UNIVERSITY OF SOUTHAMPTON

Robust Stability and Performance for Multiple Model Switched Adaptive Control

by

Dominic Buchstaller

A thesis submitted for the degree of Doctor of Philosophy

in the
Faculty of Engineering, Science and Mathematics
School of Electronics and Computer Science
University of Southampton

January 2010

ABSTRACT

While the concept of switching between multiple controllers to achieve a control objective is not new, the available analysis to date imposes various structural and analytical assumptions on the controlled plant. The analysis presented in this thesis, which is concerned with an Estimation-based Multiple Model Switched Adaptive Control (EMMSAC) algorithm originating from Fisher-Jeffes (2003); Vinnicombe (2004), is shown not to have such limitations. As the name suggests, the key difference between EMMSAC and common multiple model type switching schemes is that the switching decision is based on the outcome of an optimal estimation process. The use of such optimal estimators is the key that allows for a simplified, axiomatic approach to analysis. Also, since estimators may be implemented by standard optimisation techniques, their construction is feasible for a broad class of systems.

The presented analysis is the first of its kind to provide comprehensive robustness and performance guarantees for a multiple model control algorithm, in terms of $l_p,\ 1 \leq p \leq \infty$ bounds on the closed loop gain, and is applicable to the class of minimal MIMO LTI plants. A key feature of this bound is that it permits the on-line alteration of the plant model set (dynamic EMMSAC) in contrast to the usual assumption that the plant model set is constant (static EMMSAC). It is shown that a static EMMSAC algorithm is conservative whereas a dynamic EMMSAC algorithm, based on the technique of dynamically expanding the plant model set, can be universal. It is also shown that the established gain bounds are invariant to a refinement of the plant model set, e.g. as a successive increasing fidelity sampling of a continuum of plants. Dynamic refinement of the plant model set is considered with the view to increase expected performance.

Furthermore, the established bounds — which are also a measure of performance — have the property that they are explicit in the free variables of the algorithm. It is shown that this property of the bound forms the basis for a principled, performance-orientated approach to design. Explicit, performance-orientated design examples are given and the trade off between dynamic and static constructions of plant model sets are investigated with respect to prior information on the acting disturbances and the uncertainty.

Contents

List of Figures

List of Tables

Declaration of authorship

I, DOMINIC BUCHSTALLER declare that the thesis entitled Robust Stability and Performance for Multiple Model Switched Adaptive Control and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- parts of this work have been published as: Buchstaller and French (2007, 2008, 2009).

Thursday 28$^{\text{th}}$ January, 2010

Acknowledgements

I would like to thank a few special people I had the honour to meet in my life, and who made this thesis (and much more) possible. Firstly, I would like to thank my father Manfred, who introduced me to the world of technology the right way — by letting me explore it; each screw, resistor and dissected household appliance at a time. I would like to thank my mother Uta, for being an inexhaustible source of support for everything I do. Sorry for all the minor mishaps and near catastrophes along the way. I would like to thank Franz Dreher, my science teacher, who understood to teach what science is all about — curiosity and imagination. The recurring courage to abandon the electronics lab to a couple of 17-year-olds with soldering irons will never be forgotten. I am also indebted to Florian Friesdorf and Christoph Hackl, great thinkers and friends, who made my time in Munich special. Night-long debates over Martini are dearly missed. However, most importantly, I would like to thank my adviser Mark French, an inspiring individual with a great mind, who always seems to ask the right questions. Mark recognised my potential, believed in me and provided the best supervision I could hope for. Thank you for all of that. I am looking forward to many pictures still to be drawn. Finally, I would like to thank my partner Bess and my sister Isabelle, indispensable sources of sanity and support, who fill my life with joy. Richard Bradley and Ivan Markovsky are dearly thanked for proof-reading parts of this thesis. I still have to get hold of that yacht I promised you Rich. Also many thanks to the control group at ECS Southampton for providing a friendly and inspiring research environment. Thanks to Banafshe Arbabzavar and Gabriele Gherbaz for being good friends.

Nomenclature

α	Bound on the growth of w_2 for the atomic closed loop $[P_p, C_{K(p)}]$
$B_\chi(p,r)$	Neighbourhood of plants around $P_p,\ p \in U$ of size $< r$ (as measured by χ)
χ	Measure of distance between plants in U
$\mathcal{C}$	Parametrised set of all controllers
$d_p[k]$	$d_p[k] = E(w_2)(p)(k)$, disturbance estimate to P_p at time k
$\vec{\delta}, \delta$	Directed gap, gap metric
D	Delay operator
Δ	Transition delay function
E	Disturbance estimator (operator)
$E(w_2)(p)(k)$	$E(w_2)(p)(k) = d_p[k]$, disturbance estimate to P_p at time k
F_k	Final switching times up to time k
$F_k(p, \nu(k)(p))$	Final switching times to plants in the neighbourhood $B(p, \nu(k)(p))$ around P_p up to time k
G	Plant-generating operator specifying the plant model set
H	Plant-generating operator defining with ν a cover for U
J	Norm weighting function
k_i	Switching time (physical or virtual)
$k_s(k)$	Last switching time up to time k
K	Controller design procedure
l	Attenuation function
λ	Horizon length of the estimator
l_i	Physical switching time
L_k	Set of physical switching times up to time k
m	Dimension of the input signal space
$map(A,B)$	Relation between elements in A and B, e.g. a function if this relation is unique.
μ	Estimator constant governed by the horizon length λ
ν	Defines the size/radius of the sub-covers in the cover (H, ν)
M	Minimisation operator
$\mathcal{M}_p$	Graph of P_p
$\mathcal{M}_p^{[a,b]}$	Graph of P_p defined on the interval $[a,b]$

$\mathcal{N}_p^{[a,b]}(w_2)$	Set of disturbances consistent with w_2 and P_p over the interval $[a, b]$
N	Norm operator - takes the norm of disturbance signals
o	Dimension of the output signal space
O_k	Ongoing switching times up to time k
$O_k(p, \nu(k)(p))$	Ongoing switching times to plants in the neighbourhood $B(p, \nu(k)(p))$ around P_p up to time k
p_*	Parametrisation corresponding to the true plant $P_{p_*} = P$
P_{p_*}	$P_{p_*} = P$, true, physical plant
$\mathcal{P}$	Parametrised set of all plants
$\mathcal{P}_i$	A plant model set
$\mathcal{P}^*$	Powerset of $\mathcal{P}$
$\mathcal{P}^G$	Union of plant sets G possibly maps to
$\mathcal{P}^H$	Union of plant sets H possibly maps to
$\mathcal{P}^U$	Union of plant sets U possibly maps to
Π_x	Projection onto the subspace x
$\Pi_{P//C}$	Map from the disturbances w_0 to the controller signals w_2
$\Pi_{C//P}$	Map from the disturbances w_0 to the plant signals w_1
Φ_j	Extraction operator; extracts signals of length $j \in \mathbb{N}$
q	Switching signal
q_f	'Free', undelayed switching signal
Q_k	Set of switching times (physical and virtual) up to time k
$Q_k(p, \nu(k)(p))$	Switching times to plants in the neighbourhood $B(p, \nu(k)(p))$ around P_p up to time k
R	Union of sub-covers of (H, ν)
$\mathscr{R}_{\sigma,k}v$	Restriction of the signal v to the interval $[k - \sigma, k]$
σ	$\sigma = \max_{p_1,p_2 \in \mathcal{P}^U} \max\{\sigma(p), \sigma(c)\}$
$\sigma(c)$	Required interval length to uniquely determine the initial condition of C_c
$\sigma(p)$	Required interval length to uniquely determine the initial condition of P_p
S	Switching operator
S_f	Undelayed, 'free' switching operator
$\mathscr{T}_k v$	Truncation of the signal v at time $k \in \mathbb{N}$
$\mathbb{T}$	Either $\mathbb{R}, \mathbb{R}^+$ or $\mathbb{Z}, \mathbb{N}$
u_0^p	Input disturbance corresponding to plant model P_p
u_0	Input disturbance corresponding to the true plant P_{p_*}
u_1^p	Plant input corresponding to plant model P_p
u_2	Controller output
U	Plant-generating operator specifying the uncertainty
$\mathcal{U}$	Space of $L_p, l_p,\ 1 \le p \le \infty$ norm bounded input signals
Υ	Disturbance weight, defined by the estimator
$\mathcal{V}$	Space of $L_p, l_p,\ 1 \le p \le \infty$ norm bounded signals

$\mathcal{V}_e$	Space of $L_p, l_p,\ 1 \leq p \leq \infty$ norm bounded signals on finite intervals
w_0	Disturbance signals $w_0 = (u_0, y_0)^\top$
w_1	Plant signals $w_1 = (u_1, y_1)^\top$
w_2	Controller signals, observation $w_2 = (u_2, y_2)^\top$
$\mathcal{W}$	$\mathcal{U} \times \mathcal{Y}$
$\mathcal{W}_e$	$\mathcal{U}_e \times \mathcal{Y}_e$
X	Residual generator (operator), factorises to $X = NE$
$X(w_2)(p)(k)$	$X(w_2)(p)(k) = r_p[k] = \|d_p[k]\|$, residual to P_p at time k
ξ	$\xi = \begin{cases} r & \text{for} \quad 1 \leq r < \infty \\ 1 & \text{for} \quad\quad r = \infty \end{cases}$
y_0^p	Output disturbance corresponding to a plant model P_p
y_0	Output disturbance corresponding to the true plant P_{p*}
y_1^p	Plant output corresponding to a plant model P_p
y_2	Controller input
$\mathcal{Y}$	Space of $L_p, l_p,\ 1 \leq p \leq \infty$ norm bounded output signals

To my parents

Chapter 1

Introduction

When Harold Stephen Black invented the feedback amplifier in 1927 (see Black (1934)) he revolutionised the telecommunications industry and created a prime example which highlights the key aspects of control theory. The problem at the time was that variations in supply voltage and amplifier gain resulted in large variations of the transmission characteristics of the then used feed-forward amplifiers. In contrast, by feeding back the output, he created a device which preserved its transmission characteristic even in the presence of disturbances in the supply voltage and uncertainty of the open-loop gain.

