$ be arbitrary. Let $\hat{t}_{\underline{\ell}}$ and $\hat{t}_{\bar{\ell}}$ be two consecutive reset times.

We prove (7.30) by contradiction; assume that

$$J_{\sigma(\hat{t}_j)}(\hat{t}_j) > J_{i^*}(\hat{t}_j), \quad \text{for all } j \in [\underline{\ell}, \bar{\ell}). \tag{7.31}$$

Then, according to (7.29a), we should have

$$i^* \in \hat{\mathcal{I}}(\hat{t}_j), \qquad j \in [\underline{\ell}, \bar{\ell}). \tag{7.32}$$

We know by the definition of index resets that for all $j \in (\underline{\ell}, \bar{\ell})$ we have $\mathcal{I}(\hat{t}_j) \neq \mathcal{I}^*$, which means that by (7.29b)

$$\mathcal{I}(\hat{t}_j) = \mathcal{I}(\hat{t}_{j-1}) \cap \hat{\mathcal{I}}(\hat{t}_{j-1}), \qquad j \in (\underline{\ell}, \bar{\ell});$$

then by induction we see that

$$\mathcal{I}(\hat{t}_j) = \mathcal{I}(\hat{t}_{\underline{\ell}}) \cap \hat{\mathcal{I}}(\hat{t}_{\underline{\ell}}) \cap \hat{\mathcal{I}}(\hat{t}_{\underline{\ell}+1}) \cap \cdots \cap \hat{\mathcal{I}}(\hat{t}_{j-2}) \cap \hat{\mathcal{I}}(\hat{t}_{j-1}), \qquad j \in (\underline{\ell}, \bar{\ell}).$$

But $\mathcal{I}(\hat{t}_{\underline{\ell}}) = \mathcal{I}^*$, so using (7.32) in the above, we see that

$$i^* \in \mathcal{I}(\hat{t}_j), \qquad j \in [\underline{\ell}, \bar{\ell}) \tag{7.33}$$

as well. So according to this and to (7.32) we have $i^* \in \mathcal{I}(\hat{t}_{\bar{\ell}-1}) \cap \hat{\mathcal{I}}(\hat{t}_{\bar{\ell}-1})$. However, we know by the definition of index resets and (7.29b) that $\mathcal{I}(\hat{t}_{\bar{\ell}-1}) \cap \hat{\mathcal{I}}(\hat{t}_{\bar{\ell}-1}) = \varnothing$, which is a contradiction, so it must be that (7.31) does not hold. ∎

As discussed throughout the thesis about similar results, in the above we do not make any claim that $\theta^* \in \mathcal{S}_{\sigma(t)}$ at any time; it only makes an indirect statement about the size of the prediction error. It turns out that this is enough to ensure that desired closed-loop behavior is attained.

7.4 The Main Result

We will define a vector $\bar{\phi}(t) \in \mathbb{R}^{2(\bar{n}+g)}$

$$\bar{\phi}(t) := \begin{bmatrix} y(t) \\ y(t-1) \\ \vdots \\ y(t-\bar{n}-g+1) \\ u(t) \\ u(t-1) \\ \vdots \\ u(t-\bar{n}-g+1) \end{bmatrix}$$

to serve as the "plant state"; while this is longer than what is needed for a minimal state representation of (7.1), the choice will facilitate our analysis. Recall that the vectors ψ_i, $i \in \mathcal{I}^*$, and $\bar{\psi}$ contain values of the tracking error and the auxiliary control input, while the vector $\bar{\phi}$ contains values of the plant input and output. Before proceeding, it is convenient to define a weighted sum of past values of y^*:

$$\tilde{y}^*(t) := \sum_{j=0}^{\bar{n}+g-1} |y^*(t-j)|. \tag{7.34}$$

Theorem 7.1. *Suppose that the adaptive controller consisting of the parameter estimators* (7.18)*, control gains* (7.23)*, switching signal* (7.24) *with switching times* (7.27)*, performance signal* (7.28)*, switching algorithm* (7.29)*, and control law* (7.25) *and* (7.26)*, is applied to the plant* (7.1)*. Then for every* $\lambda \in (0,1)$*,* $\delta \in (0,\infty]$ *and* $N \geq 2\bar{n}+g$*, there exists a constant* $\gamma > 0$ *so that, for every* $n \in \{1,2,\ldots,\bar{n}\}$ *and* $\theta \in \mathbf{\Theta}_n$*,* $t_0 \in \mathbb{Z}$*,* $\bar{\phi}(t_0) \in \mathbb{R}^{2(\bar{n}+g)}$*,* $\sigma_0 \in \mathcal{I}^*$*,* $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ *(*$i \in \mathcal{I}^*$*), and* $w, y^* \in \boldsymbol{\ell}_\infty$*,*

i) the following bound holds:

$$\|\bar{\phi}(t)\| \leq \gamma\lambda^{t-\tau}\|\bar{\phi}(\tau)\| + \gamma\sum_{j=\tau}^{t-1}\lambda^{t-1-j}(|w(j)| + |\tilde{y}^*(j+1)|), \qquad t > \tau \geq t_0; \qquad (7.35)$$

ii) if $\mathbf{Q}(z^{-1})Y^*(z) = 0$ *and* $\mathbf{Q}(z^{-1})W(z) = 0$*, then* $y(t) \to y^*(t)$ *exponentially fast, in the sense that*

$$|\varepsilon(t)| \leq \gamma\lambda^{t-t_0}\left(\|\bar{\phi}(t_0)\| + \|y^*\|_\infty\right), \qquad t \geq t_0.$$

Remark 7.9. *The above result shows that the closed-loop system experiences linear-like behavior. There is a* uniform *exponential decay bound on the effect of the initial condition, and a convolution bound on the effect of the exogenous signals. This implies that the system has a bounded gain (from* w *and* y^* *to* y*) in every p-norm; in particular, for* $p = \infty$ *we see from the above bound that*

$$\begin{aligned}\|\bar{\phi}(t)\| &\leq \gamma\lambda^{t-t_0}\|\bar{\phi}(t_0)\| + \frac{\gamma}{1-\lambda}\sup_{j\in[t_0,t)}(|w(j)| + |\tilde{y}^*(j+1)|)\\ &\leq \frac{\gamma(\bar{n}+g)}{1-\lambda}\left(\lambda^{t-t_0}\|\bar{\phi}(t_0)\| + \|w\|_\infty + \|y^*\|_\infty\right), \qquad t \geq t_0.\end{aligned}$$

Hence, if $w, y^* \in \boldsymbol{\ell}_\infty$*, then* $y, u \in \boldsymbol{\ell}_\infty$*, so* $\varepsilon, v, e_i\,(i \in \mathcal{I}^*)$ *lie in* $\boldsymbol{\ell}_\infty$ *as well.*

We emphasize here that we are able to prove Theorem 7.1 using a switching control law **without** assuming that the switching stops. As far as the author knows, only a few similar results are found in the literature, e.g. the Supervisory Control approach of Morse [51], although convolution bounds are not proven there.

Remark 7.10. *Observe that the stability result (with a slightly different controller) of Theorem 6.1 of the previous chapter is subsumed by the result of Theorem 7.1 by putting* $\mathbf{Q}(z^{-1}) = 1$ *(i.e.* $g = 0$*).*

Proving Theorem 7.1 requires two steps:

- First, we analyze the adaptive control system to obtain a desired bound on the key quantity $\bar{\psi}(t)$ (which consists of present and past values of the tracking error and the auxiliary input), which plays a key role in the auxiliary model (7.12), the parameter estimator (7.18), and the control law (7.25); this requires a careful analysis of the closed-loop system.
- Second, we use linear system theory to translate the bound on $\bar{\psi}$ to a bound on $\bar{\phi}$ (which consists of present and past values of the plant's input and output).

To enhance readability and to focus the reader's attention on the most important aspects of the approach, we will present the first part in the form of a Proposition.

Proposition 7.3 (**Main Proposition**)**.** *Suppose that the adaptive controller (7.18), (7.23)–(7.25), and (7.27)–(7.29) is applied to the auxiliary plant (7.12). Then, for every $\lambda \in (0,1)$, $\delta \in (0,\infty]$ and $N \geq 2\bar{n}+g$, there exists a constant $c>0$ so that, for every $n \in \{1,2,\ldots,\bar{n}\}$ and $\theta^* \in \tilde{\boldsymbol{\Theta}}_n$, $t_0 \in \mathbb{Z}$, $\bar{\psi}(t_0) \in \mathbb{R}^{2\bar{n}+g}$, $\sigma_0 \in \mathcal{I}^*$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ ($i \in \mathcal{I}^*$), and $w, y^* \in \boldsymbol{\ell}_\infty$, the following holds*

$$\|\bar{\psi}(t)\| \leq c\lambda^{t-\tau}\|\bar{\psi}(\tau)\| + \sum_{j=\tau}^{t-1} c\lambda^{t-j-1}|\bar{w}(j)|, \quad t > \tau \geq t_0. \tag{7.36}$$

Before presenting the proof of this Proposition, we need a crude bound on the closed-loop behavior.

Lemma 7.2. *Suppose that the adaptive controller (7.18), (7.23)–(7.29) is applied to the plant (7.1). Then for every $p \geq 0$, there exist constants $\bar{c}_1, \bar{c}_2 \geq 1$ so that, for every $n \in \{1,2,\ldots,\bar{n}\}$ and $\theta \in \boldsymbol{\Theta}_n$, $t_0 \in \mathbb{Z}$, $t \geq t_0$, $N \geq 1$, $\sigma_0 \in \mathcal{I}^*$, $\bar{\phi}(t_0) \in \mathbb{R}^{2(\bar{n}+g)}$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ ($i \in \mathcal{I}^*$), and $w, y^* \in \boldsymbol{\ell}_\infty$, the following hold:*

i) $\|\bar{\psi}(t+p)\| \leq \bar{c}_1\|\bar{\psi}(t)\| + \bar{c}_1 \sum_{j=0}^{p-1} |\bar{w}(t+j)|$.

ii) $\|\bar{\phi}(t+p)\| \leq \bar{c}_2\|\bar{\phi}(t)\| + \bar{c}_2 \sum_{j=0}^{p-1} (|w(t+j)| + |\tilde{y}^*(t+j)|)$.

Proof. See Appendix B. ■

Proof of Proposition 7.3:

Fix $\lambda \in (0,1)$, $\delta \in (0,\infty]$ and $N \geq 2\bar{n}+g$. Let $n \in \{1,2,\ldots,\bar{n}\}$, $\theta^* \in \tilde{\mathbf{\Theta}}_n$, $t_0 \in \mathbb{Z}$, $\bar{\psi}(t_0) \in \mathbb{R}^{2(\bar{n}+g)}$, $\sigma_0 \in \mathcal{I}^*$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ $(i \in \mathcal{I}^*)$, and $w, y^* \in \boldsymbol{\ell}_\infty$ be arbitrary. We denote the sequence of index set reset times by $\hat{t}_{\ell_0}, \hat{t}_{\ell_1}, \hat{t}_{\ell_2}, \ldots$ (see Fig. 7.2).

Step 1: Obtain a state-space model describing $\bar{\psi}(t)$ for $t \in [\hat{t}_\ell, \hat{t}_{\ell+1})$.

It will be convenient for analysis to have all of the parameter estimates and controller gains to be of the same length. To this end, we will pad $\hat{\theta}_i(t)$ and $\hat{K}_i(t)$ by zeros in the appropriate locations: we define $\hat{\Theta}_i(t) \in \mathbb{R}^{2\bar{n}+g}$ by

$$\hat{\Theta}_i(t) := \begin{bmatrix} \hat{\hat{a}}_{i,1}(t) & \hat{\hat{a}}_{i,2}(t) & \cdots & \hat{\hat{a}}_{i,n_i+g}(t) & \mathbf{0}^\top_{\bar{n}-n_i} & \hat{b}_{i,1}(t) & \hat{b}_{i,2}(t) & \cdots & \hat{b}_{i,n_i}(t) & \mathbf{0}^\top_{\bar{n}-n_i} \end{bmatrix}^\top$$

and $\hat{\hat{K}}_i(t) \in \mathbb{R}^{2\bar{n}+g}$ by

$$\hat{\hat{K}}_i(t) :=$$
$$\begin{bmatrix} -\hat{p}_{i,1}(t) & -\hat{p}_{i,2}(t) & \cdots & -\hat{p}_{i,n_i+g}(t) & \mathbf{0}^\top_{\bar{n}-n_i} & -\hat{l}_{i,1}(t) & -\hat{l}_{i,2}(t) & \cdots & -\hat{l}_{i,n_i}(t) & \mathbf{0}^\top_{\bar{n}-n_i} \end{bmatrix};$$

so by definition of the prediction error (7.16) and from the control law in (7.25) we have

$$\begin{aligned} \varepsilon(t+1) &= \hat{\theta}_{\sigma(t)}(k)^\top \psi_{\sigma(t)}(k) + e_{\sigma(t)}(t+1) \\ &= \hat{\Theta}_{\sigma(t)}(t)^\top \bar{\psi}(t) + e_{\sigma(t)}(t+1), \end{aligned} \tag{7.37}$$

$$\begin{aligned} v(t+1) &= \hat{K}_{\sigma(t)}(t)\psi_{\sigma(t)}(t) \\ &= \hat{\hat{K}}_{\sigma(t)}(t)\bar{\psi}(t). \end{aligned} \tag{7.38}$$

Next, define the matrix $\mathcal{A}_i(t) \in \mathbb{R}^{(2\bar{n}+g)\times(2\bar{n}+g)}$ by

$$\mathcal{A}_i(t) := \begin{bmatrix} \hat{\Theta}_i(t)^\top \\ \begin{bmatrix} I_{\bar{n}+g-1} & \mathbf{0}_{(\bar{n}+g-1)\times(\bar{n}+1)} \end{bmatrix} \\ \hat{\hat{K}}_i(t) \\ \begin{bmatrix} \mathbf{0}_{(\bar{n}-1)\times(\bar{n}+g)} & I_{\bar{n}-1} & \mathbf{0}_{\bar{n}-1} \end{bmatrix} \end{bmatrix} = \begin{bmatrix} & & & \hat{\Theta}_i(t)^\top & & & & \\ 1 & 0 & \cdots & 0 & 0 & \cdots & \cdots & 0 \\ & \ddots & & \vdots & \vdots & \cdots & \cdots & \vdots \\ & & 1 & 0 & 0 & \cdots & \cdots & 0 \\ & & & \hat{\hat{K}}_i(t) & & & & \\ 0 & \cdots & \cdots & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \cdots & \cdots & \vdots & & \ddots & & \vdots \\ 0 & \cdots & \cdots & 0 & & & 1 & 0 \end{bmatrix}; \tag{7.39}$$

using (7.21) it is easy to verify that the characteristic equation of the matrix $\mathcal{A}_i(t)$ (for frozen time t) satisfies

$$\begin{aligned}\det\Big(zI_{2\bar{n}+g} - \mathcal{A}_i(t)\Big) &= z^{2\bar{n}+g}[\hat{\mathbf{A}}_i(t,z^{-1})\hat{\mathbf{L}}_i(t,z^{-1}) + \hat{\mathbf{B}}_i(t,z^{-1})\hat{\mathbf{P}}_i(t,z^{-1})] \\ &= z^{2\bar{n}+g}.\end{aligned}$$

This means that for every $i \in \mathcal{I}^*$, for each time t the matrix $\mathcal{A}_i(t)$ has all of its eigenvalues at zero. Also define $B_1 := \mathbf{e}_1 \in \mathbb{R}^{2\bar{n}+g}$ and

$$\Delta_i(t) := \rho_i(t)\frac{e_i(t+1)}{\|\bar{\psi}(t)\|^2}B_1\bar{\psi}(t)^\top \tag{7.40}$$

so we have

$$B_1 e_i(t+1) = \Delta_i(t)\bar{\psi}(t) + B_1\underbrace{[1-\rho_i(t)]e_i(t+1)}_{=:\eta_i(t)}.$$

From (7.37) and (7.38), the fact that the switching signal is constant on $[\hat{t}_\ell, \hat{t}_{\ell+1})$, and the definition of $\bar{\psi}$, we have that

$$\begin{aligned}\bar{\psi}(t+1) &= \mathcal{A}_{\sigma(t)}(t)\bar{\psi}(t) + B_1 e_{\sigma(t)}(t+1) \\ &= \mathcal{A}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\bar{\psi}(t) + \Big[\mathcal{A}_{\sigma(t)}(t) - \mathcal{A}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\Big]\bar{\psi}(t) + B_1 e_{\sigma(t)}(t+1) \\ &= \mathcal{A}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\bar{\psi}(t) + \Big[\mathcal{A}_{\sigma(\hat{t}_\ell)}(t) - \mathcal{A}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\Big]\bar{\psi}(t) + B_1 e_{\sigma(\hat{t}_\ell)}(t+1) \\ &= \mathcal{A}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\bar{\psi}(t) + \Big[\mathcal{A}_{\sigma(\hat{t}_\ell)}(t) - \mathcal{A}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell) + \Delta_{\sigma(\hat{t}_\ell)}(t)\Big]\bar{\psi}(t) + B_1\eta_{\sigma(\hat{t}_\ell)}(t), \\ &\qquad t \in [\hat{t}_\ell, \hat{t}_{\ell+1}),\ \ell \in \mathbb{Z}^+.\end{aligned} \tag{7.41}$$

Step 2: Obtain a bound on $\|\bar{\psi}(\hat{t}_{\ell+1})\|$ in terms of $\|\bar{\psi}(\hat{t}_\ell)\|$.