Explicit technological applications of feedback can be found in antiquity. For example feedback control ensured stability of the outrigger canoe prior to 1500 BC (see Abramovitch (2005)) despite the odds of wind, waves and inconsistent load; making safer and further sea travel possible. It also allowed the construction of the water clocks of the ancient Greeks and Arabs around 200 BC (see Mayr (1970)) where the water flowing from a vessel was taken as a measure of time and a feed-back control system ensured the constance of the water level/pressure inside the vessel to improve the accuracy.

Notwithstanding these significant ancient applications, it was the feedback amplifier of Black which provided much of the stimulus for the development of classical feedback control theory, which in turn formed the foundation of robust control theory (H_∞ etc.) in the 1980's. Note that the roots of modern control theory itself, rather than the uncertainty/robustness aspects, can be traced to the works of Maxwell and Routh on the stability of governors (see Maxwell (1868); Routh (1877)).

Abstractly we can describe a feed-back control system as the interconnection between a physical process or plant $P : u_1 \mapsto y_1$ and controller $C : y_2 \mapsto u_2$ as in Figure 1.1, where $(u_1, y_1)^\top$ represents the plant signals, $(u_2, y_2)^\top$ the controller signals, $(u_0, y_0)^\top$ the external disturbances acting on the system, and P and C are operators typically modelled by differential or difference equations.

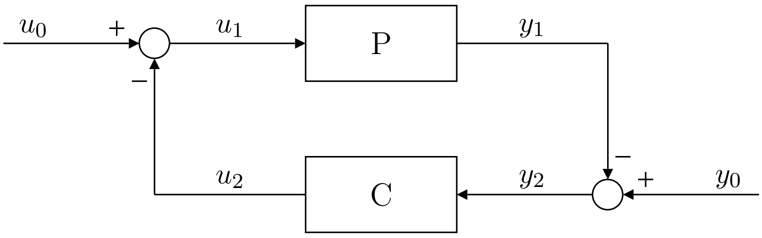

FIGURE 1.1: Closed loop system $[P, C]$

We denote P a plant 'model' since the equations describing P are a simplified representation of the true physical plant denoted by P_1. These equations are usually obtained by analytical or empirical techniques. P therefore only represents an approximation of the physical system P_1.

The required accuracy of the model P, and hence the mismatch between the plant model P and the physical system P_1, is strongly influenced by the nature of the dynamics of P, and the requirements of the control objective. To see this observe that the development of a controller for an air conditioning system within a building does not require a very accurate model P of the physical properties of the room P_1; a coarse model is sufficient to construct a corresponding controller C that keeps the temperature variation within reasonable bounds. However, when positioning the head of a hard disk drive accurately within milliseconds, an accurate model P of the drive's arm P_1 is indispensable to construct a sufficiently good controller C. The knowledge of P therefore determines the achievable performance of the controller C.

To quantify the mismatch of the physical system and the model, or the uncertainty in the physical system, we typically invoke a suitable so-called uncertainty model. In this thesis this is described by introducing an appropriate measure of distance $\delta(P, P_1)$ between plants P and P_1. The uncertainty around a nominal model P is then described by the set of plants P_1 lying within some specified distance $\epsilon > 0$ of P, e.g.
$\Delta_\epsilon = \{P_1 \ : \ \delta(P, P_1) < \epsilon\}$. Throughout this thesis, we take the distance δ to be the gap metric.

The second major factor influencing stability and performance of dynamical systems are external disturbances acting on the system. A stable system has to remain so even in the presence of disturbances (disturbance rejection). For example an airplane guided by an auto pilot shows good disturbance rejection since it maintains its course and altitude even in the presence of disturbances, e.g. wind or air pockets. In contrast, an example of unwanted disturbance amplification is the Tacoma Narrows suspension bridge, where disturbances in the form of strong winds caused resonant oscillations of increasing magnitude in the bridge structure and ultimately led to its destruction.

Observe that in L_2 or l_2 the norm of a signal relates to its energy content:

$$\|x\|_2 = \left(\int_0^\infty |x(t)|^2 dt\right)^{1/2}, \text{ in continuous time } (L_2)$$
$$\|x\|_2 = \left(\sum_{i=0}^\infty |x_i|^2\right)^{1/2}, \text{ in discrete time } (l_2).$$

A good test for stability would then be to check whether a system fed with signals of finite energy responds with signals of finite energy; or in other words, that the amplification or gain from $\|(u_0, y_0)^\top\|_2$ to $\|(u_2, y_2)^\top\|_2$ is finite (we will later see that this also implies that the gain from $\|(u_0, y_0)^\top\|_2$ to $\|(u_1, y_1)^\top\|_2$ is finite).

A system is therefore said to be gain stable if the operator $\Pi_{P//C} : \begin{pmatrix} u_0 \\ y_0 \end{pmatrix} \mapsto \begin{pmatrix} u_2 \\ y_2 \end{pmatrix}$ is bounded, i.e. if

$$\gamma = \|\Pi_{P//C}\|_2 = \sup_{(u_0,y_0)^\top \neq 0} \frac{\|(u_2, y_2)^\top\|_2}{\|(u_0, y_0)^\top\|_2} < \infty.$$

The quantity γ thus denotes the gain from the external disturbances to the internal signals and hence if a closed-loop system is gain stable, then γ is a sensible measure of nominal performance.

Under a technical assumption of well-posedness (see Chapter 2) we can now inter-relate disturbances, uncertainty, stability, robustness and performance in the following way:

Theorem 1.1. *Let P, P_1, C be linear and time invariant. If the closed loop $[P, C]$ is gain stable and*

$$\delta(P, P_1) < \frac{1}{\|\Pi_{P//C}\|_2} = b_{P,C}$$

then the closed loop $[P_1, C]$ is also gain stable.

Proof The proof can be found in Georgiou and Smith (1990) which is based on Zames and El-Sakkary (1980). □

Therefore, if a controller C is able to stabilise a plant P it will also stabilise all plants in the neighbourhood $\delta(P, P_1) < b_{P,C}$ where $b_{P,C}$ is the inverse of the maximum gain from the external disturbances to the internal signals and is denoted the robust stability margin. Hence $b_{P,C}$ is both a measure of nominal performance and robust stability. This relationship will be of major importance to us since by analysis of the nominal plant model P in closed loop with controller C we can then show by the above theorem that if the closed loop $[P, C]$ is gain stable and the mismatch between the plant model P and the physical plant P_1 is smaller then the robust stability margin $b_{P,C}$, then the closed loop $[P_1, C]$ will also be gain stable.

It can be shown that the given robust stability framework also extends to the non-linear domain and to general signal spaces (see Georgiou and Smith (1997)). However in a general non-linear setting the worst case signal amplification from $(u_0, y_0)^\top$ to $(u_2, y_2)^\top$ can vary with magnitude of the signal $(u_0, y_0)^\top$. Hence Georgiou and Smith (1997) also establish robust stability results where the gain is measured by a so-called gain function:

$$\gamma(r) = \sup\{\|(u_2, y_2)^\top\| : \|(u_0, y_0)^\top\| \leq r\},\ r \geq 0.$$

The gain function $\gamma(r)$ measures the maximum size of the internal signals, given a disturbance of size smaller then $r \geq 0$. This gives us a comprehensive set of tools to analyse the robustness properties of (non-linear) closed-loop systems.

The remainder of the introduction has the purpose of motivating the class of algorithms considered in this thesis. We will first show that a single, fixed, linear time invariant (LTI) controller C is generally insufficient to control a plant P if the uncertainty in P is large (i.e. if $\delta(P, P_1)$ is large, where P represents the plant model and P_1 the physical system). This arises when P lies in some known, but potentially large set Δ, for example if Δ describes a parametrically uncertain system with a large parameter variation. One solution to such a problem is to make the controller adaptive. We will discuss various (classical) adaptive algorithms and their limitations in terms of robust stability. Such robust stability considerations will then motivate the class of Multiple Model Switched Adaptive Control (MMSAC) algorithms and Estimation-based MMSAC (EMMSAC) algorithms - the latter will be the focus of this thesis.

1 Insufficiency of LTI control theory

Although LTI control theory gives good design methodologies for LTI control problems where the uncertainties in the system are small, there are many applications where it cannot give sufficient performance and robustness guarantees or even fails to give them at all. Here we detail two such scenarios:

Conservativeness: Consider a plant P given by the transfer function

$$P : u_1 \mapsto y_1 : y_1 = \frac{1}{s-a} u_1.$$

Let $a > 0$ be a fixed but uncertain parameter, for example an unknown mass, and consider a proportional controller

$$C : y_2 \mapsto u_2 : u_2 = -l y_2, l > 0$$

to be in a closed loop interconnection with P as in Figure 1.1. The resulting closed-loop transfer function from y_0 to u_1 is given by

$$S = \frac{u_1}{y_0} = -\frac{C}{1-CP} = \frac{l(s-a)}{s+(l-a)}$$

where u_1, y_0 are the Laplace transforms of the corresponding time domain signals. All poles must reside in the left complex half plane in order to ensure bounded input - bounded output (BIBO) stability of S. This requires us to choose the controller gain l larger than a. Consequently for a large uncertainty in a, i.e. if all that is known is that $|a| \le a_{max}$ where a_{max} is large, we will have to choose a large l, i.e. $l > a_{max}$, to ensure stability of S.

We can now establish a lower bound for the closed loop gain (with Theorem 2.3) in the following way:

$$\|\Pi_{P//C}\|_2 = b_{P,C}^{-1} = \sup_{(u_0,y_0)^\top \neq 0} \frac{\|(u_2,y_2)^\top\|_2}{\|(u_0,y_0)^\top\|_2} \ge \sup_{y_0\neq 0} \frac{\|u_1\|_2}{\|y_0\|_2} = \|S\|_\infty = \sup_{w\in\mathbb{R}} |S(jw)|$$

where

$$|S(j\omega)|^2 = \frac{|j\omega l - la|^2}{|j\omega + (l-a)|^2} = \frac{\omega^2 l^2 + l^2 a^2}{\omega^2 + (l-a)^2}.$$

A simple calculation shows that $|S(j\omega)|^2$ reaches its maximum at $\omega = \infty$, where

$$\lim_{\omega\to\infty} |S(j\omega)|^2 = l^2 > a_{max}^2.$$

This shows that the maximum gain from the external disturbances to the plant signals scales with $l > a_{max}$; and therefore its inverse, the robust stability margin $b_{P,C}$, shrinks to zero as a_{max} becomes large — as depicted in Figure 1.2.

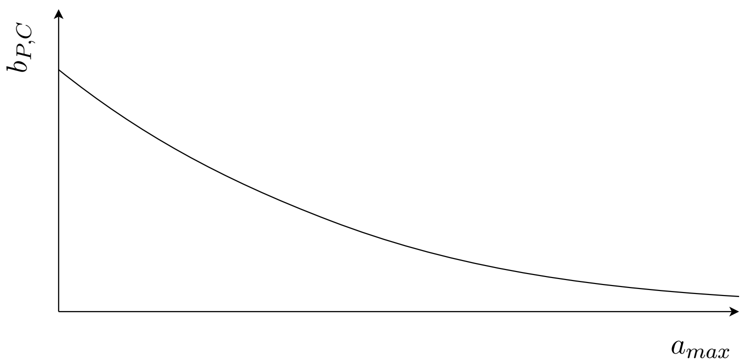

FIGURE 1.2: Stability margin $b_{P,C}$ vs. uncertainty a_{max}

Controllers with this property are regarded as conservative:

Definition 1.2. *A controller C is said to be conservative if the closed loop performance degrades with an increasingly large uncertainty in $P = P_{p_*}$. (See Figure 1.3.)*

Definition 1.3. *A controller C is said to be universal if it maintains a constant level of performance invariant to the uncertainty in $P = P_{p_*}$. (See Figure 1.3.)*

One way of showing that a controller is non-conservative is therefore to show that it is universal. It can be shown that all LTI controllers and also non-linear memoryless controllers are conservative (see French (2008)) with respect to $b_{P,C}$. In this thesis we will present necessarily non-linear and dynamic control designs that are universal.

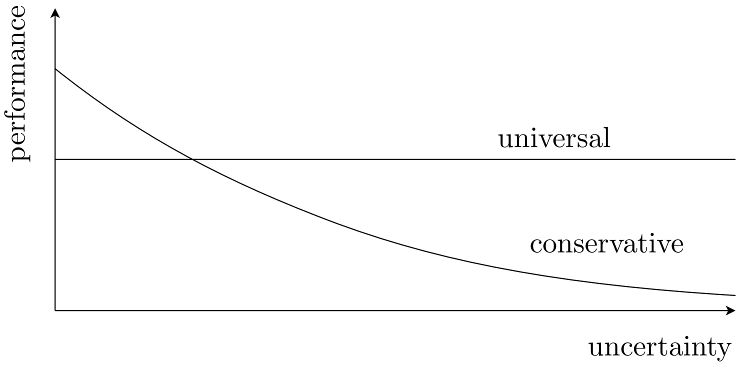

FIGURE 1.3: Closed loop performance for conservative and universal controllers under increasing uncertainty

Simultaneous stabilisation: In practise we often need to consider two distinct plant models P_1 and P_2 and ask the question whether a given controller C can ensure closed-loop stability for both of them, i.e. ensure that $[P_1, C]$ and $[P_2, C]$ are stable. This arises if we seek to control systems with different dynamics with the same controller or the control of systems that can abruptly change their dynamic behaviour over time. For example in fault tolerant control we are interested if a controller, designed to control a nominal system P_1, also controls a faulty system P_2 with a different dynamic behaviour.

It is well known that the problem of simultaneous stabilisation is related to the problem of strong stabilisation. A plant P is said to be strongly stabilisable if a controller C can be found such that the closed loop $[P, C]$ is stable and C is itself stable.

In the case where plants and controllers are LTI, Youla et al. (1974) and Saeks and Murray (1982) showed that two plants, given by the real-rational transfer functions P_1 and P_2 are simultaneously stabilisable if and only if

$$P = \frac{N_2 M_1 - N_1 M_2}{N_2 X_1 + M_2 Y_1}$$

is strongly stabilisable where M_1, M_2, N_1, N_2 are coprime factors of P_1 and P_2. Furthermore we have from Saeks and Murray (1982) and Vidyasagar (1985) that P is strongly stabilisable if and only if it has an even number of real poles between every pair of real zeros in Re $s \geq 0$.

We now consider an explicit example: Consider the two real rational plants

$$P_1 = \frac{1}{s},\ P_2 = -\frac{1}{s}.$$

We now claim that no LTI controller can simultaneously stabilise P_1 and P_2.

We express P_1 and P_2 in a coprime factor form:

$$P_1 = \frac{M_1}{N_1},\ P_2 = \frac{M_2}{N_2},$$

where $M_i, N_i, X_i, Y_i,\ i \in \{1, 2\}$ are given by

$$\begin{aligned} M_1 &= \frac{1}{s+1},\ N_1 = \frac{s}{s+1},\ X_1 = 1,\ Y_1 = 1 \\ M_2 &= -\frac{1}{s+1},\ N_2 = \frac{s}{s+1},\ X_2 = 1,\ Y_2 = -1 \end{aligned}$$

and satisfy the required Bezout identities

$$N_i X_i + M_i Y_i = 1,\ i \in \{1, 2\}.$$

Hence P_1 and P_2 are simultaneous stabilisable if and only if

$$P = \frac{N_2 M_1 - N_1 M_2}{N_2 X_1 + M_2 Y_1} = \frac{2s(s+1)}{(s+1)^2(s-1)} = \frac{2s}{(s+1)(s-1)}$$

is strongly stabilisable.

Since P has zeros at $s = 0$ and $s = \infty$ and only one intermediate pole ($s = 1$), P is not strongly stabilisable and therefore P_1 and P_2 are not simultaneous stabilisable by a linear controller. The non-linear control designs considered in this thesis are designed to handle such scenarios.

2 Adaptive control

We have shown in the previous section that LTI controllers have difficulties in some situations. In particular they do not pose a solution to the simultaneous stabilisation problem and they are conservative.

These limitations motivate the field of adaptive control. The basic idea behind adaptive control is that a learning component in the controller gathers information from the on-line observation of closed loop signals of an uncertain physical system P in order to learn about the uncertainty. This information is then utilised to generate control signals promising better performance then a fixed, non-learning controller. Whilst adaptive control has a long history, and whilst such controllers have the potential to be non-conservative and handle non-simultaneously stabilisable plants there is currently a relatively poor understanding of their robustness properties.

The substantive body of research on ‘robust adaptive control’ confines uncertainty models to additive or multiplicative classes (see Ioannou and Sun (1996) for a comprehensive review of these approaches). However more recent work French (2008) and French et al. (2006) has established robust stability margins for classical schemes in the context of gap metric uncertainty models — this thesis builds on the approaches therein.

We start with the discussion of continuously tuned adaptive controllers and then turn to multiple model type algorithms where the concept of switching is introduced along the way.

3 Continuous adaptive control

3.1 Nominal stability

Assume that no disturbances are acting on the system for now, i.e. $(u_0, y_0)^\top = 0$. Consider the plant

$$P : u_1 \mapsto y_1 : y_1 = \frac{1}{s-a} u_1$$

equally defined by the corresponding differential equation

$$P : \dot{y}_1 = a y_1 + u_1, \; y_1(-t) = 0, \; \forall t \in \mathbb{R} \tag{1.1}$$

where a is an uncertain parameter. A typical non-switched adaptive control implementation is given by the equations

$$C : y_2 \mapsto u_2 : \begin{cases} u_2 = -y_2(1+\hat{a}) \\ \dot{\hat{a}} = y_2^2 \\ \hat{a}(0) = 0. \end{cases} \tag{1.2}$$

The time-varying parameter $\hat{a}$ is thought of the estimate of the parameter a, since in the case where $a = \hat{a}$ we have by equations (1.1),(1.2) and with $(u_0, y_0)^\top = 0 = (u_1, y_1)^\top + (u_2, y_2)^\top$ from Figure 1.1 that

$$\dot{y}_2 = a y_2 + u_2 = a y_2 - \hat{a} y_2 - y_2 = -y_2$$

which is asymptotically stable. However for the case where $a \neq \hat{a}$ we have to consider the mismatch $\theta = a - \hat{a}$.

Consider the Lyapunov function

$$V(y_2, \theta) = \frac{1}{2} y_2^2 + \frac{1}{2} \theta^2. \tag{1.3}$$

By equations (1.1)–(1.3) and since $\dot{\theta} = -\dot{\hat{a}} = -y_2^2$, we obtain:

$$\dot{V}(y,\theta) = y_2\dot{y}_2 + \theta\dot{\theta} = y_2(ay_2+u_2) - \theta y_2^2 = y_2(ay_2 - \hat{a}y_2 - y_2) - \theta y_2^2 = \theta y_2^2 - y_2^2 - \theta y_2^2 = -y_2^2.$$

Since $\dot{V}$ is negative semidefinite, by La-Salle's theorem, we have $y_2 \to 0$ for $t \to \infty$. It is then straightforward to verify that the signals $\hat{a}, u_1, y_1, u_2, y_2$ are all bounded. Also observe these properties hold for any value of a. A controller with this property is denoted a universal controller and it is therefore non-conservative.

Designs of this type share the deficiencies of LTI controllers in simultaneously stabilising $\frac{1}{s}, -\frac{1}{s}$. In fact it has been shown in French (2004) that no smooth controller can simultaneously stabilise $\frac{1}{s}, -\frac{1}{s}$.

A continuously tuned controller that can cope with such plants is the Nussbaum universal controller (Nussbaum (1983)) in equation (1.4)

$$C : y_2 \mapsto u_2 : \begin{cases} u_2 = y_2\hat{a}^2 \cos\hat{a} \\ \dot{\hat{a}} = y_2^2 \\ \hat{a}(0) = 0, \end{cases} \tag{1.4}$$

which stabilises any plant $P \in \Delta$ where $\Delta = \{\frac{\pm 1}{s-a} \ : \ a \in \mathbb{R}\}$. This is accomplished by introducing the oscillatory function $\cos\hat{a}$. Hence if $y_2 \neq 0, \dot{\hat{a}} > 0$ then $\hat{a}$ is increasing and u_2 will oscillate in sign. Therefore the controller will 'try out' negative and positive signs. Now, as the generated control output u_2 manages to stabilise the plant, i.e. y_2 becomes small, then $\dot{\hat{a}}$ becomes small and $\hat{a}$ settles to a constant value. This means that the oscillation will slow down and the sign of $\cos\hat{a}$ will remain constant over increasingly long intervals, essentially giving the individual controllers an increasing amount of time to stabilise (and destabilise) the system. The algorithm will then eventually settle on the correct sign since the oscillation stops if $y_2 = 0$. This result was later generalised to systems with arbitrary relative degree in continuous time by Mudgett and Morse (1985) and discrete-time by Lee and Narendra (1986).