We now are going to analyze the key equation (7.41) in detail; we make the following observations. For $t \in [\hat{t}_\ell, \hat{t}_{\ell+1})$, we have $\mathcal{A}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell) \in \mathbb{R}^{(2\bar{n}+g)\times(2\bar{n}+g)}$ to be a constant matrix with all eigenvalues equal to zero; since $N \geq 2\bar{n}+g$,

$$\Big[\mathcal{A}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\Big]^{\hat{t}_{\ell+1}-\hat{t}_\ell} = \Big[\mathcal{A}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\Big]^N = \mathbf{0}. \tag{7.42}$$

Next, note that for $t \in [\hat{t}_\ell, \hat{t}_{\ell+1})$ we have that the dimension of $\hat{\theta}_{\sigma(t)}(t)$ is constant; by utilizing part (1) of Proposition 7.2 to provide a bound on the difference between parameter estimates at two different point in time, the fact that the controller gains are analytic functions of the parameter estimates, and (7.39), we conclude that there exists a constant c_1 such that

$$\begin{aligned}
\left\|\mathcal{A}_{\sigma(\hat{t}_\ell)}(t)-\mathcal{A}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\right\| &\le \left\|\hat{\Theta}_{\sigma(\hat{t}_\ell)}(t)-\hat{\Theta}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\right\|+\left\|\hat{\bar{K}}_{\sigma(\hat{t}_\ell)}(t)-\hat{\bar{K}}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\right\| \\
&= \left\|\hat{\theta}_{\sigma(\hat{t}_\ell)}(t)-\hat{\theta}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\right\|+\left\|\hat{K}_{\sigma(\hat{t}_\ell)}(t)-\hat{K}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\right\| \\
&\le \left\|\hat{\theta}_{\sigma(\hat{t}_\ell)}(t)-\hat{\theta}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\right\|+c_1\left\|\hat{\theta}_{\sigma(\hat{t}_\ell)}(t)-\hat{\theta}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\right\| \\
&= (1+c_1)\left\|\hat{\theta}_{\sigma(\hat{t}_\ell)}(t)-\hat{\theta}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\right\| \\
&\le (1+c_1)\sum_{j=\hat{t}_\ell}^{t-1}\rho_{\sigma(\hat{t}_\ell)}(j)\frac{|e_{\sigma(\hat{t}_\ell)}(j+1)|}{\|\bar{\psi}(j)\|}, \qquad t\in[\hat{t}_\ell,\hat{t}_{\ell+1}),\ \ell\in\mathbb{Z}^+.
\end{aligned} \tag{7.43}$$

From (7.40) we obtain

$$\|\Delta_{\sigma(\hat{t}_\ell)}(t)\| = \rho_{\sigma(\hat{t}_\ell)}(t)\frac{|e_{\sigma(\hat{t}_\ell)}(t+1)|}{\|\bar{\psi}(t)\|}. \tag{7.44}$$

So from (7.43), (7.44) and definition of the performance signal (7.28), there exists a constant c_2 so that for all $t\in[\hat{t}_\ell,\hat{t}_{\ell+1})$:

$$\begin{aligned}
\left\|\mathcal{A}_{\sigma(\hat{t}_\ell)}(t)-\mathcal{A}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)+\Delta_{\sigma(\hat{t}_\ell)}(t)\right\| &\le \left\|\mathcal{A}_{\sigma(\hat{t}_\ell)}(t)-\mathcal{A}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\right\|+\|\Delta_{\sigma(\hat{t}_\ell)}(t)\| \\
&\le (1+c_1)\left(\sum_{j=\hat{t}_\ell}^{t-1}\rho_{\sigma(\hat{t}_\ell)}(j)\frac{|e_{\sigma(\hat{t}_\ell)}(j+1)|}{\|\bar{\psi}(j)\|}\right)+ \\
&\qquad \rho_{\sigma(\hat{t}_\ell)}(t)\frac{|e_{\sigma(\hat{t}_\ell)}(t+1)|}{\|\bar{\psi}(t)\|} \\
&\le \underbrace{(1+c_1)}_{=:c_2}\sum_{j=\hat{t}_\ell}^{t}\rho_{\sigma(\hat{t}_\ell)}(j)\frac{|e_{\sigma(\hat{t}_\ell)}(j+1)|}{\|\bar{\psi}(j)\|} \\
&\le c_2\sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+1}-1}\rho_{\sigma(\hat{t}_\ell)}(j)\frac{|e_{\sigma(\hat{t}_\ell)}(j+1)|}{\|\bar{\psi}(j)\|} \\
&= c_2 J_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell), \qquad t\in[\hat{t}_\ell,\hat{t}_{\ell+1}),\ \ell\in\mathbb{Z}^+.
\end{aligned} \tag{7.45}$$

To proceed, we need a bound on $\eta_i(t)$.

Claim 7.1. *There exists a c_3 such that for all $i \in \mathcal{I}^*$:*

$$|\eta_i(t)| \leq c_3|\bar{w}(t)|, \qquad t \geq t_0. \tag{7.46}$$

Proof of Claim 7.1. If $\rho_i(t) = 1$, then $\eta_i(t) = 0$. If $\rho_i(t) = 0$, then $\eta_i(t) = e_i(t+1)$ and from the estimator definition

$$|e_i(t+1)| \geq (2\bar{\mathbf{s}} + \delta)\|\bar{\psi}(t)\|;$$

but notice that

$$\begin{aligned} e_i(t+1) &= \psi_{i^*}(t)^\top \theta^* - \psi_i(t)^\top \hat{\theta}_i(t) + \bar{w}(t) \\ \Rightarrow |e_i(t+1)| &\leq \|\psi_{i^*}(t)\|\|\theta^*\| + \|\psi_i(t)\|\|\hat{\theta}_i(t)\| + |\bar{w}(t)| \\ &\leq \|\mathcal{S}_{i^*}\|\|\psi_{i^*}(t)\| + \|\mathcal{S}_i\|\|\psi_i(t)\| + |\bar{w}(t)| \\ &\leq \bar{\mathbf{s}}\left(\|\psi_{i^*}(t)\| + \|\psi_i(t)\|\right) + |\bar{w}(t)| \\ &\leq 2\bar{\mathbf{s}}\|\bar{\psi}(t)\| + |\bar{w}(t)|. \end{aligned}$$

Combining the above two statements:

$$2\bar{\mathbf{s}}\|\bar{\psi}(t)\| + |\bar{w}(t)| \geq (2\bar{\mathbf{s}} + \delta)\|\bar{\psi}(t)\| \Rightarrow \|\bar{\psi}(t)\| \leq \frac{1}{\delta}|\bar{w}(t)|;$$

this means that $|e_i(t+1)| \leq \frac{2\bar{\mathbf{s}}}{\delta}|\bar{w}(t)| + |\bar{w}(t)|$, so define $c_3 := \frac{2\bar{\mathbf{s}}}{\delta} + 1$. $\square$

Now we return to analyzing the key equation (7.41). Solving for $\bar{\psi}(\hat{t}_{\ell+1})$ yields

$$\begin{aligned} \bar{\psi}(\hat{t}_{\ell+1}) = \left[\mathcal{A}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\right]^{\hat{t}_{\ell+1}-\hat{t}_\ell} \bar{\psi}(\hat{t}_\ell) + \sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+1}-1} \left[\mathcal{A}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\right]^{\hat{t}_{\ell+1}-j-1} \Bigg(\Bigg[\mathcal{A}_{\sigma(\hat{t}_\ell)}(j) - \mathcal{A}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell) + \\ \Delta_{\sigma(\hat{t}_\ell)}(j)\Bigg]\bar{\psi}(j) + B_1\eta_{\sigma(\hat{t}_\ell)}(j)\Bigg). \end{aligned} \tag{7.47}$$

It follows from the compactness of the $\mathcal{S}_i$'s that $\left\|\left[\mathcal{A}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\right]^j\right\|$, $j = 0, 1 \ldots, N-1$, is bounded above by a constant, which we label c_4. So incorporating this and the observations of (7.42), (7.45) and (7.46) into (7.47), we obtain

$$\|\bar{\psi}(\hat{t}_{\ell+1})\| \leq c_4 \sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+1}-1} \left(c_2 J_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\|\bar{\psi}(j)\| + c_3|\bar{w}(j)|\right)$$

$$= c_4 c_2 J_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell) \sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+1}-1} \|\bar{\psi}(j)\| + c_4 c_3 \sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+1}-1} |\bar{w}(j)|. \tag{7.48}$$

It follows from Lemma 7.2 (applied for $p = 1, 2, \ldots, N-1$) that there exists a constant c_5 so that the following holds:

$$\sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+1}-1} \|\bar{\psi}(j)\| \leq c_5 \|\bar{\psi}(\hat{t}_\ell)\| + c_5 \sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+1}-2} |\bar{w}(j)|; \tag{7.49}$$

so substituting (7.49) into (7.48) it follows that there exists a constant c_6 so that for all $\ell \in \mathbb{Z}^+$:

$$\begin{aligned} \|\bar{\psi}(\hat{t}_{\ell+1})\| &\leq c_4 c_2 J_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell) \left(c_5 \|\bar{\psi}(\hat{t}_\ell)\| + c_5 \sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+1}-2} |\bar{w}(j)| \right) + c_4 c_3 \sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+1}-1} |\bar{w}(j)| \\ &\leq c_6 J_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell) \|\bar{\psi}(\hat{t}_\ell)\| + c_6 \left(1 + J_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\right) \sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+1}-1} |\bar{w}(j)|. \end{aligned} \tag{7.50}$$

Step 3: Obtain a bound on $\bar{\psi}$ between index set reset times which depend solely on J_{i^*}.

Let $\hat{t}_{\ell_j}$ be an arbitrary reset time. From Lemma 7.1 we know that there exists an ℓ^* satisfying $\ell_j \leq \ell^* < \ell_{j+1}$ such that

$$J_{\sigma(\hat{t}_{\ell^*})}(\hat{t}_{\ell^*}) \leq J_{i^*}(\hat{t}_{\ell^*}). \tag{7.51}$$

So here we will analyze the closed-loop behavior for

$$[\hat{t}_{\ell_j}, \hat{t}_{\ell_{j+1}}) = [\hat{t}_{\ell_j}, \hat{t}_{\ell^*}) \cup [\hat{t}_{\ell^*}, \hat{t}_{\ell^*+1}) \cup [\hat{t}_{\ell^*+1}, \hat{t}_{\ell_{j+1}}).$$

The behavior on $[\hat{t}_{\ell^*}, \hat{t}_{\ell^*+1})$ can be analyzed by combining (7.50) with (7.51):

$$\|\bar{\psi}(\hat{t}_{\ell^*+1})\| \leq c_6 J_{i^*}(\hat{t}_{\ell^*}) \|\bar{\psi}(\hat{t}_{\ell^*})\| + c_6 \left(1 + J_{i^*}(\hat{t}_{\ell^*})\right) \sum_{j=\hat{t}_{\ell^*}}^{\hat{t}_{\ell^*+1}-1} |\bar{w}(j)|. \tag{7.52}$$

The behavior for the other two intervals can be analyzed by utilizing Lemma 7.2. To this end, from the switching algorithm, it is clear that

$$\ell_{j+1} - \ell_j \le m, \qquad j \in \mathbb{Z}^+;$$

so $\hat{t}_{\ell_{j+1}} - \hat{t}_{\ell_j} \le Nm$, which means that $\hat{t}_{\ell^*} - \hat{t}_{\ell_j} \le Nm$ and $\hat{t}_{\ell_{j+1}} - \hat{t}_{\ell^*+1} \le Nm$ as well; so we can utilize Lemma 7.2 with $p \le Nm$: in particular, there exists a constant c_7 so that

$$\|\bar{\psi}(\hat{t}_{\ell_{j+1}})\| \le c_7 \|\bar{\psi}(\hat{t}_{\ell^*+1})\| + c_7 \sum_{j=\hat{t}_{\ell^*+1}}^{\hat{t}_{\ell_{j+1}}-1} |\bar{w}(j)| \tag{7.53}$$

and

$$\|\bar{\psi}(\hat{t}_{\ell^*})\| \le c_7 \|\bar{\psi}(\hat{t}_{\ell_j})\| + c_7 \sum_{j=\hat{t}_{\ell_j}}^{\hat{t}_{\ell^*}-1} |\bar{w}(j)|. \tag{7.54}$$

Now define

$$\alpha(\ell_j) := \max_{\tau \in [\ell_j, \ell_{j+1})} J_{i^*}(\hat{t}_\tau) \tag{7.55}$$

and

$$\tilde{w}(j) := \sum_{q=\hat{t}_{\ell_j}}^{\hat{t}_{\ell_{j+1}}-1} |\bar{w}(q)|;$$

by combining these bounds with (7.52) we conclude that there exists a constant c_8 such that

$$\|\bar{\psi}(\hat{t}_{\ell_{j+1}})\| \le c_8 \alpha(\ell_j) \|\bar{\psi}(\hat{t}_{\ell_j})\| + c_8 (1 + \alpha(\ell_j)) \tilde{w}(j), \qquad j \in \mathbb{Z}^+. \tag{7.56}$$

Step 4: Analyze the first-order difference inequality (7.56).

We will analyze (7.56) to obtain a bound on the closed-loop behavior of the system for the whole time horizon. The first step is to analyze the square sum of $\alpha(\cdot)$ over an interval; from the definition of $\alpha(\cdot)$ and the Cauchy-Schwarz property, we have for all $j_2 > j_1 \ge 0$:

$$\sum_{q=j_1}^{j_2-1} \alpha(\ell_q)^2 = \sum_{q=j_1}^{j_2-1} \left(\max_{p \in [\ell_q, \ell_{q+1})} J_{i^*}(\hat{t}_p) \right)^2$$

$$
\begin{aligned}
&\leq \sum_{q=j_1}^{j_2-1} \left(\sum_{p=\ell_q}^{\ell_{q+1}-1} J_{i^*}(\hat{t}_p) \right)^2 \\
&\leq \sum_{q=j_1}^{j_2-1} \left([\ell_{q+1} - \ell_q] \sum_{p=\ell_q}^{\ell_{q+1}-1} J_{i^*}(\hat{t}_p)^2 \right) \\
&\leq m \sum_{q=j_1}^{j_2-1} \sum_{p=\ell_q}^{\ell_{q+1}-1} J_{i^*}(\hat{t}_p)^2 \\
&= m \sum_{q=j_1}^{j_2-1} \sum_{p=\ell_q}^{\ell_{q+1}-1} \left(\sum_{\tau=\hat{t}_p}^{\hat{t}_{p+1}-1} \rho_{i^*}(\tau) \frac{|e_{i^*}(\tau+1)|}{\|\bar{\psi}(\tau)\|} \right)^2 \\
&\leq m \sum_{q=j_1}^{j_2-1} \sum_{p=\ell_q}^{\ell_{q+1}-1} \left([\hat{t}_{p+1} - \hat{t}_p] \sum_{\tau=\hat{t}_p}^{\hat{t}_{p+1}-1} \rho_{i^*}(\tau) \frac{|e_{i^*}(\tau+1)|^2}{\|\bar{\psi}(\tau)\|^2} \right) \\
&= Nm \sum_{q=j_1}^{j_2-1} \sum_{p=\ell_q}^{\ell_{q+1}-1} \sum_{\tau=\hat{t}_p}^{\hat{t}_{p+1}-1} \rho_{i^*}(\tau) \frac{|e_{i^*}(\tau+1)|^2}{\|\bar{\psi}(\tau)\|^2} \\
&= Nm \sum_{\tau=\hat{t}_{\ell_{j_1}}}^{\hat{t}_{\ell_{j_2}}-1} \rho_{i^*}(\tau) \frac{|e_{i^*}(\tau+1)|^2}{\|\bar{\psi}(\tau)\|^2}.
\end{aligned}
$$

Using the above together with part (2) of Proposition 7.2, we obtain, for any $p > q \geq 0$:

$$
\sum_{j=q}^{p-1} \alpha(\ell_j)^2 \leq 2Nm \|\tilde{\theta}_{i^*}(\hat{t}_{\ell_q})\|^2 + 4Nm \sum_{j=\hat{t}_{\ell_q}}^{\hat{t}_{\ell_p}-1} \rho_{i^*}(j) \frac{|\bar{w}(j)|^2}{\|\bar{\psi}(j)\|^2}. \tag{7.57}
$$

To proceed, let $\tau \geq t_0$ be arbitrary. We define

$$
\lambda_1 := \frac{\lambda^{Nm}}{\max\{1, c_8\}} \in (0,1).
$$

We now partition the timeline into two parts: one in which $\bar{w}(\cdot)$ is small versus $\bar{\psi}(\cdot)$ and one where it is not; with

$$
\nu := \left(\frac{\lambda_1}{4Nm} \right)^2, \tag{7.58}
$$

we define

$$S_{good} = \left\{ j \geq \tau : \bar{\psi}(j) \neq 0 \text{ and } \frac{|\bar{w}(j)|^2}{\|\bar{\psi}(j)\|^2} < \nu \right\},$$

$$S_{bad} = \left\{ j \geq \tau : \bar{\psi}(j) = 0 \text{ or } \frac{|\bar{w}(j)|^2}{\|\bar{\psi}(j)\|^2} \geq \nu \right\};$$

clearly $\{j \in \mathbb{Z} : \; j \geq \tau\} = S_{good} \cup S_{bad}$. Notice that if $\bar{w} = 0$, then S_{good} could be the whole timeline $[\tau, \infty)$. We can clearly define a (possibly infinite) sequence of intervals of the form $[k_l, k_{l+1})$ which satisfy:
(i) $k_0 = \tau$ serves as the initial instant of the first interval;
(ii) $[k_l, k_{l+1})$ either belongs to S_{good} or S_{bad}; and
(iii) if $k_{l+1} \neq \infty$ and $[k_l, k_{l+1})$ belongs to S_{good} then $[k_{l+1}, k_{l+2})$ belongs to S_{bad}, and vice versa.

Now we analyze the behavior during each interval.

Step 4.1: $[k_l, k_{l+1}) \subset S_{bad}$.

Let $j \in [k_l, k_{l+1})$ be arbitrary. In this case $\frac{|\bar{w}(j)|^2}{\|\bar{\psi}(j)\|^2} \geq \nu$ or $\|\bar{\psi}(j)\| = 0$; in either case

$$\|\bar{\psi}(j)\| \leq \tfrac{1}{\sqrt{\nu}} |\bar{w}(j)|.$$

Also, applying Lemma 7.2 for one step, there exists a constant c_9 so that

$$\begin{aligned} \|\bar{\psi}(j+1)\| &\leq c_9 \|\bar{\psi}(j)\| + c_9 |\bar{w}(j)| \\ &\leq c_9 \tfrac{1}{\sqrt{\nu}} |\bar{w}(j)| + c_9 |\bar{w}(j)|, \quad j \in [k_l, k_{l+1}). \end{aligned}$$

This, in turn, implies that

$$\|\bar{\psi}(j)\| \leq \begin{cases} \frac{1}{\sqrt{\nu}} |\bar{w}(j)| & j = k_l \\ c_9 (\frac{1}{\sqrt{\nu}} + 1) |\bar{w}(j-1)| & j = k_l + 1, \ldots, k_{l+1}. \end{cases} \tag{7.59}$$

Step 4.2: $[k_l, k_{l+1}) \subset S_{good}$.

First suppose that $k_{l+1} - k_l \leq 2Nm$; then by Lemma 7.2 it can be easily proven that there exists a constant c_{10} so that

$$\|\bar{\psi}(t)\| \leq c_{10} \lambda^{t-k_l} \|\bar{\psi}(k_l)\| + c_{10} \sum_{j=k_l}^{t-1} \lambda^{t-j-1} |\bar{w}(j)|, \qquad t \in [k_l, k_{l+1}].$$

Now suppose that $k_{l+1} - k_l > 2Nm$. This means that in the interval of interest, namely $[k_l, k_{l+1})$, there are at least two reset times: there exist $q < p$ so that

$$k_l \leq \hat{t}_{\ell_q} < \hat{t}_{\ell_p} \leq k_{l+1};$$

in fact, there may be many choices of q and p; so let $\underline{q}$ be the smallest such q and $\bar{p}$ be the largest such p. To proceed, observe that $\|\bar{\psi}(j)\| \neq 0$ and $\frac{|\bar{w}(j)|^2}{\|\bar{\psi}(j)\|^2} < \nu$. Using this bound which holds on $[k_l, k_{l+1})$, together with the fact that $\|\tilde{\theta}_{i^*}(\hat{t}_{\ell_q})\| \leq 2\|\mathcal{S}_{i^*}\| \leq 2\bar{\mathbf{s}}$, we rewrite (7.57) to yield

$$\begin{aligned}\sum_{j=q}^{p-1} \alpha(\ell_j)^2 &\leq 8Nm\bar{\mathbf{s}}^2 + 4Nm(\hat{t}_{\ell_p} - \hat{t}_{\ell_q})\nu \\ &= 8Nm\bar{\mathbf{s}}^2 + 4N^2m(\ell_p - \ell_q)\nu \\ &\leq 8Nm\bar{\mathbf{s}}^2 + 4N^2m^2(p-q)\nu, \qquad \underline{q} \leq q < p \leq \bar{p}. \end{aligned} \tag{7.60}$$

From the definition of ν in (7.58), the above bound can be simplified to

$$\sum_{j=q}^{p-1} \alpha(\ell_j)^2 \leq 8Nm\bar{\mathbf{s}}^2 + (p-q)\frac{\lambda_1^2}{4}, \qquad \underline{q} \leq q < p \leq \bar{p}. \tag{7.61}$$

Now we will analyze the difference inequality in (7.56). First, we use (7.61) to bound the second occurrence of α in (7.56); from (7.61) we see that $\alpha(\ell_j) \leq \sqrt{8Nm\bar{\mathbf{s}}^2 + \lambda_1^2/4} \leq \sqrt{8Nm\bar{\mathbf{s}}^2 + 1} =: c_{11}$, $\underline{q} \leq j \leq \bar{p}$. So we can rewrite (7.56) to yield

$$\|\bar{\psi}(\hat{t}_{\ell_{j+1}})\| \leq c_8\alpha(\ell_j)\|\bar{\psi}(\hat{t}_{\ell_j})\| + \underbrace{c_8(1+c_{11})}_{=:c_{12}}\tilde{w}(j), \qquad \underline{q} \leq j \leq \bar{p}. \tag{7.62}$$

We now proceed to solve the above difference inequality; we will utilize the inequality of arithmetic and geometric means.