3.2 Instabilities due to lack of robustness

One major problem of the above approaches is that the effects of (input and output) disturbances and unmodeled dynamics on stability and robustness were neglected due to the initial belief that analogously to the LTI case the control system would tolerate them if sufficiently small. Unfortunately this belief was proven wrong by Rohrs et al. (1985) which showed that virtually all continuously tuning implementations at that time could in fact become unstable in the presence of seemingly harmless unmodeled dynamics and arbitrarily small disturbances.

For example assume the nominal plant

$$P = \frac{2}{s+1}$$

is perturbed multiplicatively by

$$\Phi = \frac{229}{s^2 + 30s + 229}$$

to give

$$P_1 = P\Phi = \frac{2}{s+1}\frac{229}{s^2 + 30s + 229}.$$

As before, let the adaptive controller C be given by

$$C : \begin{cases} u_2 = -y_2(1+\hat{a}) \\ \dot{\hat{a}} = y_2^2 \\ \hat{a}(0) = 0 \end{cases}$$

and assume it to be in closed-loop configuration $[P_1, C]$ with the perturbed plant P_1 where the input and output disturbances are constant and given by $u_0 = 0,\ y_0 = 3$. This setup is commonly known as Rohrs counter example.

Observe that the open-loop transfer function P_1 is stable and the perturbation Φ has a unity DC gain and two well damped complex poles distinct from P at -15Rad/s. Such a perturbation would in the LTI case not be considered as a problematic unmodeled dynamic.

However, as the simulation in Figure 1.4 shows, the given adaptive control algorithm becomes unstable.

This is due to the following mechanism of instability: 1. $\hat{a}$ diverges as time increases (this is known as parameter drift) and 2. the closed loop becomes unstable for high closed-loop gains.

To see 1. assume that $\hat{a}$ remains bounded, i.e. $\hat{a}(t) < A < \infty,\ \forall t > 0$. Since $\dot{\hat{a}} = y_2^2$ we have that

$$\hat{a}(t) = \int_0^t y_2^2 dt = \|y_2|_{[0,t]}\|_{L_2}^2$$

This implies that $y_2 \in L_2$ as $\|y_2\|_{L_2} \leq \sqrt{A}$. From $y_0 = y_1 + y_2$ we therefore have that $(y_0 - y_1)$ is in L_2. However, since $y_0 = 3 \notin L_2$ it follows that $y_1 \notin L_2$.

Now observe that since P_1 is stable ($u_1 \in L_2 \Rightarrow P_1 u_1 = y_1 \in L_2$) it follows that if $y_1 \notin L_2$ then $u_1 \notin L_2$. However, since $u_1 = -u_2 = -y_2(1+\hat{a})$, $A < \infty$ and $y_2 \in L_2$, it

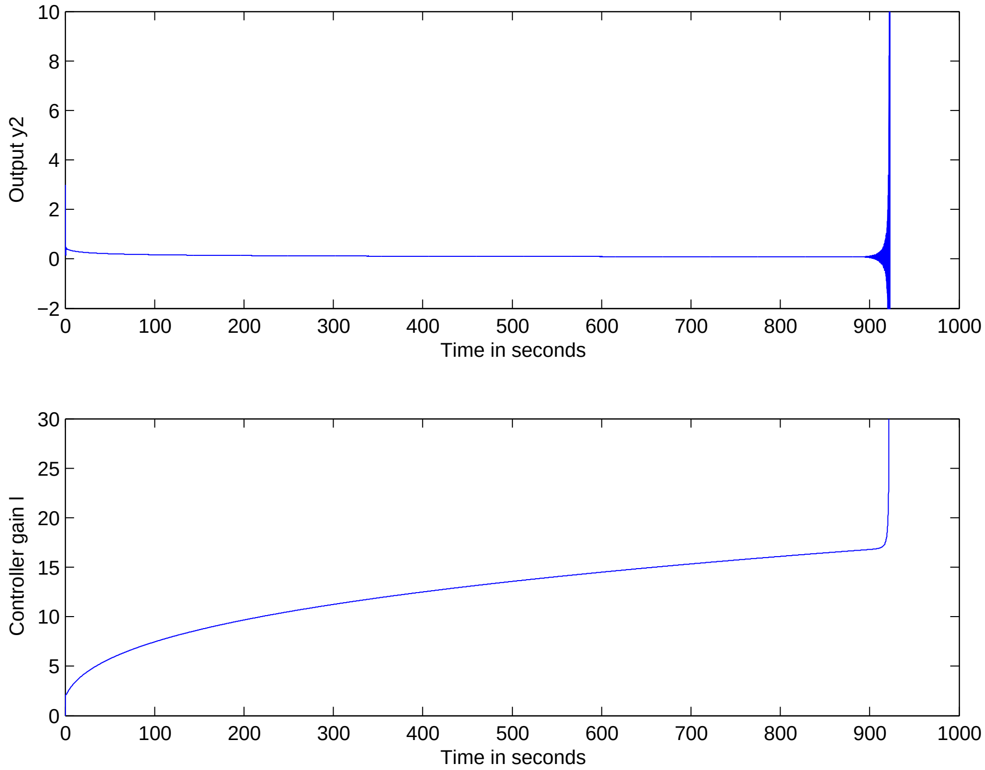

FIGURE 1.4: Classical adaptive controller in the presence of a minor perturbation and a constant output disturbance ($l = \hat{a}$)

follows that:

$$\begin{aligned}\|u_1\|_{L_2}^2 &= \int_0^\infty u_1^2 dt = \int_0^\infty y_2^2(1+\hat{a})^2 dt \\ &\leq \int_0^\infty y_2^2 2(1+\hat{a}^2)dt \\ &\leq 2(1+A^2)\int_0^\infty y_2^2 dt \\ &< \infty.\end{aligned}$$

Hence $u_1 \in L_2$, which is a contradiction and therefore $\hat{a} \to \infty$.

To see 2. consider the root-locus plot of P_1 in Figure 3.2 which displays the loci of the closed-loop poles in the complex plane in respect to the feed-back gain $\hat{a} > 0$. Since the open-loop equation P_1 has no zeros and three poles, its closed-loop poles diverge to infinity separated by a 120 degree angle for increasing feed-back gains. Since we have shown above that $\hat{a}$ (and therefore the feed-back gain) grows over all bounds, i.e. $\hat{a}(t) \to \infty$ as $t \to \infty$, the closed-loop will eventually become unstable.

This insight now allows an intuitive interpretation of the plot in Figure 1.4. Up to time $t = 900$ the closed-loop remains stable since $\hat{a}$ is still within reasonable bounds.

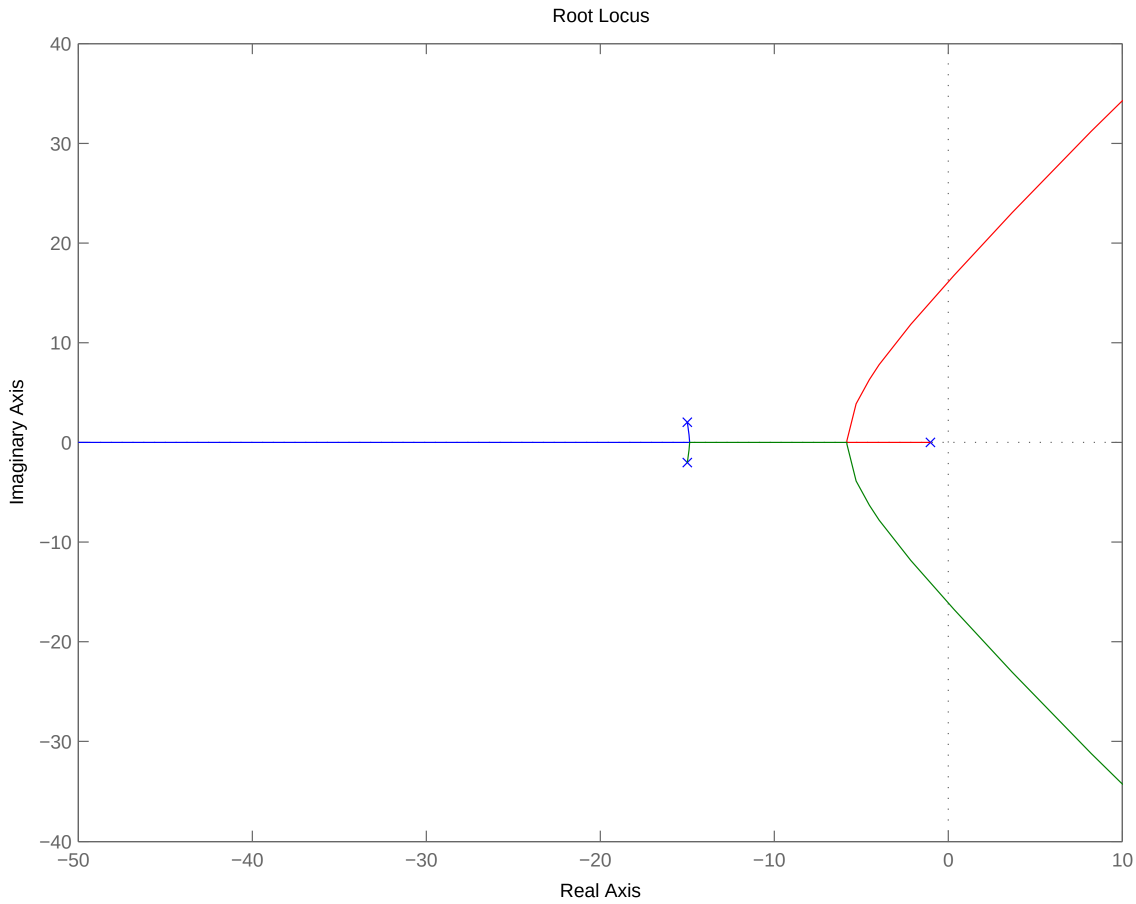

FIGURE 1.5: Root locus of the open loop transfer function P

However, shortly after $t = 900$, the growing feed-back gain $\hat{a} > 0$ forces the roots of the closed-loop to cross over to the complex right half plane and the closed-loop system becomes unstable, which leads to a rapid growth in size of the involved signals.

See also French et al. (2006), where a similar mechanism of instability is rigorously proved for a multiplicative all pass perturbation $\Phi(s) = \frac{M-s}{M+s}$, $M >> 1$. Related results showing parameter drift for nominal plants can be found e.g. in Georgiou and Smith (1997, 2001).

3.3 Robust adaptive control

Note that in the previous example the maximum gain from the external disturbances to the internal signals is in this (non-linear) case not a simple gain of a transfer function but depends on the (non-linear) worst case behaviour of the system. For the Rohrs counter example we can simply observe, at least in an L_∞ setting, that this gain is indeed infinite since we managed to produce with finite disturbances $u_0 = 0, y_0 = 3$ an unbounded y_2. Hence the closed loop system is not gain stable.

At the time (in the 1980s) the control community was increasingly aware of a systematic problem and that neglecting robustness was no longer viable. However, there was no analytical framework available — like the one of Georgiou and Smith (1997) — to study the robustness properties of the algorithms in a systematic kind of way.

What followed was the investigation into so-called robust adaptive control algorithms where robustness referred more to avoiding the phenomena of parameter drift than developing a principled robustness theory. Since the unbounded increase of $\hat{a}$ was identified as the problematic part that ultimately caused the overall instability, typical robust adaptive control approaches are aimed at keeping $\hat{a}$ within reasonable bounds. In particular, modifications to the algorithm were proposed such as the inclusion of dead-zone, σ modifications, or projection operators to ensure robustness in the presence of disturbances and unmodeled dynamics. Assumptions that the reference signal is sufficiently rich to ensure parameter convergence or the injection of artificial probing signals had a similar purpose. A summary of robust adaptive control ideas and corresponding robustness proofs for plants can be found in standard text books on the topic, e.g. Ioannou and Sun (1996), Narendra and Annaswamy (1989) and Sastry and Bodson (1989). These approaches apply to plants perturbed by multiplicative and additive uncertainty only. For a recent example of this style of analysis see Ikhouane and Krstic (1998), where the authors are able to show for a continuously tuned adaptive controller that all closed loop signals are bounded and the tracking error is proportional to the size of the disturbance; however only allow multiplicative uncertainties and output disturbances.

We will now discuss some of these modifications, however note that since most robust adaptive controllers provide for zero output disturbances ($y_0 = 0$) an infinite robustness margin for the parametric uncertainty, a measurement of performance in terms of such margins is not meaningful. We therefore consider alternative, non-singular, measures of performance that are related to $b_{P,C}$. For example in French (2002) the performance is evaluated by an integral costs functional that penalises the state and the control effort, where in Sanei and French (2006) the costs functional is a sum of L_∞ measures of the state trajectory, the control signal and its derivative.

Dead zones: The idea behind dead zones is to monitor the measurable signals and to disable the parameter update law for $\hat{a}$ if they enter the dead-zone region Ω, where a poor signal to noise ratio could destabilise the system. The region Ω naturally depends on an a priori knowledge of the size of the disturbances and introduces some conservativeness to the design.

The modified controller from equation (1.2) then reads

$$C_{dead}: \begin{cases} u_2 = -y_2(1+\hat{a}) \\ \dot{\hat{a}} = \begin{cases} y_2^2 & \text{if} \quad (u_2, y_2)^\top \notin \Omega, \\ 0 & \text{if} \quad (u_2, y_2)^\top \in \Omega \end{cases} \\ \hat{a}(0) = 0 \end{cases} .$$

Projections: The aim of the projection modification is to directly bound the size of the tuning parameter $\hat{a}$ and is given by

$$C_{proj}: \begin{cases} u_2 = -y_2(1+\hat{a}) \\ \dot{\hat{a}} = \begin{cases} y_2^2 & \text{if} \quad \hat{a} < \hat{a}_{max}, \\ 0 & \text{if} \quad \hat{a} \geq \hat{a}_{max} \end{cases} \\ \hat{a}(0) = 0 \end{cases}$$

where we observe that a priori knowledge of an upper bound $\hat{a}_{max}$ of $\hat{a}$ is required to construct C_{proj}.

See Sanei and French (2006) for a direct performance comparison between the dead zone and the projection modification, where the authors show that if the bound on the uncertainty is sufficiently conservative then a dead-zone modified controller outperforms its projection modified counterpart. The converse holds when the a priori information on the disturbance level is sufficiently conservative.

Sigma modification: A further possibility to prevent $\hat{a}$ from drifting to infinity is to add an additional term to the parameter update law and penalise large values of $\hat{a}$. With σ being a small, positive constant we would then have a controller

$$C_\sigma: \begin{cases} u_2 = -y_2(1+\hat{a}) \\ \dot{\hat{a}} = y_2^2 - \sigma\hat{a} \\ \hat{a}(0) = 0. \end{cases}$$

If however the true parameter a is large then $\hat{a}$ is large hence via the parameter update equation $\hat{a}$ will be forced away from its equilibrium point $\hat{a}_{eq}$. The introduction of an offset, i.e. $\dot{\hat{a}} = y_2^2 - \sigma(\hat{a}_{eq} - \hat{a})$, would solve this problem however implies a priori knowledge about $\hat{a}_{eq}$ which, if it exists, would question the use of an adaptive controller in the first place. Also uncertainties in the knowledge of $\hat{a}_{eq}$ would lead to conservativeness.

These and other modifications to standard adaptive algorithms all follow the same basic principle: to suppress the parameter drift of $\hat{a}$; where unfortunately a certain amount of conservativeness is introduced along the way. Although initial robustness results for additive and multiplicative uncertainty in terms of Lyapunov stability theory do exist, there remains the lack of a coherent theory capturing the robustness properties of the algorithms in the presence of general disturbances and fully unstructured uncertainty.

Due to this lack of theory and the often unpredictable behaviour of the algorithm, the community's interest in adaptive control cooled down noticeably over the years. Additionally, reports of failed practical tests with sometimes devastating results — like the flight of the X-15 (see Staff of the Flight Research Center (1971)), which disintegrated in mid-air due to controller induced high-gain instability — left a permanent mark on the adaptive approach in general. Note that the investigation of continuously tuned adaptive controllers continues up until today and various successful applications have been reported, e.g. see Guan and Pan (2008) for the control of an uncertain electro-hydraulic actuator or Hung et al. (2008) for the control of robot manipulators with non-linearly parametrised uncertainties, to cite only two recent ones. However the discussed difficulties remain.

3.4 Robust stability theory

When 27 experts were asked in survey about "[...] major open problems in control theory" (Blondel et al. (1995)), one of them gave a particularly revealing answer for the area of adaptive control:

> "There is not as yet an adequate robust adaptive control theory; this may be due to the fact that there is a complete mismatch between the current mathematical formulations of robust and adaptive control." (P. E. Caines)

More recently, attempts were made to overcome this problem and re-investigate the robustness properties of adaptive control algorithms from the perspective of robust control theory. French (2008) analyses the robustness properties of a continuously tuned adaptive controller in the framework of Georgiou and Smith (1997) for the case of fully unstructured uncertainties (in the gap metric) and the disturbance model as depicted in Figure 1.1. The author then shows that there exists a class of non-conservative, continuously tuned adaptive controllers that robustly stabilise finite-dimensional, minimum phase plants P perturbed to P_1 where the gap distance between P and P_1, the initial condition and the disturbances $(u_0, y_0)^\top$ are sufficiently small.

Although the given robustness guarantees only allow local disturbances and the established gain function bounds grow rapidly with the bound on the size of the disturbances, these results are important from the perspective of this thesis as:

1. They are the first of their kind that establish comprehensive robustness results in terms of fully unstructured uncertainties in the gap metric for an adaptive algorithm; this inspired the type of robustness analysis conducted in this thesis.
2. They provide insight and motivation for the non-conservative extensions of the algorithm considered in this thesis (in Chapter 6).

3. They demonstrate that it is possible to achieve a robust stability margin, even when the closed loop gain is infinite, and that this scenario is typical in adaptive control.

4. The author establishes gain (function) bounds and robustness margins which are compatible with the disturbance model in Figure 2.1. Therefore the two very different approaches of classical adaptive control and the multiple model switching method considered in this thesis become comparable in terms of their robustness and performance properties.

French et al. (2006) then extend on work from French (2008), considering more standard robust adaptive control designs, and revisit a specific example from Georgiou and Smith (1997) (Example 9), where a plant $P = \frac{1}{s-a}$ is perturbed by the all pass factor $\frac{M-s}{M+s},\ M > 0$ (i.e. in series connection). The robustness result shows that the closed loop system is stable if the gap distance between the perturbed and nominal plant as well as the L_2 disturbances are sufficiently small. Various mechanism of instability are then illustrated. In particular they show L_2 instability of the closed loop system for large initial conditions or large L_2 disturbances. Finally they show that L_∞ disturbances imply that the internal signals do not remain in L_∞ hence the system is considered L_∞ unstable.

4 Multiple Model Adaptive Control

Multiple model type algorithms represent an alternative to the continuously tuned algorithms discussed in the previous section. The name refers to the fact that control is performed on the basis of having a number of (plant) hypotheses, represented in a so-called plant model set, rather than working directly with a parametrised model.

Every multiple model algorithm incorporates three basic building blocks:

1. Plant model set:
   For example for the plant
   $$P_p = \frac{1}{s-p}$$
   with uncertain parameter $1 \leq p \leq 10$. We might choose the plant model set to be
   $$\left\{ \frac{1}{s-1}, \frac{1}{s-2}, \cdots, \frac{1}{s-10} \right\}.$$
   To represent such sets efficiently we let $\mathcal{P}$ denote a parametrisation set, e.g. coefficients of transfer function or state space matrices (A, B, C, D), corresponding to a model. In the case of our example we would write $\mathcal{P}_i = \{1, 2, \cdots, 10\} \subset \mathcal{P}$ where the models are given by $P_p,\ p \in \mathcal{P}_i$.

2. Controller set:
The controller set is related to the plant model set via a controller design procedure $K : \mathcal{P} \to \mathcal{C}$ where $\mathcal{C}$ is the parametrised set of all controllers. We usually require that each 'atomic' plant-controller pair is closed loop stable, i.e. $[P_p, C_{K(p)}]$ is gain stable.