Claim 7.2. *There exists a constant $\gamma_1 > 1$ such that*

$$\prod_{j=q}^{p-1} \alpha(\ell_j) \leq \gamma_1 \lambda_1^{p-q}, \qquad \underline{q} \leq q < p \leq \bar{p}. \tag{7.63}$$

Proof of Claim 7.2. Let $q, p \in \mathbb{Z}^+$ be arbitrary such that $\underline{q} \leq q < p \leq \bar{p}$. By the fact that

$\alpha(\ell_j) \geq 0$, we obtain

$$\prod_{j=q}^{p-1} \alpha(\ell_j) \leq \left[\frac{1}{p-q}\sum_{j=q}^{p-1} \alpha(\ell_j)^2\right]^{\frac{p-q}{2}}.$$

Using (7.61) we obtain

$$\prod_{j=q}^{p-1} \alpha(\ell_j) \leq \left[\frac{8Nm\bar{\mathbf{s}}^2}{p-q} + \frac{\lambda_1^2}{4}\right]^{\frac{p-q}{2}}.$$

So it is enough to prove that there exists a constant γ_1 so that

$$\left(\underbrace{\left[\frac{8Nm\bar{\mathbf{s}}^2}{j} + \frac{\lambda_1^2}{4}\right]^{\frac{1}{2}}}_{=:\beta(j)}\right)^j \leq \gamma_1 \lambda_1^j, \quad j > 0.$$

We can easily show that with $\bar{\jmath} := 16Nm \times \left\lceil \left(\frac{\bar{\mathbf{s}}}{\lambda_1}\right)^2 \right\rceil$, we have

$$\frac{8Nm\bar{\mathbf{s}}^2}{\bar{\jmath}} \leq \frac{\lambda_1^2}{2},$$

which means that

$$\beta(j)^j \leq \lambda_1^j \leq 1, \qquad j \geq \bar{\jmath}.$$

So if we define $\gamma_1 := \max\left\{1, \left(\frac{\beta(1)}{\lambda_1}\right)^{\bar{\jmath}}\right\}$, then the result in (7.63) is proven. □

Using the bound in (7.63) and the definition of λ_1 we obtain

$$\begin{aligned}\prod_{j=q}^{p-1} [c_8 \alpha(\ell_j)] &\leq \gamma_1 \lambda_1^{p-q} c_8^{p-q}, \\ &\leq \gamma_1 \lambda^{Nm(p-q)}, \qquad \underline{q} \leq q < p \leq \bar{p}.\end{aligned} \tag{7.64}$$

We can now proceed to solve (7.62) iteratively; if we use the bound in (7.64), we see that

$$\|\bar{\psi}(\hat{t}_{\ell_p})\| \le \gamma_1 \lambda^{Nm(p-q)} \|\bar{\psi}(\hat{t}_{\ell_q})\| + \sum_{j=q}^{p-1} \gamma_1 c_{12} \left(\lambda^{Nm}\right)^{p-j-1} \tilde{w}(j), \qquad \underline{q} \le q < p \le \bar{p}.$$

We can now use Lemma 7.2 (for no more than Nm steps at a time):

- to provide a bound on $\|\bar{\psi}(t)\|$ between consecutive index set reset times, i.e. between $\hat{t}_{\ell_j}$ and $\hat{t}_{\ell_{j+1}}$;
- to provide a bound on $\|\bar{\psi}(t)\|$ on the beginning part of the interval $[k_l, k_{l+1})$, until we get to the first admissible index set reset time $\hat{t}_{\ell_{\underline{q}}}$;
- to provide a bound on $\|\bar{\psi}(t)\|$ on the last part of the interval $[k_l, k_{l+1})$, after the last admissible index set reset time $\hat{t}_{\ell_{\bar{p}}}$.

After simplification, we conclude that there exists a constant $\gamma_2 \ge c_{10}$ so that

$$\|\bar{\psi}(t)\| \le \gamma_2 \lambda^{t-k_l} \|\bar{\psi}(k_l)\| + \gamma_2 \sum_{j=k_l}^{t-1} \lambda^{t-j-1} |\bar{w}(j)|, \qquad t \in [k_l, k_{l+1}]. \tag{7.65}$$

Step 4.3: Combine the bounds on S_{good} and S_{bad}.

Now we combine Step 4.1 and Step 4.2 into a general bound on $\bar{\psi}$: we glue the bounds of Step 4.1 and Step 4.2 together. Define

$$\bar{\gamma} := \max\left\{\gamma_2, c_9(1+\tfrac{1}{\sqrt{\nu}}), \gamma_2 c_9(1+\tfrac{1}{\sqrt{\nu}})\right\}.$$

Claim 7.3. *The following bound holds:*

$$\|\bar{\psi}(t)\| \le \bar{\gamma} \lambda^{t-\tau} \|\bar{\psi}(\tau)\| + \sum_{j=\tau}^{t-1} \bar{\gamma} \lambda^{t-j-1} |\bar{w}(j)|, \;\; t \ge \tau. \tag{7.66}$$

Proof of the Claim 7.3. If $[k_0, k_1) = [\tau, k_1) \subset S_{good}$, then (7.66) is true for $t \in [k_0, k_1]$ by (7.65). If $[k_0, k_1) \subset S_{bad}$, then from (7.59) we obtain

$$\|\bar{\psi}(j)\| \le \begin{cases} \|\bar{\psi}(\tau)\| & j = k_0 = \tau \\ c_9(1+\frac{1}{\sqrt{\nu}})|\bar{w}(j-1)| & j = k_0+1, \dots, k_1. \end{cases}$$

which means that (7.66) holds on $[k_0, k_1]$ for this case as well.

We now use induction: suppose that (7.66) is true for $t \in [k_0, k_l]$; we need to prove it holds for $t \in (k_l, k_{l+1}]$ as well. If $k \in [k_l, k_{l+1}) \subset S_{bad}$, then from (7.59) we see that

$$\|\bar{\psi}(j)\| \leq c_9(1 + \tfrac{1}{\sqrt{\nu}})|\bar{w}(j-1)|, \qquad j = k_l + 1, k_l + 2, \ldots, k_{l+1},$$

which means (7.66) holds on $(k_l, k_{l+1}]$. On the other hand, if $[k_l, k_{l+1}) \subset S_{good}$, then $k_l - 1 \in S_{bad}$; from (7.59) we have that

$$\|\bar{\psi}(k_l)\| \leq c_9(1 + \tfrac{1}{\sqrt{\nu}})|\bar{w}(k_l - 1)|.$$

Using (7.65) to analyze the behavior on $[k_l, k_{l+1}]$, we have

$$\begin{aligned}\|\bar{\psi}(k)\| &\leq \gamma_2 \lambda^{k-k_l}[c_9(1 + \tfrac{1}{\sqrt{\nu}})|\bar{w}(k_l - 1)|] + \sum_{j=k_l}^{k-1} \gamma_2 \lambda^{k-j-1}|\bar{w}(j)|, \\ &\leq \bar{\gamma} \sum_{j=k_l-1}^{k-1} \lambda^{k-j-1}|\bar{w}(j)|, \qquad k \in [k_l, k_{l+1}],\end{aligned} \tag{7.67}$$

which implies that (7.66) holds. □

Finally, as $\tau \geq t_0$ is arbitrary, it follows that the proof of Proposition 7.3 is concluded. ■

Now we proceed to prove the main result in Theorem 7.1.

Proof of Theorem 7.1. Fix $\lambda \in (0, 1)$, $\delta \in (0, \infty]$ and $N \geq 2\bar{n} + g$. Let $n \in \{1, 2, \ldots, \bar{n}\}$, $\theta \in \mathbf{\Theta}_n$, $t_0 \in \mathbb{Z}$, $\bar{\phi}(t_0) \in \mathbb{R}^{2(\bar{n}+g)}$, $\sigma_0 \in \mathcal{I}^*$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ $(i \in \mathcal{I}^*)$, and $w, y^* \in \boldsymbol{\ell}_\infty$ be arbitrary.

Let $\tau \geq t_0$ be arbitrary. Then applying the adaptive controller to the associated auxiliary plant (7.12), by Proposition 7.3 there exists a constant c such that

$$\|\bar{\psi}(t)\| \leq c\lambda^{t-\tau}\|\bar{\psi}(\tau)\| + \sum_{j=\tau}^{t-1} c\lambda^{t-j-1}|\bar{w}(j)|, \quad t > \tau. \tag{7.68}$$

Step 1: Finding a bound on $y(\cdot)$.

It turns out to be easy to leverage (7.68) to provide a desired bound on the output y and its past values. Using the definition of $\bar{\psi}$ given in (7.14), the definition of $\tilde{y}^*$ given in

(7.34), and the fact that $y(t) = \varepsilon(t) + y^*(t)$, with a change in the indexes it follows from (7.68) that

$$\left\| \begin{bmatrix} y(t) \\ y(t-1) \\ \vdots \\ y(t-\bar{n}-g+1) \end{bmatrix} \right\| \leq c\lambda^{t-\tau-\bar{n}-g} \|\bar{\psi}(\tau+\bar{n}+g)\| + |\tilde{y}^*(t)| + \sum_{j=\tau+\bar{n}+g}^{t-1} c\lambda^{t-j-1} |\bar{w}(j)|,$$
$$t > \tau + \bar{n} + g; \tag{7.69}$$

the reason for choosing a starting time of $\tau + \bar{n} + g$ rather than τ will become clear later in the proof. Although this provides a bound on the top part of $\bar{\phi}$, the quantity on the RHS differs from that on the RHS of the desired bound (7.36). We will now proceed to get a similar kind of bound on the bottom part of $\bar{\phi}$, after which we convert the quantity in the RHS to one of the desired form.

Step 2: Finding a bound on $u(\cdot)$.

Now we will derive a desirable bound for plant input u and its past values. The analysis here is more involved than the one for finding the bound on y.

We start by constructing a state-space model of the plant (7.1); we will choose one of dimension n which is in controllable canonical form:

$$x(t+1) = Ax(t) + Bu(t) \tag{7.70a}$$
$$y(t) = Cx(t) + w(t-1). \tag{7.70b}$$

Corresponding to our coprimeness and compactness assumptions, the set of all such (A, B, C) triples lies in a compact set.

From the plant control input defined in (7.26), we can view u as the output of the following g^{th}-order system. In fact, with $Q \in \mathbb{R}^{g\times g}$ defined by

$$Q := \begin{bmatrix} q_1 & q_2 & \cdots & q_{g-1} & q_g \\ 1 & 0 & \cdots & & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix},$$

$C_c := \mathbf{e}_1^\top \in \mathbb{R}^g$, $B_c := \mathbf{e}_1 \in \mathbb{R}^g$ and

$$\xi(t) := \begin{bmatrix} u(t) \\ u(t-1) \\ \vdots \\ u(t-g+1) \end{bmatrix},$$

it follows that

$$\xi(t+1) = Q\xi(t) + B_c v(t+1), \tag{7.71a}$$

$$u(t) = C_c \xi(t). \tag{7.71b}$$

Since $\varepsilon(t) = y(t) - y^*(t)$, by combining (7.70) with (7.71) we obtain the augmented $(n+g)^{\text{th}}$-order state-space system

$$\begin{bmatrix} x(t+1) \\ \xi(t+1) \end{bmatrix} = \underbrace{\begin{bmatrix} A & BC_c \\ \mathbf{0}_{g\times n} & Q \end{bmatrix}}_{=:\bar{A}} \underbrace{\begin{bmatrix} x(t) \\ \xi(t) \end{bmatrix}}_{=:\bar{x}(t)} + \underbrace{\begin{bmatrix} 0 \\ B_c \end{bmatrix}}_{=:\bar{B}} v(t+1) \tag{7.72a}$$

$$\varepsilon(t) = \underbrace{\begin{bmatrix} C & \mathbf{0}_g^\top \end{bmatrix}}_{=:\bar{C}} \begin{bmatrix} x(t) \\ \xi(t) \end{bmatrix} + w(t-1) - y^*(t). \tag{7.72b}$$

Since (7.70) is controllable and observable and does not have common zeros with the eigenvalues of Q (Assumption 7.2), it follows that $(\bar{C}, \bar{A})$ is observable; hence, there exists a unique $\bar{H}$ such that the eigenvalues of $\bar{A} + \bar{H}\bar{C}$ are all zero and it is well-known that $\bar{H}$ is a continuous function of $\bar{A}$ and $\bar{C}$. Now rewrite (7.72) as

$$\bar{x}(t+1) = \left[\bar{A} + \bar{H}\bar{C}\right]\bar{x}(t) - \bar{H}\varepsilon(t) + \bar{B}v(t+1) + \bar{H}w(t-1) - \bar{H}y^*(t);$$

noting that $[\bar{A} + \bar{H}\bar{C}]^j = 0$ for all $j \geq n+g$, the solution of the above equation is

$$\bar{x}(t) = \sum_{j=1}^{n+g} \left[\bar{A} + \bar{H}\bar{C}\right]^{j-1} \left(\left[-\bar{H}\varepsilon(t-j) + \bar{B}v(t-j+1)\right] + \bar{H}[w(t-j-1) - y^*(t-j)] \right),$$
$$t \geq \tau + n + g.$$

We now want to analyze the behavior of $\bar{x}$ in terms of $\bar{\psi}$. But

$$\varepsilon(t-j) = \mathbf{e}_j^\top \bar{\psi}(t-1), \qquad j = 1, 2, \ldots, n+g,$$

$$v(t-j+1) = \begin{cases} \mathbf{e}_{\bar{n}+g+j}^{\top}\bar{\psi}(t), & j = 1,2,\dots,n \\ \mathbf{e}_{2\bar{n}+g}^{\top}\bar{\psi}(t+n-j), & j = n+1,\dots,n+g, \end{cases}$$

and $\xi(t)$ is part of $\bar{x}(t)$, then there exists a constant γ_1 so that

$$\begin{aligned} \|\xi(t)\| &\leq \gamma_1 \sum_{j=0}^{g} \|\bar{\psi}(t-j)\| + \gamma_1 \sum_{j=t-n-g}^{t-1} (|w(j-1)| + |y^*(j)|) \\ &\leq \gamma_1 \sum_{j=0}^{g} \|\bar{\psi}(t-j)\| + \gamma_1 \sum_{j=t-n-g}^{t-1} |w(j-1)| + \gamma_1 |\tilde{y}^*(t-1)|, \quad t \geq \tau+n+g. \end{aligned} \tag{7.73}$$

We will use (7.68) to provide a bound on $\bar{\psi}(\cdot)$'s: changing indexes (with $k \geq 0$, we replace t by $t-k$ and τ by $\tau+\bar{n}+g$) we have

$$\|\bar{\psi}(t-k)\| \leq c\lambda^{t-k-\tau-\bar{n}-g}\|\bar{\psi}(\tau+\bar{n}+g)\| + \sum_{j=\tau+\bar{n}+g}^{t-k-1} c\lambda^{t-k-j-1}|\bar{w}(j)|,$$
$$t > \tau+k+\bar{n}+g, \quad k \geq 0.$$

If we use this bound in (7.73) for $k = 0,1,\dots,g$, and simplify, then we see that there exists a constant γ_2 so that

$$\begin{aligned} \|\xi(t)\| \leq{}& \gamma_2\lambda^{t-\tau-\bar{n}-g}\|\bar{\psi}(\tau+\bar{n}+g)\| + \\ & \gamma_2 \sum_{j=\tau+\bar{n}+g}^{t-1} \lambda^{t-j-1}|\bar{w}(j)| + \gamma_1 \sum_{j=t-n-g}^{t-1} |w(j-1)| + \gamma_1|\tilde{y}^*(t-1)|, \qquad t \geq \tau+\bar{n}+2g. \end{aligned} \tag{7.74}$$

While $\xi(t)$ contains $u(t), u(t-1), \dots, u(t-g+1)$, the vector $\bar{\phi}(t)$ contains $u(t), u(t-1), \dots, u(t-\bar{n}-g+1)$. However, we see that

$$\left\| \begin{bmatrix} u(t) \\ u(t-1) \\ \vdots \\ u(t-\bar{n}-g+1) \end{bmatrix} \right\| \leq \sum_{j=0}^{\bar{n}} \|\xi(t-j)\|. \tag{7.75}$$

Now we apply the bound in (7.74) to obtain bounds on $\xi(t), \xi(t-1), \dots, \xi(t-\bar{n})$; if we substitute them into (7.75) and simplify, we conclude that there exists a constant γ_3 such that

$$\left\| \begin{bmatrix} u(t) \\ u(t-1) \\ \vdots \\ u(t-\bar{n}-g+1) \end{bmatrix} \right\| \le \gamma_3 \lambda^{t-\tau-\bar{n}-g} \|\bar{\psi}(\tau+\bar{n}+g)\| + \gamma_3 \sum_{j=0}^{\bar{n}} |\tilde{y}^*(t-j-1)| + \\ \gamma_3 \sum_{j=t-\bar{n}-n-g}^{t-1} |w(j-1)| + \gamma_3 \sum_{j=\tau+\bar{n}+g}^{t-1} \lambda^{t-j-1} |\bar{w}(j)|, \\ t \ge \tau + 2\bar{n} + 2g. \tag{7.76}$$

Step 3: Obtaining the desired bound.