3. Performance function:
The performance function returns a performance signal d_p and has the purpose to assess how valid each plant model $p \in \mathcal{P}$ is, where by convention, a smaller performance signal indicates higher validity.

Typically these building blocks are then interconnected in the following way. The performance signal d_p is evaluated for a (finite) subset $\mathcal{P}_i \subset \mathcal{P}$ denoted the plant model set. Then the controller for which the corresponding plant model $p \in \mathcal{P}_i$ has the smallest performance signal is switched into closed loop (or multiple controllers are implemented in parallel and the outputs are weighted according to their corresponding performance signals). Assuming that the true plant P_{p_*} is included in the plant model set, i.e. $p_* \in \mathcal{P}_i$ and the performance signal is minimal for the plant model corresponding to p_* (or minimal for a plant model close to p_*), the implemented controller might indeed stabilise the true plant P_{p_*}. Such a design framework can also be utilised in the time-varying setting, subject to performance signals being evaluated over suitable short moving horizons.

4.1 Gain scheduled control

We will now briefly consider gain scheduled control (see Murray-Smith and Johansen (1997) for an overview) which fits into the multiple model framework in the time-varying setting; although note that the remainder of the thesis will handle the time invariant case, under significant less observation information e.g. no measurement of the 'scheduling variables'.

In the process or aviation industry one often has to deal with dynamical systems depending non-linearly on some key process variables (the so-called scheduling variables). For example the aerodynamic properties of an airplane such as lift, drag etc. are non-linear functions of altitude, speed and other variables which can be directly measured. The dynamical changes are so significant over the whole flight envelope that there is no hope in adequately controlling the system by a single LTI controller.

Typically a non-linear plant $P(a) : u \mapsto y$ is therefore linearised over a finite set of equilibrium points $a_j,\ 1 \leq j \leq i,\ i \in \mathbb{N}$ of representative operating conditions corresponding to a plant model set $\mathcal{P}_i = \{p_1, p_2, \cdots, p_i\},\ i \in \mathbb{N}$. Via the controller design procedure K, to every plant model $P_p,\ p \in \mathcal{P}_i$ a corresponding controller $C_{K(p)}$ is constructed such that the controller pairs $[P_p, C_{K(p)}],\ p \in \mathcal{P}_i$ fulfils certain performance criteria. The scheduling variables are then measured and the atomic controller outputs are interpolated (linearly) between linearisation points, e.g. see Rugh (1991).

A naive implementation of such an approach is depicted in Figure 1.6, where a is the measured scheduling variable and $a_p,\ p \in \mathcal{P}_i$ are the equilibrium points of the plants $P(a)$. The performance function to each P_p is then given by

$$d_p = |a - a_p|,\ p \in \mathcal{P}_i$$

and

$$x = \operatorname*{argmin}_{p \in \mathcal{P}_i} d_p, \quad z = \operatorname*{argmin}_{p \in \mathcal{P}_i \setminus x} d_p,$$

are the two 'closest' models where the weights $w_p,\ p \in \mathcal{P}_i$ are chosen such that u is a linear interpolation of the atomic controller outputs

$$u = \sum_{p \in \mathcal{P}_i} w_p u_p, \quad u_p = C_{K(p)} y, \quad p \in \mathcal{P}_i$$

with

$$w_x = 1 - w_z, \quad w_z = \frac{d_x}{|a_z - a_x|}, \quad a_x \leq a \leq a_z, \quad w_p = 0, \quad p \in \mathcal{P}_i \setminus \{x, z\}.$$

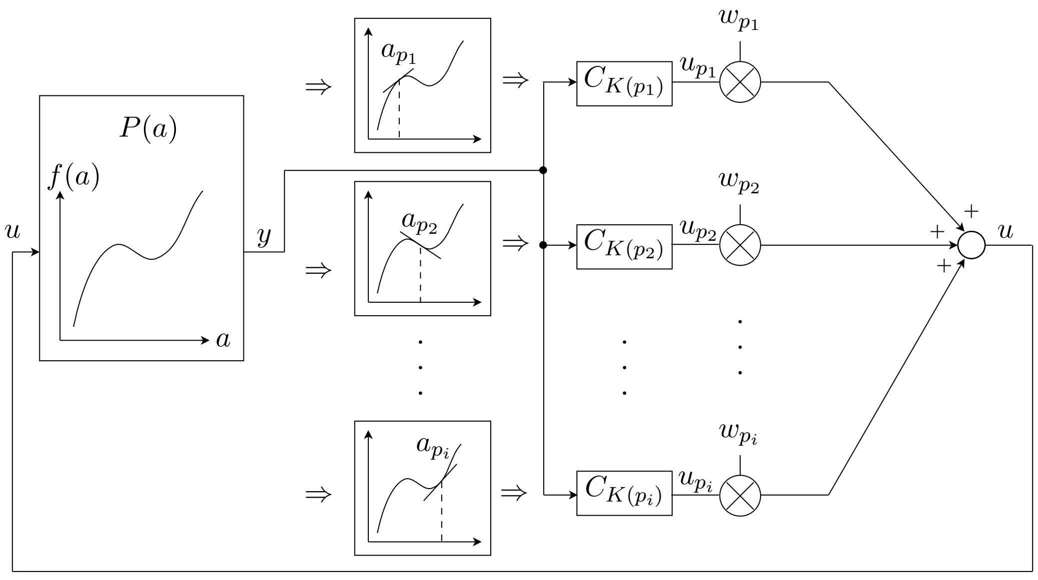

FIGURE 1.6: Gain scheduling algorithm with equilibrium points a_j and corresponding controller designs $C_{K(p_j)}$

Interpolation or blending between models has the purpose of reducing design and implementation complexity, i.e. to reduce the number of linearisations and the number of controllers. It also has the benefit that it effectively smoothens the overall control signal enabling bumpless transfer. However to ensure that that resulting blended signal is ‘sensible’ for control it must be assumed that the dynamical model and controller set satisfies continuity properties in the model parameter space. A ‘hard’ switch can be implemented e.g. by letting

$$w_x = 1, \quad w_p = 0, \quad p \in \mathcal{P}_i \setminus \{x\}.$$

Although this is less common in a gain scheduling context it has more in common with what follows.

In the gain scheduling context, stability is usually shown under the assumption that the rate of change of the involved system variables is slow. Abrupt changes in the process variables, for example by some fault in the sensor or excessive sensor noise, could lead to a fast switching / blending sequence and potentially destabilise the system. For this purpose it is sensible to restrict how fast the algorithm is allowed to blend or switch between controllers. This is enforced by a dynamical requirement that the gain scheduling variables have a slow variation. In the adaptive controllers that follow, switching delays or dwell times are explicitly introduced to prevent instabilities. For example see Liberzon (2003) for mechanisms of switching induced instability.

A particular and potentially restrictive assumption in gain scheduling is that the direct measurement of the scheduling variables is possible. For the given examples of sensor failure or excessive sensor noise this assumption might be impossible to satisfy. Also in many situations the required scheduling variables can not be measured directly in the first place and further complexity has to be introduced to estimate them. The following approaches utilise performance functions which do not depend on the measurement of scheduling variables but determine the validity of every plant model by comparing its dynamical behaviour to the observable input and output signals of the true plant $(u_2, y_2)^\top$. However note that they have a different scope to gain scheduled control since they are usually designed to control a fixed uncertain LTI plant rather than a plant with a time-varying parametric non-linearity.

4.2 Multiple Model Switched Adaptive Control (MMSAC)

A typical switched multiple model algorithms — as depicted in Figure 1.7 — is composed of two basic parts:

1. The subsystem determining the best plant candidate — switching logic S
2. The subsystem implementing the feedback controller — switching controller C

As Figure 1.7 suggests, the signal interconnecting the two subsystems S and C is to be denoted the switching signal q. To be able to deal with possible time dependent building blocks in the algorithm it is sensible to introduce a time base to the involved signals.

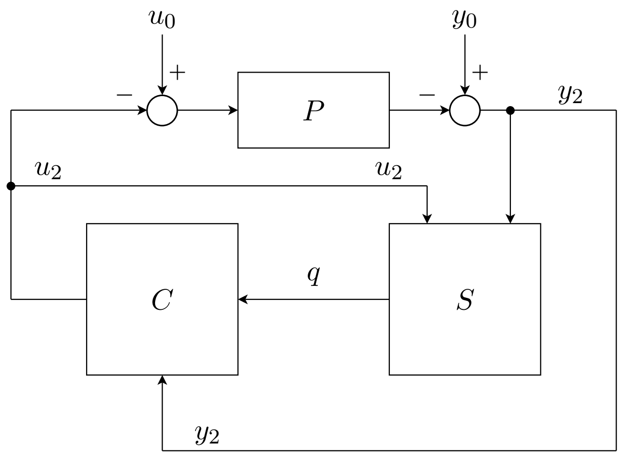

FIGURE 1.7: Multiple model switched system with switching logic S and controller C

In particular we later implement a switching delay to prevent overly fast switching and possible instability and need to quantify how long such a delay should last.

Here q then naturally becomes a piecewise constant function of time — it is discontinuous at switching times, where the switching logic chooses a new controller to be switched into closed loop, and constant everywhere else. Given some set of controllers $\mathcal{C} = \{c_1, c_2, \cdots, c_i\},\ i \in \mathbb{N}$, the switching signal $q : \mathbb{R} \rightarrow \mathcal{C}$ and switching times $\{t_1, t_2, \cdots, t_v, \cdots\},\ v \in \mathbb{N}$, a possible trajectory of q is depicted in Figure 1.8.