Now we provide the desired bound on the full vector $\bar{\phi}$: using (7.69) and (7.76) and the fact that $n + \bar{n} \le 2\bar{n}$, there exists a constant γ_4 so that

$$\left\|\bar{\phi}(t)\right\| \le \gamma_4 \lambda^{t-\tau-\bar{n}-g} \|\bar{\psi}(\tau+\bar{n}+g)\| + \gamma_4 \sum_{j=0}^{\bar{n}+1} |\tilde{y}^*(t-j)| + \\ \gamma_4 \sum_{j=t-2\bar{n}-g}^{t-1} |w(j-1)| + \gamma_4 \sum_{j=\tau+\bar{n}+g}^{t-1} \lambda^{t-j-1} |\bar{w}(j)|, \qquad t \ge \tau + 2\bar{n} + 2g. \tag{7.77}$$

The bound in (7.77) looks very similar to the desired bound, except for the use of $\bar{w}$, the use of $\bar{\psi}$ instead of $\bar{\phi}$, and the starting point; we will now deal with these issues in that order. We first replace $\bar{w}(\cdot)$ with its constituent signals; we see from the definition of $\bar{w}$ in (7.7) that for each time j, $\bar{w}(j)$ is a weighted sum of $w(j), w(j-1), \dots, w(j-g)$ and $y^*(j+1), y^*(j), y^*(t-1) \dots, y^*(j-n-g+1)$. So substituting this into (7.77) and simplifying, we see that there exists a constant γ_5 such that

$$\left\|\bar{\phi}(t)\right\| \le \gamma_5 \lambda^{t-\tau} \|\bar{\psi}(\tau+\bar{n}+g)\| + \gamma_5 \sum_{j=0}^{\bar{n}+1} |\tilde{y}^*(t-j)| + \\ \gamma_5 \sum_{j=t-2\bar{n}-g}^{t-1} |w(j-1)| + \gamma_5 \sum_{j=\tau}^{t-1} \lambda^{t-j-1} (|w(j)| + |y^*(j+1)|), \qquad t \ge \tau + 2\bar{n} + 2g. \tag{7.78}$$

Using the definition of $\tilde{y}^*$ given in (7.34), and changing the range of t slightly in order to ensure that the index on w is greater than or equal to τ, then it follows that there exists a constant γ_6 such that

$$\left\|\bar{\phi}(t)\right\| \leq \gamma_6 \lambda^{t-\tau} \|\bar{\psi}(\tau+\bar{n}+g)\| + \gamma_6 \sum_{j=\tau}^{t-1} \lambda^{t-j-1}(|w(j)| + |y^*(j+1)|),$$
$$t \geq \tau + 2\bar{n} + 2g + 1. \tag{7.79}$$

We now want to have a bound on $\|\bar{\psi}(\tau+\bar{n}+g)\|$ in terms of $\|\bar{\phi}(\tau+\bar{n}+g)\|$ and then obtain a bound in terms of $\|\bar{\phi}(\tau)\|$. From the definition of $\bar{\psi}(t)$ and definition of the auxiliary input $v(\cdot)$, we see that $\bar{\psi}(t)$ consists of $y(t)-y^*(t), y(t-1)-y^*(t-1), \ldots, y(t-\bar{n}-g+1)-y^*(t-\bar{n}-g+1)$ and weighted sums of $u(t), u(t-1), \ldots, u(t-\bar{n}-g+1)$; so by the definition of $\bar{\phi}(\cdot)$ we observe that there exists a constant γ_7 such that

$$\|\bar{\psi}(t)\| \leq \gamma_7 \|\bar{\phi}(t)\| + \sum_{j=0}^{\bar{n}+g-1} |y^*(t-j)|. \tag{7.80}$$

Applying this to (7.79), we see that there exists a constant γ_8 so that

$$\left\|\bar{\phi}(t)\right\| \leq \gamma_8 \lambda^{t-\tau} \|\bar{\phi}(\tau+\bar{n}+g)\| + \gamma_8 \sum_{j=\tau}^{t-1} \lambda^{t-j-1}(|w(j)| + |y^*(j+1)|),$$
$$t \geq \tau + 2\bar{n} + 2g + 1. \tag{7.81}$$

We now obtain a bound on $\|\bar{\phi}(\tau+\bar{n}+g)\|$ in terms of $\|\bar{\phi}(\tau)\|$ utilizing Lemma 7.2 with $p = \bar{n}+g$. We conclude that there exists γ_9 such that

$$\left\|\bar{\phi}(t)\right\| \leq \gamma_9 \lambda^{t-\tau} \|\bar{\phi}(\tau)\| + \gamma_9 \sum_{j=\tau}^{\tau+\bar{n}+g-1} |\tilde{y}^*(j)| + \gamma_9 \sum_{j=\tau}^{t-1} \lambda^{t-j-1}(|w(j)| + |y^*(j+1)|),$$
$$t \geq \tau + 2\bar{n} + 2g + 1. \tag{7.82}$$

We can obtain a bound in terms of only $\tilde{y}^*$ using the definition in (7.34); after manipulation we conclude that there exists a constant γ_{10} so that

$$\left\|\bar{\phi}(t)\right\| \leq \gamma_{10} \lambda^{t-\tau} \|\bar{\phi}(\tau)\| + \gamma_{10} \sum_{j=\tau}^{t-1} \lambda^{t-j-1}(|w(j)| + |\tilde{y}^*(j+1)|),$$
$$t \geq \tau + 2\bar{n} + 2g + 1. \tag{7.83}$$

Finally, we will use Lemma 7.2 to get the desired bound for $t \in [\tau, \tau+2\bar{n}+2g]$: there exists a constant γ_{11} such that

$$\|\bar{\phi}(t)\| \le \gamma_{11}\lambda^{t-\tau}\|\bar{\phi}(\tau)\| + \sum_{j=\tau}^{t-1}\gamma_{11}\lambda^{t-j-1}(|w(j)| + |\tilde{y}^*(j+1)|),$$
$$\tau + 2\bar{n} + 2g \ge t > \tau. \tag{7.84}$$

Combining this bound with the bound in (7.83) we obtain the desired bound for the whole timeline: there exists a constant $\bar{\gamma}$ so that

$$\left\|\bar{\phi}(t)\right\| \le \bar{\gamma}\lambda^{t-\tau}\|\bar{\phi}(\tau)\| + \bar{\gamma}\sum_{j=\tau}^{t-1}\lambda^{t-j-1}(|w(j)| + |\tilde{y}^*(j+1)|), \qquad t > \tau. \tag{7.85}$$

As $\tau \ge t_0$ is arbitrary, this concludes the proof of Part i).

Step 4: Proving asymptotic tracking.

Observe that if y^* and w satisfy $\mathbf{Q}(z^{-1})Y^*(z) = 0$ and $\mathbf{Q}(z^{-1})W(z) = 0$, then from the definition of $\bar{W}(z)$ in (7.7) we have $\bar{w}(t) = 0$. So from (7.68) and the definition of $\bar{\psi}$, we see that

$$|\varepsilon(t)| \le \|\bar{\psi}(t)\| \le c\lambda^{t-t_0}\|\bar{\psi}(t_0)\|, \qquad t \ge t_0.$$

Using (7.80) to obtain a bound on $\|\bar{\psi}(t_0)\|$, we obtain the desired bound on the tracking error. ■

7.5 Robustness Results

To proceed, we consider a time-varying version of the plant (7.1). In order to apply the results of Chapter 2, we would like the plant model to incorporate the vector $\bar{\phi}(t)$ regardless of the value of $n \in \{1, 2, \ldots, \bar{n}\}$, so we will pad θ with zeros in the obvious spots and then write the time-varying version of plant (7.1) as

$$y(t+1) = \overline{\theta}(t)^\top \bar{\phi}(t) + w(t), \quad t \in \mathbb{Z}; \tag{7.86}$$

we define $\overline{\mathbf{\Theta}}_n \subset \mathbb{R}^{2(\bar{n}+g)}$ to represent the padded elements of $\mathbf{\Theta}_n$:

$$\overline{\mathbf{\Theta}}_n := \left\{ \begin{bmatrix} a_1 \\ \vdots \\ a_n \\ \mathbf{0}_{\bar{n}+g-n} \\ b_1 \\ \vdots \\ b_n \\ \mathbf{0}_{\bar{n}+g-n} \end{bmatrix} \in \mathbb{R}^{2(\bar{n}+g)} \; : \; \theta \in \mathbf{\Theta}_n \right\},$$

and define

$$\overline{\mathbf{\Theta}} := \bigcup_{n=1}^{\bar{n}} \overline{\mathbf{\Theta}}_n,$$

which is clearly compact.

We now consider the time-varying plant (7.86) with the term $d_\Delta(t) \in \mathbb{R}$ added to represent the unmodelled dynamics:

$$y(t+1) = \bar{\theta}(t)^\top \bar{\phi}(t) + w(t) + d_\Delta(t), \quad t \in \mathbb{Z}. \tag{7.87}$$

As discussed in Chapter 2, we assume that d_Δ satisfies

$$\mathfrak{w}(t+1) = \beta \mathfrak{w}(t) + \beta \left\| \bar{\phi}(t) \right\|, \quad \mathfrak{w}(t_0) = \mathfrak{w}_0 \tag{7.88a}$$

$$|d_\Delta(t)| \leq \mu \mathfrak{w}(t) + \mu \left\| \bar{\phi}(t) \right\|, \quad t \geq t_0. \tag{7.88b}$$

Theorem 7.2. *Suppose that the adaptive controller (7.18), (7.23)–(7.29) is applied to the time-varying plant (7.87) with d_Δ satisfying (7.88). Then for every $\delta \in (0, \infty]$, $N \geq 2\bar{n} + g$, $\beta \in (0,1)$ and $\bar{c}_0 \geq 0$, there exist $\bar{\epsilon} > 0$, $\mu > 0$, $\tilde{\lambda} \in (\beta, 1)$ and $\tilde{\gamma} > 0$ such that for every $t_0 \in \mathbb{Z}$, $\bar{\phi}(t_0) \in \mathbb{R}^{2(\bar{n}+g)}$, $\sigma_0 \in \mathcal{I}^*$, $\bar{\theta} \in \mathcal{S}\left(\overline{\mathbf{\Theta}}, \bar{c}_0, \bar{\epsilon}\right)$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ ($i \in \mathcal{I}^*$), and $w, y^* \in \boldsymbol{\ell}_\infty$, the following holds:*

$$\left\| \begin{bmatrix} \bar{\phi}(t) \\ \mathfrak{w}(t) \end{bmatrix} \right\| \leq \tilde{\gamma} \tilde{\lambda}^{t-t_0} \left\| \begin{bmatrix} \bar{\phi}(t_0) \\ \mathfrak{w}_0 \end{bmatrix} \right\| + \sum_{j=t_0}^{t-1} \tilde{\gamma} \tilde{\lambda}^{t-j-1} (|w(j)| + |\tilde{y}^*(j+1)|), \quad t \geq t_0.$$

Proof of Theorem 7.2. We observe here that the plant (7.1) and the controller (7.18), (7.23)–(7.29) fit into the paradigm of Chapter 2: we set

$$\begin{aligned}
\vartheta(t) &= \bar{\phi}(t), \\
f(\vartheta(\cdot)) &= \bar{\phi}(\cdot), \\
z(t) &= \varnothing, \\
\hat{\theta}(t) &= \begin{bmatrix} \hat{\theta}_1(t) \\ \hat{\theta}_2(t) \\ \vdots \\ \hat{\theta}_m(t) \end{bmatrix}, \\
r(t) &= \tilde{y}^*(t+1), \\
\Omega &= \mathcal{S}_1 \times \mathcal{S}_2 \times \cdots \times \mathcal{S}_m.
\end{aligned}$$

In Theorem 7.1 it is proven the controller (7.18), (7.23)–(7.29) provides a convolution bound for (7.1). Then, by Theorems 2.2, 2.2 and 2.3 we immediately see that the same is true in the presence of time-variation and/or unmodelled dynamics. ■

7.6 Simulation Examples

7.6.1 Sinusoidal Tracking

In this example, we will show the efficiency of the proposed approach, mainly in dealing with plant changes and noise. We have the upper bound on the order of the plant to be $\bar{n} = 2$; consider the following family of plants:
(i) first-order plants with an uncertainty set of

$$\mathbf{\Theta}_1 = \left\{ \begin{bmatrix} a_1 \\ b_1 \end{bmatrix} \in \mathbb{R}^2 : a_1 \in [1, \tfrac{3}{2}], b_1 \in [-2, -1] \cup [1, 2] \right\}$$

and (ii) second-order plants with an uncertainty set of

$$\mathbf{\Theta}_2 = \left\{ \begin{bmatrix} a_1 \\ a_2 \\ b_1 \\ b_2 \end{bmatrix} \in \mathbb{R}^4 \, : \, a_1 = \tfrac{3}{2}, a_2 \in \left\{ \tfrac{-3}{2}, \tfrac{3}{2} \right\}, b_1 \in [-1, 0], b_2 \in [-5, -3] \right\}.$$

It is obvious that each of the sets above is compact as required; also, the coprimeness requirement is satisfied. You can see that all potential models are unstable. Also, the 2^{nd}-order models are all nonminimum phase.

The goal is to track reference signals of frequency $\frac{\pi}{25}$: so we set

$$\mathbf{Q}(z^{-1}) = 1 - 2\cos(\tfrac{\pi}{25})z^{-1} + z^{-2},$$

i.e.

$$q_1 = 2\cos(\tfrac{\pi}{25}), \qquad q_2 = -1, \qquad g = 2.$$

Observe that 2^{nd}-order plant models have a real zero that can lie in $[3, \infty)$, i.e. the associated $\mathbf{B}(z^{-1})$ and $\mathbf{Q}(z^{-1})$ are coprime as required. Next, with $n \in \{1, 2\}$ we use the definition in (7.11) to construct the uncertainty sets of the associated auxiliary plant:

$$\begin{aligned}
\tilde{\mathbf{\Theta}}_1 = \Bigg\{ & \begin{bmatrix} \bar{a}_1 & \bar{a}_2 & \bar{a}_3 & b_1 \end{bmatrix}^\top \in \mathbb{R}^4 : \bar{a}_1 \in [1 + q_1, \tfrac{3}{2} + q_1], \bar{a}_2 \in [q_2 - \tfrac{3q_1}{2}, q_2 - q_1], \bar{a}_3 \in [-q_2, -\tfrac{3q_2}{2}], \\
& b_1 \in [-2, -1] \cup [1, 2] \Bigg\}, \\
\tilde{\mathbf{\Theta}}_2 = \Bigg\{ & \begin{bmatrix} \bar{a}_1 & \bar{a}_2 & \bar{a}_3 & \bar{a}_4 & b_1 & b_2 \end{bmatrix}^\top \in \mathbb{R}^6 : \bar{a}_1 = \tfrac{3}{2} + q_1, \\
& (\bar{a}_2, \bar{a}_3, \bar{a}_4) \in \left\{ \left(q_2 + \tfrac{3}{2}(1 - q_1), -q_2 - \tfrac{3q_1}{2}, -\tfrac{3q_2}{2}\right), \left(q_2 + \tfrac{3}{2}(q_1 - 1), -q_2 + \tfrac{3q_1}{2}, \tfrac{3q_2}{2}\right) \right\}, \\
& b_1 \in [-1, 0], b_2 \in [-5, -3] \Bigg\};
\end{aligned}$$

clearly these sets are compact. We see that $\tilde{\mathbf{\Theta}}_1$ and $\tilde{\mathbf{\Theta}}_2$ are not convex; the convex hull of each could violate the coprimeness requirement; for example notice that the convex hull of $\tilde{\mathbf{\Theta}}_1$ includes the case of $b_1 = 0$, which corresponds to a non-stabilizable system, violating the coprimeness assumption. For each of $\tilde{\mathbf{\Theta}}_1$ and $\tilde{\mathbf{\Theta}}_2$, we will need a set of compact and convex sets so that their union contain $\tilde{\mathbf{\Theta}}_1$ and $\tilde{\mathbf{\Theta}}_2$ respectively and satisfy the coprimeness requirement. There is a natural choice: define

$$\begin{aligned}
\mathcal{S}_1 = \Bigg\{ & \begin{bmatrix} \bar{a}_1 & \bar{a}_2 & \bar{a}_3 & \bar{a}_4 & b_1 & b_2 \end{bmatrix}^\top \in \mathbb{R}^6 : \bar{a}_1 = \tfrac{3}{2} + q_1, \bar{a}_2 = q_2 + \tfrac{3}{2}(1 - q_1), \\
& \bar{a}_3 = -q_2 - \tfrac{3q_1}{2}, \bar{a}_4 = -\tfrac{3q_2}{2}, b_1 \in [-1, 0], b_2 \in [-5, -3] \Bigg\},
\end{aligned}$$

$$\mathcal{S}_2 = \Big\{ \begin{bmatrix} \bar{a}_1 & \bar{a}_2 & \bar{a}_3 & \bar{a}_4 & b_1 & b_2 \end{bmatrix}^\top \in \mathbb{R}^6 : \bar{a}_1 = \tfrac{3}{2} + q_1, \bar{a}_2 = q_2 + \tfrac{3}{2}(q_1 - 1),$$

$$\bar{a}_3 = -q_2 + \tfrac{3q_1}{2}, \bar{a}_4 = \tfrac{3q_2}{2}, b_1 \in [-1, 0], b_2 \in [-5, -3] \Big\},$$

$$\mathcal{S}_3 = \Big\{ \begin{bmatrix} \bar{a}_1 & \bar{a}_2 & \bar{a}_3 & b_1 \end{bmatrix}^\top \in \mathbb{R}^4 : \bar{a}_1 \in [1 + q_1, \tfrac{3}{2} + q_1], \bar{a}_2 \in [q_2 - \tfrac{3q_1}{2}, q_2 - q_1], \bar{a}_3 \in [-q_2, -\tfrac{3q_2}{2}],$$

$$b_1 \in [-2, -1] \Big\},$$

$$\mathcal{S}_4 = \Big\{ \begin{bmatrix} \bar{a}_1 & \bar{a}_2 & \bar{a}_3 & b_1 \end{bmatrix}^\top \in \mathbb{R}^4 : \bar{a}_1 \in [1 + q_1, \tfrac{3}{2} + q_1], \bar{a}_2 \in [q_2 - \tfrac{3q_1}{2}, q_2 - q_1], \bar{a}_3 \in [-q_2, -\tfrac{3q_2}{2}],$$

$$b_1 \in [1, 2] \Big\};$$

clearly $\tilde{\boldsymbol{\Theta}}_1 \subset \mathcal{S}_3 \cup \mathcal{S}_4$ and $\tilde{\boldsymbol{\Theta}}_2 \subset \mathcal{S}_1 \cup \mathcal{S}_2$. The auxiliary plant associated with the 2^{nd}-order plant models has potential poles of either complex ones, or real ones of values 2.186 or -0.686, while it has a zero that lies in $[3, \infty)$; this means that the coprimeness requirement is satisfied as well. So we are going to estimate parameters using 4 parallel estimators; however, we see that for $\mathcal{S}_1$ and $\mathcal{S}_2$, each has only one value for parameters $\bar{a}_1, \bar{a}_2, \bar{a}_3$ and $\bar{a}_4$, which means that the estimation of those parameters is trivial.

For this simulation, we set the plant to

$$y(t+1) = \begin{cases} \tfrac{3}{2}\Big[y(t) - y(t-1)\Big] - \tfrac{3}{4}u(t) - 4u(t-1), & t \leq 500 \\ -\tfrac{3}{2}\Big[y(t) + u(t)\Big], & t > 500. \end{cases}$$

We set the reference signal to

$$y^*(t) = 2\sin(\tfrac{\pi}{25}t)$$

and the noise to

$$w(t) = \begin{cases} 0.05\cos(45t), & 250 \leq t < 750 \\ 0 & \text{otherwise.} \end{cases}$$

We will apply the proposed controller (7.18) and (7.23)–(7.29); we choose $N = 6$ and $\delta = \infty$. We set the plant initial conditions to $y(0) = y(-1) = y(-2) = y(-3) = 1.75$ and $u(0) = u(-1) = u(-2) = u(-3) = 0$; we also set

$$\hat{\theta}_1(0) = \begin{bmatrix} \tfrac{3}{2} + q_1 & q_2 + \tfrac{3}{2}(1 - q_1) & -q_2 - \tfrac{3q_1}{2} & -\tfrac{3q_2}{2} & -\tfrac{1}{2} & -4.5 \end{bmatrix}^\top,$$

$$\hat{\theta}_2(0) = \begin{bmatrix} \frac{3}{2}+q_1 & q_2+\frac{3}{2}(q_1-1) & -q_2+\frac{3q_1}{2} & \frac{3q_2}{2} & -\frac{1}{2} & -4.5 \end{bmatrix}^\top,$$
$$\hat{\theta}_3(0) = \begin{bmatrix} q_1+1 & 1.25 & \frac{1}{2} & -2 \end{bmatrix}^\top,$$
$$\hat{\theta}_4(0) = \begin{bmatrix} q_1+1 & 1.25 & \frac{1}{2} & 2 \end{bmatrix}^\top,$$

and $\sigma_0 = 4$. The results are in Fig. 7.3. We see that the controller provides good tracking performance; while performance degrades temporarily when noise is added and when the plant change happens, tracking recovers.

7.6.2 Set-Point Control

In this section, another simulation example is provided to show the application to set-point control. Here, we illustrate the disturbance rejection property and tolerance to slow time-variation of the approach. Consider the 2^{nd}-order plant:

$$y(t+1) = a_1(t)y(t) + a_2(t)y(t-1) + b_1(t)u(t) + b_2(t)u(t-1) + w(t)$$

with parameters belonging to the uncertainty sets $\mathbf{\Theta}_1 = \varnothing$ and

$$\mathbf{\Theta}_2 = \left\{ \begin{bmatrix} a_1 \\ a_2 \\ b_1 \\ b_2 \end{bmatrix} \in \mathbb{R}^4 : a_1 \in [-2,0], a_2 \in [-3,-1], b_1 \in [-1,0], b_2 \in [-5,-3] \cup [3,5] \right\},$$

i.e. the order is known in this case. Hence, every admissible model is unstable and non-minimum phase, which makes this plant challenging to control; it has two complex unstable poles together with a zero that can lie in $[3,\infty)$. It is also obvious to see that $\mathbf{\Theta}_2$ is not a convex set; notice that the convex hull of it includes the case of having $b_1 = b_2 = 0$, which corresponds to a non-stabilizable system, violating the coprimeness assumption. So, we apply the proposed approach in this chapter.

With $\mathbf{Q}(z^{-1}) = 1 - z^{-1}$, i.e.

$$q_1 = 1, \qquad g = 1,$$

we define the set $\tilde{\mathbf{\Theta}}_2$ by (7.11); so we will be estimating the parameters of the auxiliary plant: $\theta^*(t) = \begin{bmatrix} \bar{a}_1(t) & \bar{a}_2(t) & \bar{a}_3(t) & b_1(t) & b_2(t) \end{bmatrix}^\top \in \tilde{\mathbf{\Theta}}_2$. We know that the set $\tilde{\mathbf{\Theta}}_2$ is also compact and satisfies the coprimeness requirement; we will need to find a set of compact and convex sets that their union contains $\tilde{\mathbf{\Theta}}_2$ and that will also satisfy the coprimeness

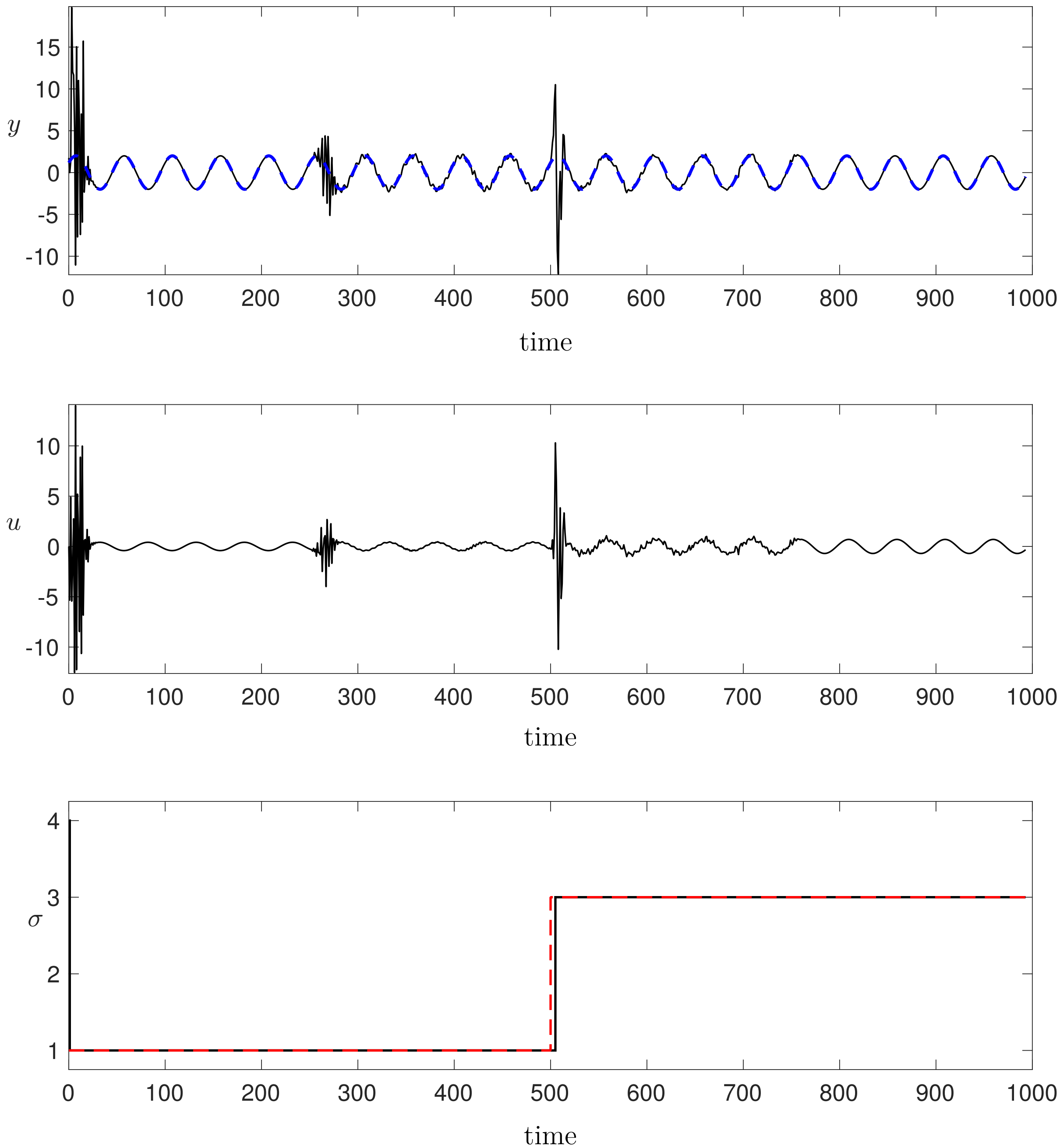


Figure 7.3: The upper plot shows both the reference (dashed) and the output (solid); the next plot shows the plant control input; the bottom plot shows the switching signal (solid) and the correct index (dashed).

requirement. We define

$$\mathcal{S}_1 = \left\{ \begin{bmatrix} \bar{a}_1 \\ \bar{a}_2 \\ \bar{a}_3 \\ b_1 \\ b_2 \end{bmatrix} \in \mathbb{R}^5 : \bar{a}_1 \in [-1,1], \bar{a}_2 \in [-3,1], \bar{a}_3 \in [1,3], b_1 \in [-1,0], b_2 \in [-5,-3] \right\},$$

$$\mathcal{S}_2 = \left\{ \begin{bmatrix} \bar{a}_1 \\ \bar{a}_2 \\ \bar{a}_3 \\ b_1 \\ b_2 \end{bmatrix} \in \mathbb{R}^5 : \bar{a}_1 \in [-1,1], \bar{a}_2 \in [-3,1], \bar{a}_3 \in [1,3], b_1 \in [-1,0], b_2 \in [3,5] \right\}.$$

Each of the sets $\mathcal{S}_1$ and $\mathcal{S}_2$ is a hyper-rectangle, which is easy to project onto; we easily see that $\tilde{\mathbf{\Theta}}_2 \subset \mathcal{S}_1 \cup \mathcal{S}_2$. It can be verified that each of $\mathcal{S}_1$ and $\mathcal{S}_2$ contain models that are coprime, as desired.

For this simulation we set

$$\begin{aligned} a_1(t) &= -\tfrac{1}{2}, \\ a_2(t) &= -2 + \tfrac{1}{2}\cos\left(\tfrac{1}{100}t\right), \\ b_1(t) &= -\tfrac{1}{2} - \tfrac{1}{2}\sin\left(\tfrac{1}{175}t\right), \\ b_2(t) &= -4. \end{aligned}$$

We will apply the proposed controller (7.18) and (7.23)–(7.29); we choose $N = 5$ and $\delta = \bar{\mathbf{s}} \approx 6.7$. We set the reference y^* to be a square wave of magnitude 2 and period 350, and initial condition $y(0) = y(-1) = y(-2) = 2$ and $u(0) = u(-1) = u(-2) = 0$; we also set $\hat{\theta}_1(0) = \begin{bmatrix} 0 & -1 & 2 & -\frac{1}{2} & -5 \end{bmatrix}^\top$, $\hat{\theta}_2(0) = \begin{bmatrix} 0 & -1 & 2 & -\frac{1}{2} & 5 \end{bmatrix}^\top$ and $\sigma_0 = 2$. We set the disturbance to be of a constant magnitude: $|w(t)| = \frac{1}{2}$, but with its sign changing every 250 steps. Figure 7.4 displays the results. We see that the controller does a good job of tracking even when parameters are time-varying; the closed-loop system experiences some transient behavior when the set-point or disturbance change signs, but the tracking recovers quickly.

Remark 7.11. *We can also compare the performance here with that which arises when we use the classical estimator* (3.4) *as part of the adaptive controller; we end up with the same sort of result as in Example 2 of the simulation section of Chapter 4, namely a degradation in performance.*

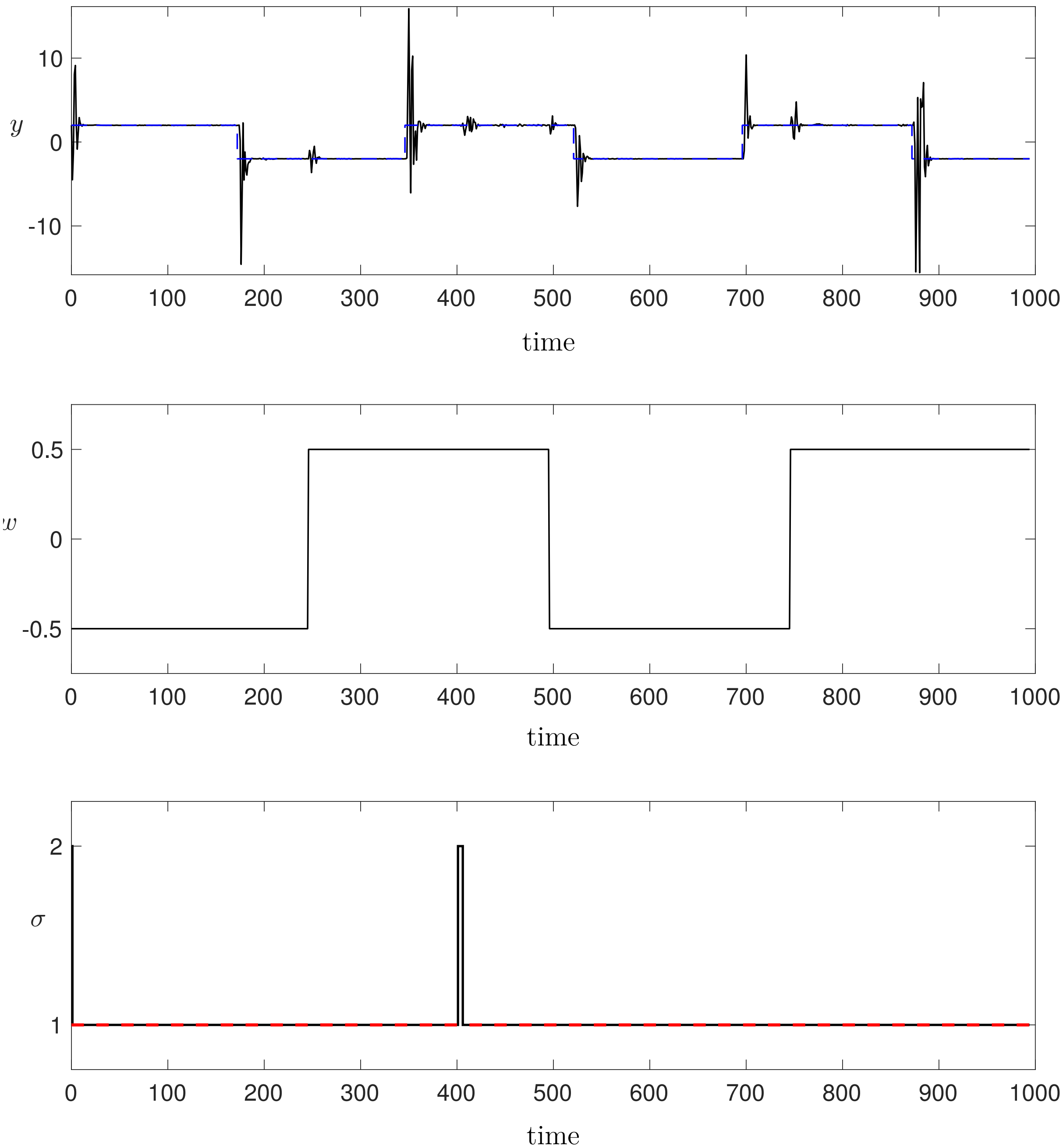


Figure 7.4: The top plot shows both the reference (dashed) and the output (solid); the middle plot shows the disturbance w; the bottom plot shows the switching signal (solid) and the correct index (dashed).

7.7 Conclusion

In this chapter, we have considered the problem of tracking for a discrete-time plant with unknown order; we assume knowledge of an upper bound on the order, and that the uncertainty set of parameters for each admissible order lies in a compact set, subject to a coprimeness requirement. Rather than directly estimating the plant parameters, we instead estimate the parameters of a suitably defined auxiliary plant model. We use compactness to prove that for each admissible order, the uncertainty set is contained in a finite union of convex sets; we use a projection-algorithm based estimator for each convex set. At each point in time, we employ a switching algorithm to determine which model and parameter estimates are used in the control law. We prove that this adaptive controller guarantees desirable linear-like closed-loop behavior: exponential stability, a bounded noise gain, and convolution bounds on the input-output behavior, as well as asymptotic tracking for certain classes of reference and noise signals; we do not assume that the switching stops.

Chapter 8

Conclusion

8.1 Summary of Results

We have developed an approach to multi-model adaptive control that guarantees linear-like closed-loop behavior: exponential stability, a bounded noise gain and a convolution bound on the exogenous inputs. In contrast, usually in the adaptive control literature, only asymptotic stability and a bounded-noise bounded-state property is proven, although there are exceptions; furthermore, the only results which yield a convolution bound are Miller [43], Miller and Shahab [46], Miller and Shahab [48]. In earlier work on this approach, namely in the aforementioned papers, the requirement of convexity on the set of uncertainty plays a crucial role in obtaining these desirable closed-loop properties. Here we have shown that we can prove the same linear-like properties without the convexity assumption: the main idea is to use multiple estimators together with a switching algorithm. The proposed discrete-time adaptive control approach includes: 1) covering the compact set of admissible parameters by a finite number of convex sets, 2) designing a parameter estimator based on the original Projection Algorithm for each convex set, and 3) using a switching algorithm to switch between the corresponding controllers.

First, we have proven a general result that exponential stability and a linear-like convolution bound on the closed-loop behavior can be leveraged to show tolerance to a degree of time-variations and unmodelled dynamics, i.e. such a bound guarantees robustness, which is proven in a modular fashion. After that, we have proven, in various contexts and with a focus on non-convex but compact sets of uncertainty, that our approach provides the desirable aforementioned linear-like properties. First we consider the simplest case of adaptive control of first-order linear plants. We then extend the approach to a special class of nonlinear plants (which have stable zero dynamics); we consider both cases of a known control/input gain sign and an unknown one. Afterwards, we turn to adaptive control of possibly non-minimum phase LTI plants; we first consider the stabilization problem given two convex sets of uncertainty. Finally, we turn to the more difficult problem of tracking the sum of a finite number of sinusoids of known frequencies subject to an unknown plant order and a general compact set of uncertainty.

8.2 Some Limitations

Here we discuss some limitations of our approach. You can see that in both the context of the one-step-ahead control laws of Chapter 4 and Chapter 5, and the context of the pole-placement based control laws of Chapter 6 and Chapter 7, we rely heavily on the deadbeat nature of the approach to provide the desirable closed-loop properties. One drawback is that this may incur larger transient control actions, which can be undesirable for some practical applications. A way around this is to investigate how to extend the approach to less demanding control law designs.

We have shown asymptotic tracking results for a general reference signal in the nonlinear systems setting when using a single estimator (Chapter 5), and for a reference signal which is a sum of sinusoids of known frequencies in the LTI plants setting when using multiple estimators (Chapter 7). However, we were not able to prove asymptotic tracking of general reference signals when using multiple estimators. One possible reason is that our proposed switching mechanism, which includes memoryless performance signals, is not proven to eventually stop even in the absence of noise; however, in all of our simulations, switching does stop and tracking is achieved. While the memoryless performance signals played a role in obtaining desirable closed-loop properties, one expects there exists a careful choice of a performance signal with memory that achieves tracking as well as the desirable linear-like closed behavior.

Observe that in the nonlinear setting, the approach considers only a special family of nonlinear systems. As mentioned earlier, this work can be considered a first step towards obtaining desirable linear-like closed-loop properties in the context of adaptive control of a more general class of discrete-time nonlinear systems.

8.3 Future Directions

Finally, we provide some possible avenues of research in relation to the results of this thesis.

- An obvious area is to extend the results to the continuous-time setting. The difficulty here is that the estimator often look much different, and in the cases which look similar, there are technical issues about existence and uniqueness of the solutions to the associated differential equation.

- As has been alluded to in the limitations section, another aspect worth studying is related to controller design. We would like to further investigate the performance of the transient behavior which would be helpful for potential design issues.

- Also related to the previous point, one would like to extend the One-Step-Ahead Adaptive Control approach to the more general Model Reference Adaptive Control (MRAC) problem; because we are seeking stronger closed-loop properties than what is normally proven in the literature, more detailed analysis is expected in dealing with the MRAC setup.

- Discussed in the limitations section, we would like to investigate more sophisticated switching algorithms while aiming for the same desirable closed-loop properties; approaches to switching which use performance signals with memory merits more study.

- We would like also to relax the coprimeness requirement in the pole-placement approach of Chapter 6 and Chapter 7 to one requiring that all common zeros be in the open unit disk. Hence, we would like to relax Assumption 7.1 to one having a single compact set in the highest dimension.

- Of course, an area of interest is potential practical applications; one example is dealing with sensor and/or actuator failures (with the approach suitably modified into the multi-input multi-output paradigm).

References

[1] B. D. O. Anderson, T. S. Brinsmead, F. De Bruyne, J. P. Hespanha, D. Liberzon, and A. S. Morse, "Multiple model adaptive control. part 1: Finite controller coverings," *International Journal of Robust and Nonlinear Control*, vol. 10, no. 11-12, pp. 909–929, Sep. 2000.

[2] B. D. O. Anderson, T. Brinsmead, D. Liberzon, and A. S. Morse, "Multiple model adaptive control with safe switching," *International Journal of Adaptive Control and Signal Processing*, vol. 15, no. 5, pp. 445–470, 2001.

[3] B. D. O. Anderson and C. R. Johnson, "Exponential convergence of adaptive identification and control algorithms," *Automatica*, vol. 18, no. 1, pp. 1 –13, 1982.

[4] S. Baldi, P. A. Ioannou, and E. Mosca, "Multiple Model Adaptive Mixing Control: The Discrete-Time Case," *IEEE Transactions on Automatic Control*, vol. 57, no. 4, pp. 1040–1045, Apr. 2012.

[5] S. Baldi, G. Battistelli, E. Mosca, and P. Tesi, "Multi-model unfalsified adaptive switching supervisory control," *Automatica*, vol. 46, no. 2, pp. 249–259, 2010.

[6] S. Baldi and P. A. Ioannou, "Stability Margins in Adaptive Mixing Control Via a Lyapunov-Based Switching Criterion," *IEEE Transactions on Automatic Control*, vol. 61, no. 5, pp. 1194–1207, May 2016.

[7] S. Baldi, P. A. Ioannou, and E. B. Kosmatopoulos, "Adaptive mixing control with multiple estimators," *International Journal of Adaptive Control and Signal Processing*, vol. 26, no. 8, pp. 800–820, Aug. 2012.

[8] G. Battistelli, E. Mosca, M. G. Safonov, and P. Tesi, "Stability of unfalsified adaptive switching control in noisy environments," *IEEE Transactions on Automatic Control*, vol. 55, no. 10, pp. 2424–2429, Oct. 2010.