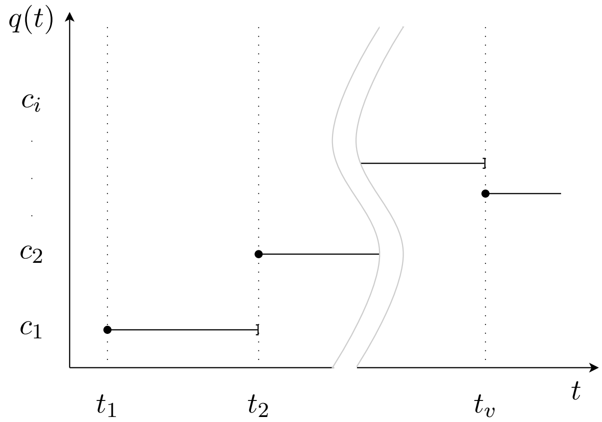

FIGURE 1.8: Switching signal q

Before the potential benefits of MMSAC systems are discussed, a simple example for a switching logic S is given. Let all plant models run in parallel to the true plant and assume that the size of each model's output error (when compared to the true plant's output) represents the performance signal, i.e. let the plant model set $\mathcal{P}_i = \{p_1, p_2, \cdots, p_i\},\ i \in \mathbb{N}$ and $d_p = |e_p|,\ p \in \mathcal{P}_i$ in Figure 1.9 where the true plant is given by P_{p_*}. This rudimentary scheme is completely deterministic and follows the underlying idea that if a model and the true plant are close to each other, their dynamical response should be similar and therefore the output error small. However this implementation proves to be problematic since arbitrarily small differences in the initial conditions lead

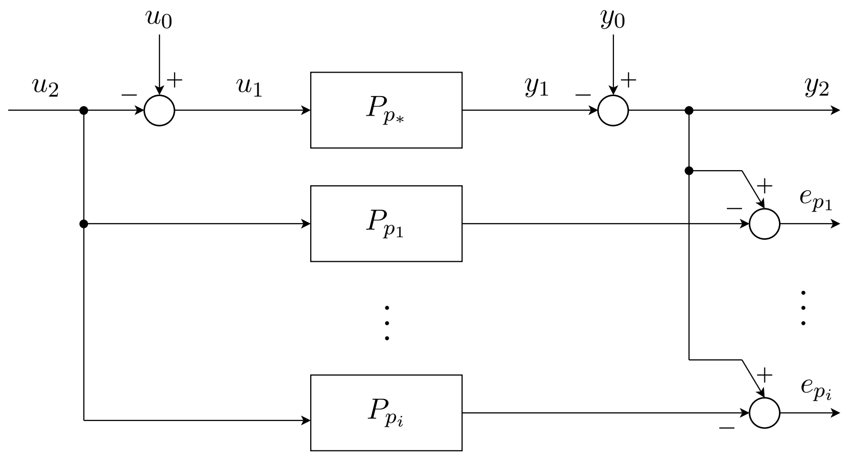

FIGURE 1.9: Free running plant models

to growing output errors even if the model and true plant are identical, which is only eliminated if the plant and model are both stable. One remedy to this problem is to utilise observers (Figure 1.10) instead of free running plants since they are known to

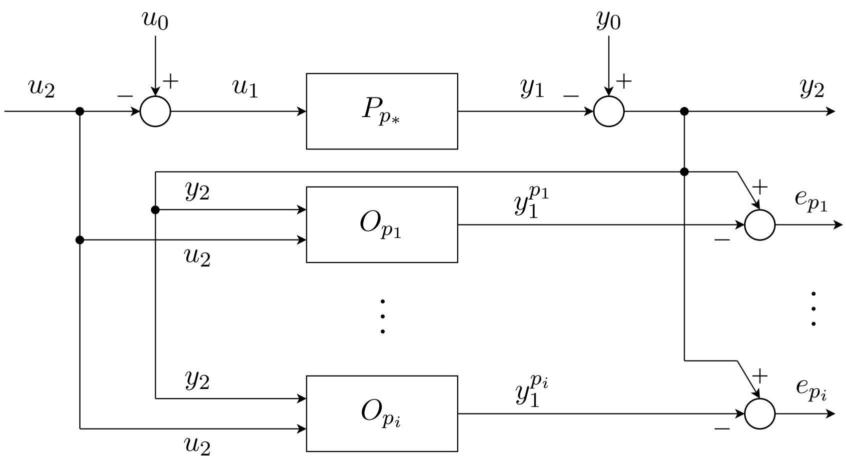

FIGURE 1.10: Observer bank

ensure error convergence even if plant and model are unstable (assuming zero disturbances). The 'free' switching signal q_f is then determined from $e_p,\ p \in \mathcal{P}_i$ for example via

$$q_f(t) = \underset{p \in \mathcal{P}_i}{\operatorname{argmin}} \int_0^t |e_p(\tau)|^2 d\tau.$$

The signal q_f is denoted free since no delay is involved in its construction. We will later (in Chapter 4) introduce a switching delay (operator) $D : q_f \mapsto q$ that delays the signal q_f to q to prevent instability effects due to overly fast switching.

Before we introduce an important algorithm adopting the idea, i.e. to utilise observers for performance evaluation, we want to emphasise that even at this basic level of discussion of MMSAC systems it can be seen that they have the potential to have a number of desirable properties, i.e. there is no conceptual reason why they should not have these properties.

- Design freedom to choose the atomic controllers:

  In principle, MMSAC allows the use of controllers built from entirely standard (off the shelf) design procedures unlike many other adaptive algorithms where the control design is constrained by the specifics of the algorithm itself. The only restriction is that for every plant in the plant (model) set $P_p,\ p \in \mathcal{P}$ the corresponding controller $C_{K(p)}$, constructed by the design procedure K, is stabilising, i.e. $[P_p, C_{K(p)}]$ is gain stable. This also allows the easy assimilation of already existing control designs into the given structure.

- Allows for non-convex parameter sets:

  In continuously tuned adaptive algorithms difficulties arise if the process model is parametrised over non-convex sets, as during the tuning of the parameter the algorithm can enter regions of undesired parametrisation. The MMSAC algorithms naturally does not have this problem since it can 'jump' to the controller corresponding to the best-performing plant straight away.

  For example consider the chemical reactor in Figure 1.11 where the overall system dynamic of the reactor is assumed to be governed by two chemicals a and b. Let the parametrisation of the actual reactor be fixed and unknown however assume it to lie withing the pictured parameter surface, symbolising all feasible combinations. Applying a continuous adaptive control algorithm to this non-convex problem, the

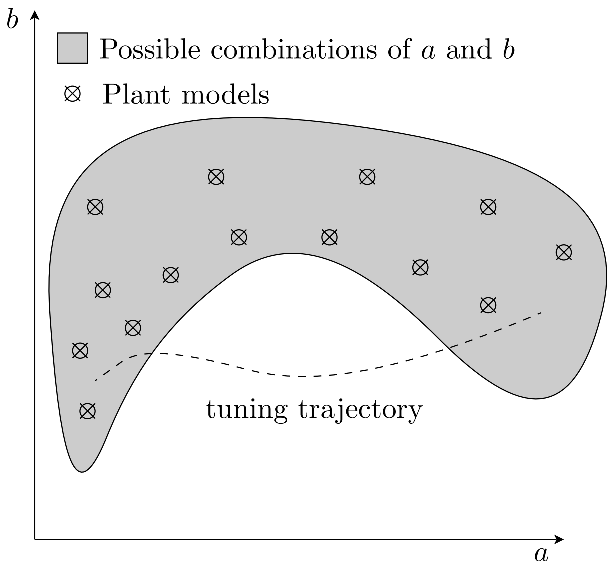

FIGURE 1.11: Tuning versus switching

  tuning trajectory might actually exit the parameter surface and therefore control the reactor based on false, potentially dangerous assumptions about a and b. On the other hand if we utilise MMSAC the algorithm has the freedom to 'jump' the concavity and avoid this problem. (for a further discussion see Hespanha et al. (2003)).

- Copes with the simultaneous stabilisation problem:

  This is another implication of the switching nature of the MMSAC algorithm and simply follows from the fact the algorithm can switch to any controller in the controller set. If we consider the simultaneous stabilisation example in Section 1 concerning two plants $P_{p_1} = \frac{1}{s}$ and $P_{p_2} = -\frac{1}{s}$, we could choose the corresponding controllers to be $C_{K(p_1)} = 1$ and $C_{K(p_2)} = -1$, where we note that the atomic closed loops $[P_{p_1}, C_{K(p_1)}]$ and $[P_{p_2}, C_{K(p_2)}]$ are stable. In the ideal case the algorithm will then switch to (or remain long enough with) the correct controller, so that the system is stable.

- Allows for Multiple Input Multiple Output Systems:
  In the multiple model framework the inclusion of MIMO systems virtually comes for free since there are inherently no restrictions on the dimension of the involved signals. The only (trivial) requirement is that the signal dimensions of the true plant, plant models, corresponding controllers and the switching logic are compatible.

  Although no conceptual limitations prevent MMSAC to be applied to MIMO problems, very little MIMO analysis is available in the literature to date, however see Mosca et al. (2001) for an exception.

- Modularised approach:
  Observe that the problems of performance evaluation and generation of the switching signal q, performed by the switching logic S and the feedback implementation given by the controller C — as depicted in Figure 1.7 — are only interlinked via q and otherwise completely separated. This allows a simplified implementation and analysis of the algorithm since changes in C do not necessarily require changes in S and vice versa; hence they can be designed and analysed separately. In practise this is of great importance since it will reduce the overall complexity of the design process.

For an enthusiastic promotion of multiple model switched adaptive control see Hespanha et al. (2003).

The structural freedom in MMSAC stands out, especially if compared to other adaptive algorithms such as the continuously tuned adaptive controllers introduced in Section 3. Their controller design is completely dictated by the structure of the algorithm itself. Also, they can experience bursting effects for some unfortunate value of the tuned parameter (see Anderson (2005)) hence potentially have difficulties if the plant is parametrised over non-convex parameter sets. The Nussbaum controller is the only continuously tuned algorithm theoretically capable of dealing with the simultaneous stabilisation problem however there is little hope to ever apply it in practice. Furthermore the analysis of classical adaptive algorithms is usually limited to the SISO case and the generalisation to MIMO is extremely cumbersome.

However note that for MMSAC it is common in the literature to impose assumptions on the plant, controller and algorithm to simplify the analysis. Standard such assumptions are that the controllers have a particular design, the plant is SISO, possibly stable, S and C are somehow interwoven, etc. Hence some of the native features are sacrificed for the simplicity of the analysis. The literature to date therefore only reflects a subset of possible MMSAC designs. We will shortly discuss multiple model type algorithms in detail in order to illustrate this.