[9] G. Battistelli, J. P. Hespanha, E. Mosca, and P. Tesi, "Model-Free Adaptive Switching Control of Time-Varying Plants," *IEEE Transactions on Automatic Control*, vol. 58, no. 5, pp. 1208–1220, May 2013.

[10] G. Battistelli, E. Mosca, and P. Tesi, "Adaptive memory in multi-model switching control of uncertain plants," *Automatica*, vol. 50, no. 3, pp. 874–882, 2014.

[11] D. Borrelli, A. S. Morse, and E. Mosca, "Discrete-time supervisory control of families of two-degrees-of-freedom linear set-point controllers," *IEEE Transactions on Automatic Control*, vol. 44, no. 1, pp. 178–181, Jan. 1999.

[12] F.-C. Chen and H. K. Khalil, "Adaptive control of a class of nonlinear discrete-time systems using neural networks," *IEEE Transactions on Automatic Control*, vol. 40, no. 5, pp. 791–801, May 1995.

[13] S. Dai, C. Yang, S. S. Ge, and T. H. Lee, "Robust adaptive output feedback control of a class of discrete-time nonlinear systems with nonlinear uncertainties and unknown control directions," *International Journal of Robust and Nonlinear Control*, vol. 23, no. 13, pp. 1472–1495, Sep. 2013.

[14] C. Desoer, "Slowly varying discrete system $x_{i+1} = A_i x_i$," *Electronics Letters*, vol. 6, no. 11, pp. 339–340, May 1970.

[15] S. Fekri, M. Athans, and A. Pascoal, "Issues, progress and new results in robust adaptive control," *International Journal of Adaptive Control and Signal Processing*, vol. 20, no. 10, pp. 519–579, Dec. 2006.

[16] A. Feuer and A. S. Morse, "Adaptive control of single-input, single-output linear systems," *IEEE Transactions on Automatic Control*, vol. 23, no. 4, pp. 557–569, Aug. 1978.

[17] M. Fu and B. Barmish, "Adaptive stabilization of linear systems via switching control," *IEEE Transactions on Automatic Control*, vol. 31, no. 12, pp. 1097–1103, Dec. 1986.

[18] S. S. Ge, C. Yang, S.-L. Dai, Z. Jiao, and T. H. Lee, "Robust adaptive control of a class of nonlinear strict-feedback discrete-time systems with exact output tracking," *Automatica*, vol. 45, no. 11, pp. 2537 –2545, 2009.

[19] S. S. Ge, C. Yang, and T. H. Lee, "Adaptive robust control of a class of nonlinear strict-feedback discrete-time systems with unknown control directions," *Systems & Control Letters*, vol. 57, no. 11, pp. 888 –895, 2008.

[20] G. C. Goodwin, P. Ramadge, and P. Caines, "Discrete-time multivariable adaptive control," *IEEE Transactions on Automatic Control*, vol. 25, no. 3, pp. 449–456, Jun. 1980.

[21] G. C. Goodwin and K. S. Sin, *Adaptive Filtering Prediction and Control*. New York, NY, USA: Dover Publications, Inc., 1984.

[22] L. Guo, "On critical stability of discrete-time adaptive nonlinear control," *IEEE Transactions on Automatic Control*, vol. 42, no. 11, pp. 1488–1499, Nov. 1997.

[23] Z. Han and K. S. Narendra, "New Concepts in Adaptive Control Using Multiple Models," *IEEE Transactions on Automatic Control*, vol. 57, no. 1, pp. 78–89, Jan. 2012.

[24] J. P. Hespanha and A. S. Morse, "Stability of switched systems with average dwell-time," in *Proceedings of the 38th IEEE Conference on Decision and Control*, vol. 3, Dec. 1999, 2655–2660 vol.3.

[25] J. P. Hespanha, D. Liberzon, A. S. Morse, B. D. O. Anderson, T. S. Brinsmead, and F. De Bruyne, "Multiple model adaptive control. Part 2: switching," *International Journal of Robust and Nonlinear Control*, vol. 11, no. 5, pp. 479–496, Apr. 2001.

[26] J. P. Hespanha, D. Liberzon, and A. S. Morse, "Hysteresis-based switching algorithms for supervisory control of uncertain systems," *Automatica*, vol. 39, no. 2, pp. 263–272, Feb. 2003.

[27] J. P. Hespanha, D. Liberzon, and A. S. Morse, "Overcoming the limitations of adaptive control by means of logic-based switching," *Systems & Control Letters*, vol. 49, no. 1, pp. 49–65, May 2003.

[28] J. P. Hespanha and A. S. Morse, "Scale-independent hysteresis switching," in *Hybrid Systems: Computation and Control*, F. W. Vaandrager and J. H. van Schuppen, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 1999, pp. 117–122.

[29] P. A. Ioannou and K. S. Tsakalis, "A robust direct adaptive controller," *IEEE Transactions on Automatic Control*, vol. 31, no. 11, pp. 1033–1043, Nov. 1986.

[30] P. A. Ioannou and B. Fidan, *Adaptive control tutorial*. Society for Industrial and Applied Mathematics, 2006.

[31] P. A. Ioannou and J. Sun, *Robust Adaptive Control*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1995.

[32] I. Kanellakopoulos, "A discrete-time adaptive nonlinear system," *IEEE Transactions on Automatic Control*, vol. 39, no. 11, pp. 2362–2365, Nov. 1994.

[33] S. Kersting and M. Buss, "How to systematically distribute candidate models and robust controllers in multiple-model adaptive control: A coverage control approach," *IEEE Transactions on Automatic Control*, vol. 63, no. 4, pp. 1075–1089, Apr. 2018.

[34] G. Kreisselmeier and B. D. O. Anderson, "Robust model reference adaptive control," *IEEE Transactions on Automatic Control*, vol. 31, no. 2, pp. 127–133, Feb. 1986.

[35] G. Kreisselmeier, "Adaptive control of a class of slowly time-varying plants," *Systems & Control Letters*, vol. 8, no. 2, pp. 97–103, Dec. 1986.

[36] M. Kuipers and P. A. Ioannou, "Multiple Model Adaptive Control With Mixing," *IEEE Transactions on Automatic Control*, vol. 55, no. 8, pp. 1822–1836, Aug. 2010.

[37] I. Landau, R. Lozano, M. M'Saad, and A. Karimi, *Adaptive Control: Algorithms, Analysis and Applications*, ser. Communications and Control Engineering. London: Springer, 2011.

[38] Y. Li and H.-F. Chen, "Robust adaptive pole placement for linear time-varying systems," *IEEE Transactions on Automatic Control*, vol. 41, no. 5, pp. 714–719, May 1996.

[39] B. Mårtensson, "The order of any stabilizing regulator is sufficient a priori information for adaptive stabilization," *Systems & Control Letters*, vol. 6, no. 2, pp. 87–91, Jul. 1985.

[40] R. H. Middleton and G. C. Goodwin, "Adaptive control of time-varying linear systems," *IEEE Transactions on Automatic Control*, vol. 33, no. 2, pp. 150–155, 1988.

[41] R. H. Middleton, G. C. Goodwin, D. J. Hill, and D. Q. Mayne, "Design issues in adaptive control," *IEEE Transactions on Automatic Control*, vol. 33, no. 1, pp. 50–58, Jan. 1988.

[42] D. E. Miller, "Adaptive stabilization using a nonlinear time-varying controller," *IEEE Transactions on Automatic Control*, vol. 39, no. 7, pp. 1347–1359, Jul. 1994.

[43] D. E. Miller, "A parameter adaptive controller which provides exponential stability: The first order case," *Systems & Control Letters*, vol. 103, pp. 23–31, May 2017.

[44] D. E. Miller, "Classical discrete-time adaptive control revisited: Exponential stabilization," in *2017 IEEE Conference on Control Technology and Applications (CCTA)*, IEEE, Aug. 2017, pp. 1975–1980.

[45] D. E. Miller and E. J. Davison, "An adaptive controller which provides Lyapunov stability," *IEEE Transactions on Automatic Control*, vol. 34, no. 6, pp. 599–609, Jun. 1989.

[46] 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*, vol. 30, no. 4, p. 19, Nov. 2018.

[47] D. E. Miller and M. T. Shahab, "Classical d-step-ahead adaptive control revisited: Linear-like convolution bounds and exponential stability," in *2019 American Control Conference*, Jul. 2019, pp. 417–422.

[48] D. E. Miller and M. T. Shahab, "Adaptive tracking with exponential stability and convolution bounds using vigilant estimation," *Mathematics of Control, Signals, and Systems*, Apr. 2020.

[49] A. S. Morse, "Global stability of parameter-adaptive control systems," *IEEE Transactions on Automatic Control*, vol. 25, no. 3, pp. 433–439, Jun. 1980.

[50] A. S. Morse, "Supervisory control of families of linear set-point controllers–Part 1: Exact matching," *IEEE Transactions on Automatic Control*, vol. 41, no. 10, pp. 1413–1431, 1996.

[51] A. S. Morse, "Supervisory Control of Families of Linear Set-Point Controllers–Part 2: Robustness," *IEEE Transactions on Automatic Control*, vol. 42, no. 11, pp. 1500–1515, 1997.

[52] A. S. Morse, D. Q. Mayne, and G. C. Goodwin, "Applications of hysteresis switching in parameter adaptive control," *IEEE Transactions on Automatic Control*, vol. 37, no. 9, pp. 1343–1354, 1992.

[53] E. Mosca, *Optimal, predictive, and adaptive control*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1995.

[54] D. Mudgett and A. S. Morse, "Adaptive stabilization of linear systems with unknown high-frequency gains," *IEEE Transactions on Automatic Control*, vol. 30, no. 6, pp. 549–554, Jun. 1985.

[55] S. Naik, P. Kumar, and B. Ydstie, "Robust continuous-time adaptive control by parameter projection," *IEEE Transactions on Automatic Control*, vol. 37, no. 2, pp. 182–197, 1992.

[56] K. S. Narendra and J. Balakrishnan, "Improving transient response of adaptive control systems using multiple models and switching," *IEEE Transactions on Automatic Control*, vol. 39, no. 9, pp. 1861–1866, 1994.

[57] K. S. Narendra and J. Balakrishnan, "Adaptive control using multiple models," *IEEE Transactions on Automatic Control*, vol. 42, no. 2, pp. 171–187, 1997.

[58] K. S. Narendra and Z. Han, "A new approach to adaptive control using multiple models," *International Journal of Adaptive Control and Signal Processing*, vol. 26, no. 8, pp. 778–799, Aug. 2012.

[59] K. S. Narendra and Y.-H. Lin, "Stable discrete adaptive control," *IEEE Transactions on Automatic Control*, vol. 25, no. 3, pp. 456–461, Jun. 1980.

[60] K. S. Narendra, Y.-H. Lin, and L. Valavani, “Stable adaptive controller design, part II: Proof of stability,” *IEEE Transactions on Automatic Control*, vol. 25, no. 3, pp. 440–448, Jun. 1980.

[61] K. S. Narendra and C. Xiang, “Adaptive control of discrete-time systems using multiple models,” *IEEE Transactions on Automatic Control*, vol. 45, no. 9, pp. 1669–1686, 2000.

[62] R. Nussbaum, “Some remarks on a conjecture in parameter adaptive control,” *Systems & Control Letters*, vol. 3, no. 5, pp. 243–246, Nov. 1983.

[63] F. M. Pait and F. Kassab, “On a class of switched, robustly stable, adaptive systems,” *International Journal of Adaptive Control and Signal Processing*, vol. 15, no. 3, pp. 213–238, May 2001.

[64] J.-B. Pomet and L. Praly, “Adaptive nonlinear regulation: Estimation from the lyapunov equation,” *IEEE Transactions on Automatic Control*, vol. 37, no. 6, pp. 729–740, Jun. 1992.

[65] 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*, vol. 30, no. 9, pp. 881–889, Sep. 1985.

[66] H. L. Royden, *Real analysis*. New York, NY, USA: Macmillan, 1988.

[67] M. G. Safonov and T.-C. Tsao, “The unfalsified control concept and learning,” *IEEE Transactions on Automatic Control*, vol. 42, no. 6, pp. 843–847, Jun. 1997.

[68] 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*, Dec. 2018, pp. 2223–2228.

[69] M. T. Shahab and D. E. Miller, “Adaptive Set-Point Regulation using Multiple Estimators,” in *2019 IEEE Conference on Decision and Control*, Dec. 2019, pp. 84–89.

[70] M. T. Shahab and D. E. Miller, “Exponential stability for adaptive control of a class of first-order nonlinear systems,” *IFAC-PapersOnLine*, vol. 52, no. 29, pp. 168 –173, 2019, 13th IFAC Workshop on Adaptive and Learning Control Systems ALCOS 2019.

[71] M. T. Shahab and D. E. Miller, “Asymptotic Tracking and Linear-like Behavior Using Multi-Model Adaptive Control,” *IEEE Transactions on Automatic Control*, 2020, (Submitted, under revision).

[72] 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)*, (Submitted), 2020.

[73] Y. Song and J. W. Grizzle, "Adaptive output-feedback control of a class of discrete-time nonlinear systems," in *1993 American Control Conference*, Jun. 1993, pp. 1359–1364.

[74] M. Stefanovic and M. G. Safonov, "Safe Adaptive Switching Control: Stability and Convergence," *IEEE Transactions on Automatic Control*, vol. 53, no. 9, pp. 2012–2021, Oct. 2008.

[75] 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*, vol. 34, no. 10, pp. 1038–1046, 1989.

[76] L. Vu and D. Liberzon, "Supervisory Control of Uncertain Linear Time-Varying Systems," *IEEE Transactions on Automatic Control*, vol. 56, no. 1, pp. 27–42, Jan. 2011.

[77] C. Wen, "A robust adaptive controller with minimal modifications for discrete time-varying systems," *IEEE Transactions on Automatic Control*, vol. 39, no. 5, pp. 987–991, May 1994.

[78] C. Wen and D. J. Hill, "Global boundedness of discrete-time adaptive control just using estimator projection," *Automatica*, vol. 28, no. 6, pp. 1143–1157, Nov. 1992.

[79] J. C. Willems and G. I. Byrnes, "Global adaptive stabilization in the absence of information on the sign of the high frequency gain," in *Analysis and Optimization of Systems*, Berlin/Heidelberg: Springer-Verlag, 1984, pp. 49–57.

[80] B. Ydstie, "Transient performance and robustness of direct adaptive control," *IEEE Transactions on Automatic Control*, vol. 37, no. 8, pp. 1091–1105, 1992.

[81] B. Ydstie, "Stability of discrete model reference adaptive control — revisited," *Systems & Control Letters*, vol. 13, no. 5, pp. 429 –438, 1989.

[82] G. Zames, "On the input-output stability of time-varying nonlinear feedback systems Part one: Conditions derived using concepts of loop gain, conicity, and positivity," *IEEE Transactions on Automatic Control*, vol. 11, no. 2, pp. 228–238, Apr. 1966.

[83] Y. Zhang, W. H. Chen, and Y. C. Soh, "Improved robust backstepping adaptive control for nonlinear discrete-time systems without overparameterization," *Automatica*, vol. 44, no. 3, pp. 864 –867, 2008.

[84] Y. Zhang, C. Wen, and Y. C. Soh, “Robust adaptive control of nonlinear discrete-time systems by backstepping without overparameterization,” *Automatica*, vol. 37, no. 4, pp. 551 –558, 2001.

[85] J. Zhao and I. Kanellakopoulos, “Active identification for discrete-time nonlinear control. i. output-feedback systems,” *IEEE Transactions on Automatic Control*, vol. 47, no. 2, pp. 210–224, Feb. 2002.

[86] J. Zhao and I. Kanellakopoulos, “Active identification for discrete-time nonlinear control. ii. strict-feedback systems,” *IEEE Transactions on Automatic Control*, vol. 47, no. 2, pp. 225–240, Feb. 2002.

[87] P. V. Zhivoglyadov, R. H. Middleton, and M. Fu, “Localization based switching adaptive control for time-varying discrete-time systems,” *IEEE Transactions on Automatic Control*, vol. 45, no. 4, pp. 752–755, Apr. 2000.

[88] P. V. Zhivoglyadov, R. H. Middleton, and M. Fu, “Adaptive stabilization of uncertain discrete-time systems via switching control: the method of localization,” in *Adaptive Control Systems: Advanced Techniques for Practical Application*, G. Feng and R. Lozano, Eds., Oxford: Newnes, 1999.

[89] Q. Zhu, J.-X. Xu, S. Yang, and G.-D. Hu, “Adaptive backstepping repetitive learning control design for nonlinear discrete-time systems with periodic uncertainties,” *International Journal of Adaptive Control and Signal Processing*, vol. 29, no. 4, pp. 524–535, 2015.

APPENDICES

Appendix A

Pole-Placement using Polynomials: Placing all Poles at the Origin

Here we briefly discuss, details about the pole-placement design process to be carried out in solving (6.6) and (7.21) at every time step.

Existence of A Solution

Given the following polynomials associated with a plant

$$\mathbf{A}(z^{-1}) = a_0 + a_1 z^{-1} + \cdots + a_{n_A-1} z^{-n_A+1} + a_{n_A} z^{-n_A}$$

and

$$\mathbf{B}(z^{-1}) = b_1 z^{-1} + b_2 z^{-2} + \cdots + b_{n_B-1} z^{-n_B+1} + b_{n_B} z^{-n_B},$$

we want to design a *strictly* proper controller by choosing polynomials

$$\mathbf{L}(z^{-1}) = l_0 + l_1 z^{-1} + \cdots + l_{n_A-1} z^{-n_L+1} + l_{n_A} z^{-n_L}$$

and

$$\mathbf{P}(z^{-1}) = p_1 z^{-1} + p_2 z^{-2} + \cdots + p_{n_P-1} z^{-n_P+1} + p_{n_P} z^{-n_P},$$

so that to place all closed-loop poles at the origin, i.e. we want to satisfy the following equation:

$$\mathbf{A}(z^{-1})\mathbf{L}(z^{-1}) + \mathbf{B}(z^{-1})\mathbf{P}(z^{-1}) = 1. \tag{A.1}$$

The following result presents the conditions on obtaining $\mathbf{L}(z^{-1})$ and $\mathbf{P}(z^{-1})$.

Theorem A.1 (**Theorem 2.3.1 of Ioannou and Sun** [31]). *If polynomials $\mathbf{a}(z)$ and $\mathbf{b}(z)$ are coprime and of degree n_a and n_b, respectively, where $n_a > n_b$, then for any given arbitrary polynomial $\mathbf{a}^*(z)$ of degree $n_{a^*} \geq n_a$, the polynomial equation*

$$\mathbf{a}(z)\mathbf{l}(z) + \mathbf{b}(z)\mathbf{p}(z) = \mathbf{a}^*(z)$$

has a unique solution $\mathbf{l}(z)$ and $\mathbf{p}(z)$ whose degrees n_l and n_p, respectively, satisfy the constraints $n_p < n_a$, $n_l \geq \max\{n_{a^*} - n_a, n_b - 1\}$.

To this end, if we define the following:

$$\begin{aligned}
\mathbf{A}'(z) &:= z^{n_A}\mathbf{A}(z^{-1}), \\
\mathbf{B}'(z) &:= z^{n_B}\mathbf{B}(z^{-1}), \\
\mathbf{L}'(z) &:= z^{n_L}\mathbf{L}(z^{-1}), \\
\mathbf{P}'(z) &:= z^{n_P}\mathbf{P}(z^{-1}),
\end{aligned}$$

and $n_{A^*} := \max\{n_A + n_L, n_B + n_P - 2\}$, observe that satisfying (A.1) is equivalent to satisfying the following equation:

$$\mathbf{A}'(z)\mathbf{L}'(z) + \mathbf{B}'(z)\mathbf{P}'(z) = z^{n_{A^*}}.$$

Observe that polynomials $\mathbf{A}'(z)$, $\mathbf{B}'(z)$, $\mathbf{L}'(z)$ and $\mathbf{P}'(z)$ have degrees of

$$n_A,\ n_B - 1,\ n_L \ \text{ and } \ n_P - 1,$$

respectively. This means that we have flexibility on the choices of n_L and n_P and therefore n_{A^*} to meet the conditions of Theorem A.1; this flexibility is of course a special case because, generally, in pole-placement design we do not necessarily place all poles at zero.

Examples:

- For the case of a known plant order and a stability objective, i.e. Chapter 6, we have $n_A = n_B = n$ and the the natural choices of $n_L = n_P = n$ as well; so $n_{A^*} = 2n$. If we check the conditions of Theorem A.1 we know that $\mathbf{a}(z) \equiv \mathbf{A}'(z)$ and $\mathbf{b}(z) \equiv \mathbf{B}'(z)$ are coprime, and $n_a \equiv n$, $n_b \equiv n - 1$. Then the choices of $\mathbf{a}^*(z) \equiv z^{2n}$, $\mathbf{l}(z) \equiv \mathbf{L}'(z)$ and $\mathbf{p}(z) \equiv \mathbf{P}'(z)$ are appropriate as $n_{a^*} \equiv 2n$, $n_l \equiv n$ and $n_p \equiv n - 1$.

- For the case of a unknown plant order and tracking objective, i.e. Chapter 7, with $n_i \in \{1, 2, \dots, \bar{n}\}$, we have $n_A = n_i + g$ and $n_B = n_i$ and the natural choices of $n_L = n_i$ and $n_P = n_i + g$; so we can choose $n_{A^*} = 2n_i + g$. If we check the conditions of Theorem A.1 we see that $\mathbf{a}(z) \equiv \mathbf{A}'(z)$ and $\mathbf{b}(z) \equiv \mathbf{B}'(z)$ are coprime, and $n_a \equiv n_i + g$, $n_b \equiv n_i - 1$. Then the choices of $\mathbf{a}^*(z) \equiv z^{2n_i+g}$, $\mathbf{l}(z) \equiv \mathbf{L}'(z)$ and $\mathbf{p}(z) \equiv \mathbf{P}'(z)$ are appropriate as $n_{a^*} \equiv 2n_i + g$, $n_l \equiv n_i$ and $n_p \equiv n_i + g - 1$.

Design Steps

Given polynomials $\mathbf{A}(z^{-1})$ and $\mathbf{B}(z^{-1})$ as defined in the previous section, and n_L and n_P chosen appropriately, then we can construct the non-singular matrix $M \in \mathbb{R}^{(n_L+n_P+1)\times(n_L+n_P+1)}$:

$$M := \underbrace{\begin{bmatrix} a_0 & 0 & \cdots & 0 \\ a_1 & a_0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n_A-1} & a_{n_A-2} & \ddots & a_0 \\ a_{n_A} & a_{n_A-1} & \ddots & a_1 \\ 0 & a_{n_A} & \ddots & a_2 \\ \vdots & 0 & \ddots & \vdots \\ & \vdots & & a_{n_A} \\ 0 & & & \end{bmatrix}}_{n_L+1 \text{ columns}} \hspace{-0.5em} \underbrace{\begin{bmatrix} 0 & \cdots & 0 \\ b_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ b_{n_B-1} & \ddots & b_1 \\ b_{n_B} & \ddots & \vdots \\ 0 & \ddots & b_{n_B-1} \\ \vdots & \ddots & b_{n_B} \\ & & \\ 0 & & \end{bmatrix}}_{n_P \text{ columns}};$$

with $n_{A^*} = \max\{n_A + n_L, n_B + n_P - 2\}$, define

$$\boldsymbol{\alpha}^* := \mathbf{e}_1 \in \mathbb{R}^{n_{A^*}},$$

so if we define the vector of the unknown coefficients of $\mathbf{L}(z^{-1})$ and $\mathbf{P}(z^{-1})$ as

$$\mathbf{k} := \begin{bmatrix} l_0 \\ l_1 \\ \vdots \\ l_{n_L} \\ p_1 \\ p_2 \\ \vdots \\ p_{n_P} \end{bmatrix},$$

then we can find the solution by calculating the following:

$$\mathbf{k} = M^{-1}\boldsymbol{\alpha}^*.$$

Appendix B

Crude Bounds Proofs

Proof of Proposition 5.2. Fix $p \geq 0$. Let $t_0 \in \mathbb{Z}$, $t \geq t_0$, $\mathcal{X}(t_0) \in \mathbb{R}^n$, $\hat{\theta}(t_0) \in \mathcal{S}$, $\theta^* \in \mathcal{S}^*$, and $v, w, y^* \in \boldsymbol{\ell}_\infty$ be arbitrary.

From the definition of $\mathcal{Y}(t+1)$ it follows immediately that

$$\|\mathcal{Y}(t+1)\| \leq \|\mathcal{Y}(t)\| + |y(t+1)|. \tag{B.1}$$

From (5.10) and (5.11) we obtain

$$|y(t+1)| \leq \|\mathcal{S}^*\| \|\phi(t)\| + |w(t)| + |v(t+1)| + \|\mathcal{S}^*\| c_\varphi \sum_{q=1}^{n} |v(t-q+1)|. \tag{B.2}$$

From (5.21) and Assumptions 5.2, we obtain

$$\begin{aligned} |u(t)| &\leq c_\theta \left(\left\| \hat{\bar{\theta}}(t) \right\| \left\| \varphi\big(\mathcal{Y}(t)\big) \right\| + |y^*(t+1)| \right) \\ &\leq c_\theta \|\mathcal{S}\| \|\varphi\big(\mathcal{Y}(t)\big)\| + c_\theta |y^*(t+1)| \\ &\leq c_\theta c_\varphi \|\mathcal{S}\| \|\mathcal{Y}(t)\| + c_\theta |y^*(t+1)|; \end{aligned} \tag{B.3}$$

we now combine this with (5.9) to obtain a bound on $\phi(t)$:

$$\|\phi(t)\| \leq c_\varphi (1 + c_\theta \|\mathcal{S}\|) \|\mathcal{Y}(t)\| + c_\theta |y^*(t+1)|.$$

Substituting this bound into (B.2) and then combining with (B.1), we obtain

$$\|\mathcal{Y}(t+1)\| \leq \underbrace{(1+\|\mathcal{S}^*\|c_\varphi(1+c_\theta\|\mathcal{S}\|))}_{=:c_1}\|\mathcal{Y}(t)\| + c_\theta\|\mathcal{S}^*\||y^*(t+1)|+$$
$$|w(t)| + |v(t+1)| + c_\varphi\|\mathcal{S}^*\|\sum_{q=1}^{n}|v(t-q+1)|; \tag{B.4}$$

then using (5.13) and simplifying the above we end up with

$$\|\mathcal{Y}(t+1)\| \leq c_1\|\mathcal{Y}(t)\| + \underbrace{(c_\theta\|\mathcal{S}^*\| + (n+2)(1+c_\varphi\|\mathcal{S}^*\|))}_{=:c_2}[|y^*(t+1)| + |\bar{w}(t)|]. \tag{B.5}$$

Using the fact that $c_1, c_2 \geq 1$, if we solve the above iteratively for p steps and define $\gamma_1 := c_1^p c_2$, then we conclude that (5.24) holds. ■

Proof of Proposition 6.2. Fix $p \geq 0$. Let $t_0 \in \mathbb{Z}$, $t \geq t_0$, $\phi(t_0) \in \mathbb{R}^{2n}$, $\sigma_0 \in \{1, 2\}$, $N \geq 1$, $\theta^* \in \mathcal{S}^*$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ $(i = 1, 2)$ and $w \in \boldsymbol{\ell}_\infty$ be arbitrary.

From (6.1) we see that

$$|y(t+1)| \leq \bar{\mathbf{s}}\|\phi(t)\| + |w(t)|.$$

From (6.13) and compactness (Assumption 6.1), we have that there exists a constant γ so that

$$|u(t+1)| \leq \gamma\|\phi(t)\|.$$

From the definition of $\|\phi(t+1)\|$, we have that

$$\|\phi(t+1)\| \leq \|\phi(t)\| + |y(t+1)| + |u(t+1)|.$$

Combining these three bounds, we end up with

$$\|\phi(t+1)\| \leq \underbrace{(1+\bar{\mathbf{s}}+\gamma)}_{=:\bar{a}}\|\phi(t)\| + |w(t)|.$$

Solving iteratively, we have

$$\|\phi(t+p)\| \leq \bar{a}^p\|\phi(t)\| + \sum_{j=0}^{p-1}\bar{a}^{p-j-1}|w(t+j)|$$

$$\leq \bar{a}^p \|\phi(t)\| + \bar{a}^{p-1} \sum_{j=0}^{p-1} |w(t+j)|.$$

Put $\bar{c} := \bar{a}^p$ to conclude the proof. ■

Proof of Lemma 7.2. Fix $p \geq 0$. Let $n \in \{1, 2, \ldots, \bar{n}\}$ and $\theta \in \mathbf{\Theta}_n$, $t_0 \in \mathbb{Z}$, $t \geq t_0$, $N \geq 1$, $\sigma_0 \in \mathcal{I}^*$, $\bar{\phi}(t_0) \in \mathbb{R}^{2(\bar{n}+g)}$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ $(i \in \mathcal{I}^*)$ and $w, y^* \in \boldsymbol{\ell}_\infty$ be arbitrary.

Part (i): From the associated auxiliary plant (7.12) we see that

$$|\varepsilon(t+1)| \leq \|\theta^*\| \|\psi^*(t)\| + |\bar{w}(t)| \leq \bar{\mathbf{s}} \|\bar{\psi}(t)\| + |\bar{w}(t)|.$$

From (7.25) and compactness, we have that there exists a constant γ_1 so that

$$|v(t+1)| \leq \gamma_1 \|\bar{\psi}(t)\|. \tag{B.6}$$

From the definition of $\|\bar{\psi}(t+1)\|$ given in (7.14), we have that

$$\|\bar{\psi}(t+1)\| \leq \|\bar{\psi}(t)\| + |\varepsilon(t+1)| + |v(t+1)|.$$

Combining these three bounds, we end up with

$$\|\bar{\psi}(t+1)\| \leq \underbrace{(1 + \bar{\mathbf{s}} + \gamma_1)}_{=:\gamma_2} \|\bar{\psi}(t)\| + |\bar{w}(t)|.$$

We solve this iteratively for p steps and put $\bar{c}_1 := \gamma_2^p$ to conclude the proof of part (i).

Part (ii): First define $\gamma_3 := \max_{n \in \{1,\ldots,\bar{n}\}} \|\mathbf{\Theta}_n\|$; from the plant equation (7.1) we obtain

$$|y(t+1)| \leq \|\theta\| \|\bar{\phi}(t)\| + |w(t)| \leq \gamma_3 \|\bar{\phi}(t)\| + |w(t)|.$$

From (7.26) we see that there exists a constant γ_4 such that

$$|u(t+1)| \leq |v(t+1)| + \gamma_4 \|\bar{\phi}(t)\|.$$

From the definition of $\|\bar{\phi}(t+1)\|$, we have that

$$\|\bar{\phi}(t+1)\| \leq \|\bar{\phi}(t)\| + |y(t+1)| + |u(t+1)|.$$

Combining these three bounds and (B.6), we end up with

$$\|\bar{\phi}(t+1)\| \leq \underbrace{(1+\gamma_3+\gamma_4)}_{=:\gamma_5}\|\bar{\phi}(t)\| + \gamma_1|\bar{\psi}(t)| + |w(t)|. \tag{B.7}$$

From the definition of $\bar{\psi}(t)$ and definition of the auxiliary input $v(\cdot)$, we see that $\bar{\psi}(t)$ consists of $y(t)-y^*(t), y(t-1)-y^*(t-1), \ldots, y(t-\bar{n}-g+1)-y^*(t-\bar{n}-g+1)$ and weighted sums of $u(t), u(t-1), \ldots, u(t-\bar{n}-g+1)$; so by the definition of $\bar{\phi}(\cdot)$ we see that there exists a constant γ_6 such that

$$\|\bar{\psi}(t)\| \leq \gamma_6\|\bar{\phi}(t)\| + \sum_{j=0}^{\bar{n}+g-1} |y^*(t-j)|.$$

After substituting the above into (B.7), we obtain

$$\begin{aligned}\|\bar{\phi}(t+1)\| \leq (\gamma_5+\gamma_6\gamma_1)\|\bar{\phi}(t)\| + & \\ \gamma_1 \sum_{j=0}^{\bar{n}+g-1} |y^*(t-j)| + |w(t)|.&\end{aligned}$$

Using the definition of $\tilde{y}^*$ in (7.34), if we solve the above iteratively for p steps and put $\bar{c}_2 := (\gamma_1+1)(\gamma_5+\gamma_6\gamma_1)^p$, then we conclude that part (ii) holds. ■

Appendix C

Proofs of the Tracking Results of Chapter 5

Proof of Theorem 5.3. Fix $\delta \in (0, \infty]$ and $\lambda \in (0, 1)$. Let $t_0 \in \mathbb{Z}$, $\theta^* \in \mathcal{S}^*$, $y^* \in \boldsymbol{\ell}_\infty$, $\hat{\theta}(t_0) \in \mathcal{S}$, and $\mathcal{X}(t_0)$ be arbitrary. Now suppose that $w = v = 0$; for this case, by the definition of the function ρ, for $t \geq t_0 + 1$ we have $\bar{w} = 0$ and

$$\rho(t) = 0 \quad \Leftrightarrow \quad \|\tilde{\phi}(t)\| = 0.$$

If we incorporate the fact that $\varepsilon(t) = e(t)$ for $t \geq t_0 + 1$ into the result of Proposition 5.1, we obtain

$$\begin{aligned}
\sum_{t=t_0+1, \|\tilde{\phi}(t-1)\| \neq 0}^{T} \frac{|\varepsilon(t)|^2}{\|\tilde{\phi}(t-1)\|^2} &= \sum_{t=t_0+1, \|\tilde{\phi}(t-1)\| \neq 0}^{T} \frac{|e(t)|^2}{\|\tilde{\phi}(t-1)\|^2} \\
&\leq \sum_{t=t_0+1, \|\tilde{\phi}(t-1)\| \neq 0}^{\infty} \frac{|e(t)|^2}{\|\tilde{\phi}(t-1)\|^2} \\
&\leq 8\|\mathcal{S}\|^2.
\end{aligned}$$

When $\|\tilde{\phi}(t-1)\| = 0$ it follows that $\phi(t-1) = 0$, so it is easy to see by (5.10) that $y(t) = 0$ and by the control law (5.21) that $y^*(t) = 0$, which means that $\varepsilon(t) = 0$. Then from the above and by the definition of $\tilde{\phi}$, we obtain

$$\sum_{t=t_0+1}^{\infty} \varepsilon(t)^2 \leq 8\|\mathcal{S}\|^2 \sup_{j \geq t_0} \left[\|\phi(j)\|^2 + \|\mathcal{Y}(j)\|^2 \right].$$

We know that from the definition of the control law and by Assumption 5.2 that there exists a constant c_1 such that

$$\|\phi(j)\| \le \|\mathcal{Y}(j)\| + |u(j)| \le c_1 \left(\|\mathcal{Y}(j)\| + |y^*(j+1)|\right), \quad j \ge t_0;$$

by Theorem 5.1 we see that there exists a constant c_2 so that

$$\|\mathcal{Y}(j)\| \le c_2 \left(\|\mathcal{X}(t_0)\| + \sup_{j \ge t_0} |y^*(j+1)|\right), \quad j \ge t_0.$$

Incorporating all of the above and simplifying, we conculde that there exists a constant c so that

$$\sum_{t=t_0+1}^{\infty} \varepsilon(t)^2 \le c \left(\|\mathcal{X}(t_0)\|^2 + \|y^*\|_\infty^2\right)$$

as desired. ■

Proof of Theorem 5.4. Fix $\delta \in (0, \infty]$ and $\lambda_1 \in (0,1)$, and set $w = v = 0$. By Theorem 5.2 there exist constants $\gamma_1 > 0$ and

$$\bar{\epsilon} \in (0, 2^{\frac{3}{2}} \|\mathcal{S}\|)$$

so that for every $t_0 \in \mathbb{Z}$, $y^* \in \boldsymbol{\ell}_\infty$, $\hat{\theta}(t_0) \in \mathcal{S}$, $\mathcal{X}(t_0) \in \mathbb{R}^n$, and $\theta^* \in \mathbb{S}(\mathcal{S}^*, 0, \bar{\epsilon})$, when the adaptive controller (5.16) and (5.21) is applied to the plant (5.65), the following holds:

$$\left\| \begin{bmatrix} \mathcal{X}(t) \\ u(t) \end{bmatrix} \right\| \le \gamma_1 \lambda_1^{t-t_0} \|\mathcal{X}(t_0)\| + \frac{\gamma_1}{1-\lambda_1} \sup_{j \in [t_0, t]} |y^*(j+1)|, \qquad t \ge t_0. \tag{C.1}$$

Now, let $t_0 \in \mathbb{Z}$, $y^* \in \boldsymbol{\ell}_\infty$, $\hat{\theta}(t_0) \in \mathcal{S}$, $\mathcal{X}(t_0) \in \mathbb{R}^n$, $\epsilon \in (0, \bar{\epsilon})$ and $\theta^* \in \mathbb{S}(\mathcal{S}^*, 0, \epsilon)$ be arbitrary. Let $t_i \ge t_0$ be arbitrary; then from plant equation (5.65) we have

$$y(t+1) \quad = \quad \phi(t)^\top \theta^*(t_i) + \underbrace{\phi(t)^\top \left[\theta^*(t) - \theta^*(t_i)\right]}_{=:\Delta_i(t)}. \tag{C.2}$$

Note from the definition of $\mathbb{S}(\mathcal{S}^*, 0, \bar{\epsilon})$ given in Definition 2.2 that

$$\|\Delta_i(t)\| \le \epsilon \|\phi(t)\| \left(t - t_i\right).$$

Define $\tilde{\theta}_i(t) := \hat{\theta}(t) - \theta^*(t_i)$. Since $w = v = 0$ and $\|\phi(t)\| \le \|\tilde{\phi}(t)\|$, by applying Proposition

5.1 to the plant (C.2), we obtain[1]

$$\begin{aligned}
\sum_{j=t_i,\|\tilde{\phi}(j)\|\neq 0}^{t-1} \frac{|\varepsilon(j+1)|^2}{\|\tilde{\phi}(j)\|^2} &= \sum_{j=t_i,\|\tilde{\phi}(j)\|\neq 0}^{t-1} \frac{|e(j+1)|^2}{\|\tilde{\phi}(j)\|^2} \\
&\leq 2\|\tilde{\theta}(t_i)\|^2 + \sum_{j=t_i,\|\tilde{\phi}(j)\|\neq 0}^{t-1} \frac{4\|\Delta_i(j)\|^2}{\|\tilde{\phi}(j)\|^2} \\
&\leq 2\|\tilde{\theta}_i(t_i)\|^2 + 4\epsilon^2 \sum_{j=t_i}^{t-1} (j-t_i)^2 \\
&= 2\|\tilde{\theta}_i(t_i)\|^2 + 4\epsilon^2 \sum_{k=0}^{t-t_i-1} k^2 \\
&\leq 8\|\mathcal{S}\|^2 + 4\epsilon^2 (t-t_i-1)^3, \quad t \geq t_i+1,\ t_i \geq t_0.
\end{aligned}$$

Since the disturbance is zero here, it follows that $\|\tilde{\phi}(j)\| = 0$ implies that $\varepsilon(j+1) = 0$[2]. So from the above we conclude that

$$\sum_{j=t_i}^{t-1} |\varepsilon(j+1)|^2 \leq 4\left[2\|\mathcal{S}\|^2 + \epsilon^2(t-t_i-1)^3\right] \sup_{j\in[t_i,t)} \|\tilde{\phi}(j)\|^2, \ \ t \geq t_i+1, \quad t_i \geq t_0. \tag{C.3}$$

We now analyze the average tracking error. From (C.3) we obtain

$$\frac{1}{t-t_i}\sum_{j=t_i}^{t-1} |\varepsilon(j+1)|^2 \leq 4\left[\frac{2\|\mathcal{S}\|^2}{t-t_i} + \epsilon^2 \frac{(t-t_i-1)^3}{t-t_i}\right] \sup_{j\in[t_i,t)} \|\tilde{\phi}(j)\|^2, \quad t \geq t_i+1, \quad t_i \geq t_0. \tag{C.4}$$

Now define

$$\beta_\epsilon := \left(\epsilon\|\mathcal{S}\|^2\right)^{\frac{2}{3}} \tag{C.5}$$

and $T_\beta \in \mathbb{N}$ by

$$T_\beta := \left\lceil \frac{2\|\mathcal{S}\|^2}{\beta_\epsilon} \right\rceil;$$

[1]Observe that $\sum_{k=1}^{m} k^2 = \frac{m(m+1)(2m+1)}{6} = \frac{m^3}{3} + \frac{m^2}{2} + \frac{m}{6} \leq m^3$.

[2]Observe that this is true even when $\theta^*(t)$ is time-varying.

this means that

$$\frac{2\|\mathcal{S}\|^2}{T_\beta} \leq \beta_\epsilon.$$

We can easily check that $\frac{(T_\beta - 1)^3}{T_\beta} \leq 8T_\beta^2$. Incorporating this and the definition of T_β into (C.4), by choosing $t = t_i + T_\beta$ we have

$$\begin{aligned}
\frac{1}{T_\beta} \sum_{j=t_i}^{t_i+T_\beta-1} |\varepsilon(j+1)|^2 &\leq 4 \left[\frac{2\|\mathcal{S}\|^2}{T_\beta} + 8\epsilon^2 T_\beta^2 \right] \sup_{j \in [t_i, t_i+T_\beta)} \|\tilde{\phi}(j)\|^2 \\
&\leq 4 \left[\beta_\epsilon + 8\epsilon^2 T_\beta^2 \right] \sup_{j \in [t_i, t_i+T_\beta)} \|\tilde{\phi}(j)\|^2, \; t_i \geq t_0. \qquad \text{(C.6)}
\end{aligned}$$

We would like to obtain a bound on $\epsilon^2 T_\beta^2$. But it follows from (C.5) that

$$\beta_\epsilon^3 = \epsilon^2 \|\mathcal{S}\|^4,$$

so

$$\epsilon^2 = \beta_\epsilon \left(\frac{\beta_\epsilon}{\|\mathcal{S}\|^2} \right)^2 = 4\beta_\epsilon \left(\frac{\beta_\epsilon}{2\|\mathcal{S}\|^2} \right)^2;$$

if we define $x = \frac{\beta_\epsilon}{2\|\mathcal{S}\|^2}$, then we see that

$$\begin{aligned}
\epsilon^2 T_\beta^2 &= 4\beta_\epsilon x^2 \left(\left\lceil \frac{1}{x} \right\rceil \right)^2 \\
&\leq 4\beta_\epsilon x^2 \left(\frac{1}{x} + 1 \right)^2 \\
&= 4\beta_\epsilon (x+1)^2.
\end{aligned}$$

But $\epsilon \in (0, 2^{3/2}\|\mathcal{S}\|)$ by hypothesis, so

$$m = \frac{\beta_\epsilon}{2\|\mathcal{S}\|^2} = \frac{(\epsilon\|\mathcal{S}\|^2)^{2/3}}{2\|\mathcal{S}\|^2} < \frac{(2^{3/2}\|\mathcal{S}\| \times \|\mathcal{S}\|^2)^{2/3}}{2\|\mathcal{S}\|^2} = 1,$$

so

$$\epsilon^2 T_\beta^2 \leq 16\beta_\epsilon.$$

Substituting this into (C.6) and simplifying yields

$$\frac{1}{T_\beta}\sum_{j=t_i}^{t_i+T_\beta-1}|\varepsilon(j+1)|^2 \le 516\beta_\epsilon \sup_{j\in[t_i,t_i+T_\beta)}\|\tilde{\phi}(j)\|^2, \qquad t_i \ge t_0. \tag{C.7}$$

We now analyze the average tracking error over the whole time horizon; we do so by considering time intervals of length T_β. From (C.7) we easily obtain

$$\frac{1}{iT_\beta}\sum_{j=\bar{t}}^{\bar{t}+iT_\beta-1}|\varepsilon(j+1)|^2 \le 516\beta_\epsilon \sup_{j\in[\bar{t},\bar{t}+iT_\beta)}\|\tilde{\phi}(j)\|^2,\ i\in\mathbb{N},\ \bar{t}\ge t_0. \tag{C.8}$$

The bound in (C.8) provides a bound on the average tracking error over time intervals of lengths that are multiples of T_β. To extend this to intervals of arbitrary length, first observe that (C.8) can be rewritten as

$$\sum_{j=\bar{t}}^{\bar{t}+iT_\beta-1}|\varepsilon(j+1)|^2 \le iT_\beta(\underbrace{516}_{=:\gamma_2}\beta_\epsilon) \sup_{j\in[\bar{t},\bar{t}+iT_\beta)}\|\tilde{\phi}(j)\|^2,\ i\in\mathbb{N},\ \bar{t}\ge t_0.$$

For $k\in\{0,1,...,T_\beta-1\}$, this inequality implies that

$$\sum_{j=\bar{t}+k}^{\bar{t}+iT_\beta-1+k}|\varepsilon(j+1)|^2 \le i\gamma_2 T_\beta\beta_\epsilon \sup_{j\in[\bar{t}+k,\bar{t}+iT_\beta+k)}\|\tilde{\phi}(j)\|^2,\ i\in\mathbb{N},\ \bar{t}\ge t_0;$$

adding these two inequalities and simplifying yields

$$\sum_{j=\bar{t}}^{\bar{t}+iT_\beta-1+k}|\varepsilon(j+1)|^2 \le 2i\gamma_2 T_\beta\beta_\epsilon \sup_{j\in[\bar{t},\bar{t}+iT_\beta+k)}\|\tilde{\phi}(j)\|^2,\ i\in\mathbb{N},\ k\in\{0,1,...,T_\beta-1\},\ \bar{t}\ge t_0.$$

Changing variables to enhance clarity, we see that this implies that

$$\sum_{j=\bar{t}}^{\bar{t}+T-1}|\varepsilon(j+1)|^2 \le 2\gamma_2 T\beta_\epsilon \sup_{j\in[\bar{t},\bar{t}+T-1)}\|\tilde{\phi}(j)\|^2,\ T\ge T_\beta,\ \bar{t}\ge t_0,$$

which means that

$$\frac{1}{T}\sum_{j=\bar{t}}^{\bar{t}+T-1}|\varepsilon(j+1)|^2 \le 2\gamma_2\beta_\epsilon \sup_{j\in[\bar{t},\bar{t}+T-1)}\|\tilde{\phi}(j)\|^2,\ T\ge T_\beta,\ \bar{t}\ge t_0. \tag{C.9}$$

This means that

$$\limsup_{T\to\infty}\frac{1}{T}\sum_{j=\bar{t}}^{\bar{t}+T-1}|\varepsilon(j+1)|^2 \le 2\gamma_2\beta_\epsilon\times\limsup_{j\to\infty}\|\tilde{\phi}(j)\|^2,\ \bar{t}\ge t_0.$$

From the control law, by Assumption 5.2 and (C.1) we see that there exists a constant c_1 so that

$$\limsup_{j\to\infty}\|\tilde{\phi}(j)\| \le \limsup_{j\to\infty}(\|\phi(j)\|+\|\mathcal{X}(j)\|)\le c_1\|y^*\|_\infty;$$

the boundedness of the tracking error ensures that for every $\bar{t}\ge t_0$ we have

$$\limsup_{T\to\infty}\frac{1}{T}\sum_{j=\bar{t}}^{\bar{t}+T-1}|\varepsilon(j+1)|^2 = \limsup_{T\to\infty}\frac{1}{T}\sum_{j=t_0}^{t_0+T-1}|\varepsilon(j+1)|^2,\ \bar{t}\ge t_0,$$

so we conclude that

$$\limsup_{T\to\infty}\frac{1}{T}\sum_{j=t_0}^{t_0+T-1}|\varepsilon(j+1)|^2 \le 2c_1^2\gamma_2\beta_\epsilon\|y^*\|_\infty^2.$$

Since $\beta_\epsilon=\|\mathcal{S}\|^{4/3}\epsilon^{2/3}$, the result follows. ■

Proof of Theorem 5.5. Fix $\delta\in(0,\infty]$ and $\lambda\in(0,1)$. Let $t_0\in\mathbb{Z}$, $\theta^*\in\mathcal{S}^*$, $y^*,w,v\in\boldsymbol{\ell}_\infty$, $\hat{\theta}(t_0)\in\mathcal{S}$, and $\mathcal{X}(t_0)\in\mathbb{R}^n$ be arbitrary, but so that $\liminf_{t\to\infty}|y^*(t)|>0$. Before proceeding, choose $\underline{t}\ge t_0+1$ so that

$$\inf\{|y^*(t)|:\ t\ge\underline{t}\}>0.$$

Now, by applying Proposition 5.1, for $\bar{t}\ge t_0+1$ we obtain

$$\begin{aligned}\sum_{j=\bar{t}}^{t}\rho(j-1)\frac{|\varepsilon(j)|^2}{\|\tilde{\phi}(j-1)\|^2} &= \sum_{j=\bar{t}}^{t}\rho(j-1)\frac{|e(j)|^2}{\|\tilde{\phi}(j-1)\|^2}\\ &\le 8\|\mathcal{S}\|^2+2\bar{c}\sum_{j=\bar{t}}^{t}\rho(j-1)\frac{|\bar{w}(j-1)|^2}{\|\tilde{\phi}(j-1)\|^2}\end{aligned}$$

$$= 2\bar{c}\|\mathcal{S}\|^2 \left(\frac{4}{\bar{c}} + \frac{1}{\|\mathcal{S}\|^2} \sum_{j=\bar{t}}^{t} \rho(j-1) \frac{|\bar{w}(j-1)|^2}{\|\tilde{\phi}(j-1)\|^2} \right)$$
$$= 2\bar{c}\|\mathcal{S}\|^2 \left(\frac{4}{\bar{c}} + \sum_{j=\bar{t}}^{t} \rho(j-1) \frac{|\bar{w}(j-1)|^2}{\left(\|\mathcal{S}\|\|\tilde{\phi}(j-1)\|\right)^2} \right), \qquad t \geq \bar{t} \geq t_0 + 1. \tag{C.10}$$

From the controller equation (5.21) we have

$$y^*(t) = \hat{\theta}(t-1)^\top \phi(t-1), \qquad t \geq t_0 + 1,$$

which means that

$$|y^*(t)| \leq \|\hat{\theta}(t-1)\|\|\phi(t-1)\| \leq \|\mathcal{S}\|\|\phi(t-1)\| \leq \|\mathcal{S}\|\|\tilde{\phi}(t-1)\|, \qquad t \geq t_0 + 1;$$

if we substitute this into (C.10), then we obtain

$$\sum_{j=\bar{t}}^{t} \rho(j-1) \frac{|\varepsilon(j)|^2}{\|\tilde{\phi}(j-1)\|^2} \leq 2\bar{c}\|\mathcal{S}\|^2 \left(\frac{4}{\bar{c}} + \sum_{j=\bar{t}}^{t} \rho(j-1) \frac{|\bar{w}(j-1)|^2}{|y^*(j)|^2} \right), \qquad t \geq \bar{t} \geq \underline{t}. \tag{C.11}$$

Now we analyze the above bound for two cases: when the estimator is turned on, i.e. when $\rho(\cdot) = 1$, and when the estimator is turned off, i.e. when $\rho(\cdot) = 0$. Before proceeding, we define some notation: for $t_2 \geq t_1 \geq t_0$, we define

$$\underline{\boldsymbol{y}}^*_{[t_1,t_2]} := \inf_{j \in [t_1,t_2],\, \rho(j-1)=1} |y^*(j)|^2.$$

Case 1: The estimator is turned on: $\rho(j-1) = 1$.

From (C.11), we have

$$\sum_{j=\bar{t},\, \rho(j-1)=1}^{t} \frac{|\varepsilon(j)|^2}{\|\tilde{\phi}(j-1)\|^2} \leq 2\bar{c}\|\mathcal{S}\|^2 \left(\frac{4}{\bar{c}} + \frac{1}{\underline{\boldsymbol{y}}^*_{[\bar{t},t]}} \sum_{j=\bar{t}}^{t} |\bar{w}(j-1)|^2 \right), \qquad t \geq \bar{t} \geq \underline{t}, \tag{C.12}$$

which means that

$$\sum_{j=\bar{t},\rho(j-1)=1}^{t} |\varepsilon(j)|^2 \leq 2\bar{c}\|\mathcal{S}\|^2 \left(\sup_{j\in[\bar{t}-1,t-1]} \|\tilde{\phi}(j)\|^2\right) \left(\frac{4}{\bar{c}} + \frac{1}{\underline{\boldsymbol{y}}^*_{[\bar{t},t]}} \sum_{j=\bar{t}}^{t} |\bar{w}(j-1)|^2\right),\ t \geq \bar{t} \geq \underline{t}. \tag{C.13}$$

Case 2: The estimator is turned off: $\rho(j-1) = 0$.

In this case, we know from the definition of ρ that when $\rho(t-1) = 0$:

- if $\delta = \infty$ then $\tilde{\phi}(t-1) = 0$, so

$$\|\tilde{\phi}(t-1)\| \leq \frac{1}{\delta}|\bar{w}(t-1)|;$$

- if $\delta \in (0, \infty)$, then we have that

$$|e(t)| \geq (2\|\mathcal{S}\| + \delta)\|\tilde{\phi}(t-1)\|;$$

  using the prediction error definition in (5.14), equation (5.10) and (5.11) we conclude (by similar analysis to that done in (5.36)) that

$$|e(t)| \leq 2\|\mathcal{S}\|\|\tilde{\phi}(t-1)\| + (1 + c_\varphi\|\mathcal{S}^*\|)(n+2)|\bar{w}(t-1)|, \quad t \geq t_0 + 1; \tag{C.14}$$

  combining these two equations yields

$$\|\tilde{\phi}(t-1)\| \leq \frac{(1 + c_\varphi\|\mathcal{S}^*\|)(n+2)}{\delta}|\bar{w}(t-1)|.$$

Using (C.14) and applying bounds on $\tilde{\phi}(t-1)$ found above, we obtain

$$|\varepsilon(t)| = |e(t)| \leq \underbrace{(1 + c_\varphi\|\mathcal{S}^*\|)(n+2)\left(\frac{2\|\mathcal{S}\|}{\delta} + 1\right)}_{=:c_1}|\bar{w}(t-1)|, \quad t \geq t_0 + 1.$$

Hence,

$$\sum_{j=\bar{t},\,\rho(j-1)=0}^{t} |\varepsilon(j)|^2 \leq c_1^2 \sum_{j=\bar{t}}^{t} |\bar{w}(j-1)|^2, \quad t \geq \bar{t} \geq t_0 + 1, \tag{C.15}$$

which concludes Case 2.

We can now combine (C.13) and (C.15) of Case 1 and Case 2, respectively, to yield

$$\sum_{j=\bar{t}}^{t} |\varepsilon(j)|^2 \le 8\|\mathcal{S}\|^2 \left(\sup_{j\in[\bar{t},t]} \|\tilde{\phi}(j+1)\|^2 \right) + \\ \max\left\{ c_1^2, 2\bar{c}\|\mathcal{S}\|^2 \left(\frac{\sup_{j\in[\bar{t},t]} \|\tilde{\phi}(j+1)\|^2}{\underline{\boldsymbol{y}}^*_{[\bar{t},t]}} \right) \right\} \left(\sum_{j=\bar{t}}^{t} |\bar{w}(j-1)|^2 \right), \quad t \ge \bar{t} \ge \underline{t}. \tag{C.16}$$

By Theorem 5.1 there exists constants $c > 0$ and $\lambda \in (0,1)$ so that

$$\left\| \begin{bmatrix} \mathcal{Y}(t+1) \\ u(t+1) \end{bmatrix} \right\| \le c\lambda^{t+1-t_0} \|\mathcal{Y}(t_0)\| + \sum_{j=t_0}^{t+1} c\lambda^{t+1-j} (|y^*(j)| + |\bar{w}(j)|, \; t \ge \underline{t},$$

so we can choose $\bar{t} \ge \underline{t}$ (which depends implicitly on $\mathcal{X}(t_0)$, y^*, $\hat{\theta}(t_0)$, and θ^*) such that

$$\left\| \begin{bmatrix} \mathcal{Y}(t+1) \\ u(t+1) \end{bmatrix} \right\| \le \frac{2c}{1-\lambda} \limsup_{k\to\infty} (|y^*(\bar{t}+k)| + |\bar{w}(\bar{t}+k)|), \quad t \ge \bar{t},$$

as well as

$$|y^*(t)|^2 \ge \underbrace{\frac{1}{2} \liminf_{k\to\infty} |y^*(k)|^2}_{=:\underline{y}^*}, \quad t \ge \bar{t};$$

combining this with the definitions of $\tilde{\phi}$ and $\bar{w}$, and Assumption 5.2, we conclude that there exists a constant c_2 so that

$$\|\tilde{\phi}(t+1)\| \le c_2 \limsup_{k\to\infty} (|y^*(\bar{t}+k)| + |w(\bar{t}+k)| + |v(\bar{t}+k)|), \quad t \ge \bar{t}.$$

If we incorporate this into (C.16), then we obtain

$$\sum_{j=\bar{t}}^{t} |\varepsilon(j)|^2 \le 8\|\mathcal{S}\|^2 \left(\sup_{j\in[\bar{t},t]} \|\tilde{\phi}(j+1)\|^2 \right) + \left(\sum_{j=\bar{t}}^{t} |\bar{w}(j-1)|^2 \right) \times \\ \max\left\{ c_1^2, c_2^2 \frac{\limsup_{k\to\infty} (|y^*(\bar{t}+k)|^2 + |w(\bar{t}+k)|^2 + |v(\bar{t}+k)|^2)}{\frac{1}{2}\underline{y}^*} \right\}, \quad t \ge \bar{t},$$

which means that

$$\limsup_{T\to\infty}\frac{1}{T}\sum_{j=\bar{t}}^{\bar{t}+T-1}|\varepsilon(j)|^2 \le \limsup_{T\to\infty}\frac{1}{T}\sum_{j=\bar{t}}^{\bar{t}+T-1}|\bar{w}(j-1)|^2\times$$
$$\max\left\{c_1^2, c_2^2\frac{\limsup_{k\to\infty}(|y^*(\bar{t}+k)|^2+|w(\bar{t}+k)|^2+|v(\bar{t}+k)|^2)}{\frac{1}{2}\underline{y}^*}\right\}.$$

But $\bar{w}(t)^2$ is a weighted sum of $\{w(t)^2, v(t+1)^2, v(t)^2, \ldots, v(t+n-1)^2\}$, and the boundedness of all variables makes the starting point of the average sums irrelevant, so the desired bound (5.67) follows. ■