In contrast we will show that the class of algorithms and the corresponding analysis developed in this thesis will fully achieve the potential by incorporating all the above features. Additionally we will show that the schemes are amenable to a strong robustness analysis, i.e. we will establish explicit gain (function) bounds on the gain from the external disturbances $(u_0, y_0)^\top$ to the internal signals $(u_2, y_2)^\top$ hence we can give (by Georgiou and Smith (1997)) explicit robustness guarantees. We will also show that variations on the schemes leads to non-conservativeness.

Before a discussion of historic multiple model schemes is entered, note that although time-varying systems are not the focus of this thesis there is no inherent assumption in MMSAC that the true plant needs to be fixed, i.e. MMSAC is potentially applicable to similar system classes as the gain scheduled controllers introduced above. The only constraint that we need to satisfy on an algorithmic level is that for every frozen time instance of the plant $P_{p*}(t),\ t \in \mathbb{R}$ there exists a controller in $\mathcal{C}$ such that the atomic closed loop $[P_{p*}(t), C_{K(p)}]$ is gain stable for some $p \in \mathcal{P}_i$. Furthermore the performance of each plant model would have to be evaluated over a shorter horizon to include some kind of 'forgetting' into the algorithm since otherwise the algorithm will slow due to the accumulated history in the performance signals. There does not exist a workable theory on how to apply MMSAC to time-varying problems to date, however for later generalisations the absence of structural obstacles will be necessary.

4.3 [Robust] Multiple Model Adaptive Control ([R]MMAC)

Two historic predecessors to the MMSAC concept were the Multiple Model Adaptive Estimation (MMAE) and Multiple Model Adaptive Control (MMAC) algorithms due to Lainiotis (1971, 1976a,b), Saridis and Dao (1972) and Deshpande et al. (1973). We will now briefly discuss these historic multiple model schemes and note that they are not switched in a strict sense. The global control signal is constructed by 'blending' together various atomic control signals, as with the common practise in gain scheduled control. However the structural similarity of these algorithm in comparison to MMSAC justifies their discussion at this point.

MMAE and MMAC algorithms are set in the stochastic domain and seek to control a fixed LTI plant incorporating some uncertainty. The idea of MMAE is to utilise a bank of Kalman filters for state estimation of an unknown plant. The global state estimate is then calculated by summing over the weighed local state estimates of the Kalman filters where the weights are determined by a so-called "Posterior Probability Evaluator (PPE)" from the Kalman filter residuals (details are omitted). The (state) controller is realised by certainty equivalence with the global state estimate. MMAC is a (control) extension to MMAE where local control signals are calculated from each Kalman filter state estimate and a corresponding LQ controller. Similarly to the construction of the global state in MMAE, the global control signal is constructed by utilising the weights from the MMAE's PPE to generate a weighted sum of the local control signals. This essentially follows the same principle as the discussed gain scheduling algorithm where the performance function is now a stochastic estimation process and the performance signal indicates the probability related to each model. There are a variety of examples where MMAC has been applied successfully: in medical applications He et al. (1986) and Martin et al. (1987); aerospace applications Athans et al. (1975) and Maybeck and Stevens (1990a) and controlling flexible structures Maybeck and Stevens (1990b) and Fitch and Maybeck (1994) (to pick only a few). Maybeck also produced as series of text books in which he discusses the topic extensively Maybeck (1979, 1982a,b).

Recently, in Fekri et al. (2004), a similar scheme denoted Robust Multiple Model Adaptive Control (RMMAC) was introduced however with the difference that the local control designs are by 'state of the art' mixed μ synthesis techniques leading to output feedback controllers (instead of state controllers as in MMAC). The global control signal is constructed as a weighted sum of all local controller outputs, where the weights are generated by a PPE as in MMAE. This allows for MIMO plants. See Fekri et al. (2006) for a nice overview of MMAE, MMAC and RMMAC algorithms where the authors evaluate the performance of their algorithm through many simulations.

Although some of these stochastic algorithms work well in practice, no analytical robustness/performance results have been reported to date; where by robustness we mean that the system remains stable in the presence of input and output disturbances as well as unmodeled dynamics. This prevents a principled performance-orientated design of the algorithm, especially the design of the plant model set. Fekri et al. (2006) try to circumvent this problem and utilise the 'atomic' robustness margins of $[P_p, C_{K(p)}]$ as a measure of performance instead, which leads to a rudimentary performance-orientated design procedure for the plant model set. This is discussed in more detail later in the next section.

We now turn to two completely deterministic MMSAC schemes which are broadly applicable and open to a simplified analysis with the aim of developing hard robustness results: one where the performance function is implemented with observers and one where it is implemented utilising so-called disturbance estimators.

4.4 Observer based MMSAC

(State) observers are intriguing candidates for implementing the performance function in a deterministic setting. They are part of the basic vocabulary of control and have an intuitive interpretation: that an observer attempts to estimate the state of the observed system and that the observed state converges to the true state assuming zero disturbances. This is invariant to the stability of the corresponding plant. Furthermore, performance evaluation via observers provides an interesting link to the stochastic MMAE, MMAC and RMMAC approaches discussed above, since the Kalman filter allows both deterministic and stochastic observer interpretations.

Assume P in Figure 1.1 to be represented by the following state space equations:

$$\begin{aligned} \dot{x} &= Ax + Bu_1 \\ y_1 &= Cx + Du_1 \end{aligned}$$

and assume (A, C) to be observable and the disturbances to be zero, i.e. $(u_0, y_0) = 0$.

We then have with $(u_0, y_0)^\top = (u_1, y_1)^\top + (u_2, y_2)^\top$ that $(u_1, y_1)^\top = -(u_2, y_2)^\top$ hence

$$\begin{aligned} \dot{x} &= Ax - Bu_2 \\ y_2 &= -Cx + Du_2. \end{aligned}$$

A typical (Luenberger type) observer for P is then given by

$$\begin{aligned} \dot{\hat{x}} &= A\hat{x} + L(y_2 - \hat{y}_2) - Bu_2 \\ \hat{y}_2 &= -C\hat{x} + Du_2. \end{aligned}$$

The purpose of the second term in the observer state equation, $L(y - \hat{y})$, is to force output error convergence between the observer and the true output, where the choice of the matrix L is of major importance. To see this let the state error e be given by $e = \hat{x} - x$. We then have

$$\begin{aligned} \dot{e} &= \dot{\hat{x}} - \dot{x} = A\hat{x} + L(y_2 - \hat{y}_2) - Bu_2 - Ax + Bu_2 \\ &= A(\hat{x} - x) + L(-Cx + Du_2 + C\hat{x} - Du_2) \\ &= Ae - LCe \\ &= (A - LC)e \end{aligned}$$

hence if we choose L such that $A - LC$ has eigenvalues with strictly negative parts then the observer state asymptotically converges to the state of the true system, i.e. $\hat{x}(t) \to x(t)$ as $t \to \infty$. This nice property motivates the choice $d_p = e_p$ as the performance function, where $p = (A, B, C, D)$. For a plant model set $\mathcal{P}_i = \{p_1, p_2, \cdots, p_i\},\ i \in \mathbb{N}$ a observer bank with corresponding output errors is depicted in Figure 1.12 which we

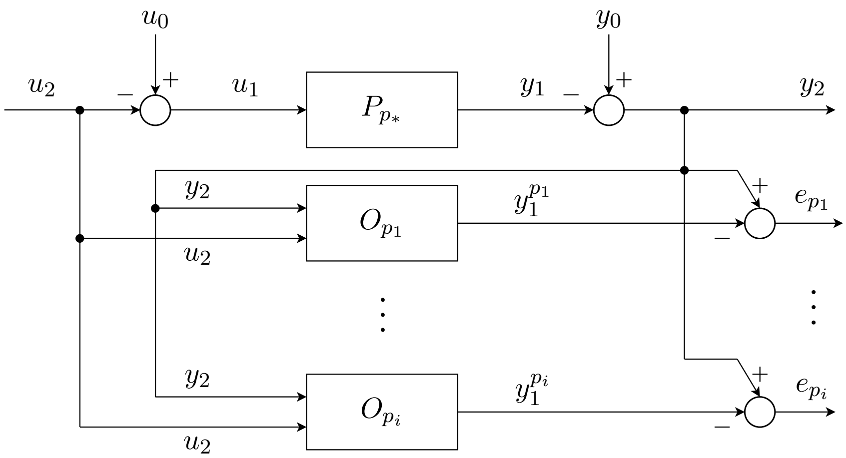

FIGURE 1.12: Observer bank

utilise to construct S. Typically the switching signal is then computed from $e_p,\ p \in \mathcal{P}_i$ by integrating the output error hence to account for the history of the error signal in the switching decision, e.g.

$$q_f(t) = \underset{p \in \mathcal{P}_i}{\operatorname{argmin}} \int_0^t |e_p(\tau)|^2 d\tau.$$

One obvious implication of utilising observers is that the class of systems we seek to control must allow the construction of observers in the first place. Although this is well understood in the linear domain with the notion of observability, things become less clear in the non-linear case, e.g. see Hespanha et al. (2002). There exist various non-linear observer designs such as high gain observers, sliding-mode observers or non-linear extended state observers, however their design and application is far from being trivial. Also fundamental questions related to error convergence or robustness often remain unanswered. This implies that the very fact that the algorithm relies on observers will complicate if not preclude a later generalisation to a wide class of non-linear systems.

Various authors have conducted analytical studies of observer based MMSAC algorithms in order to show their stability and robustness. The most prominent one is due to Morse (1996), where the author shows the asymptotic convergence of the output to a constant reference signal r of an observer based MMSAC algorithm controlling a fixed LTI SISO plant $P = \frac{\alpha_{p*}}{\beta_{P*}}$ in the presence of a constant disturbance d with $n = 0,\ \delta^a_{p_*} = \delta^m_{p_*} = 0$ (Figure 1.13 B) where $\frac{\alpha_{p*}}{\beta_{p*}}$ is proper, β_{p_*} is monic and ν_{p_*} is a polynomial of degree less then β_{p_*}. Later in Morse (1997) $\delta^a_{p_*}, \delta^m_{p_*}$ are allowed to be non-zero. The plant model set is always required to be compact.

Morse addresses the problem of implementing a large number of observers in parallel from the start. In particular he shows that observers can be written in a state shared fashion, i.e. some estimator state x_E is common to all observers which is generated by the 'multi estimator' Σ_E in Figure 1.13 (A), hence only the output equation $y_p = C_p x_E$ must

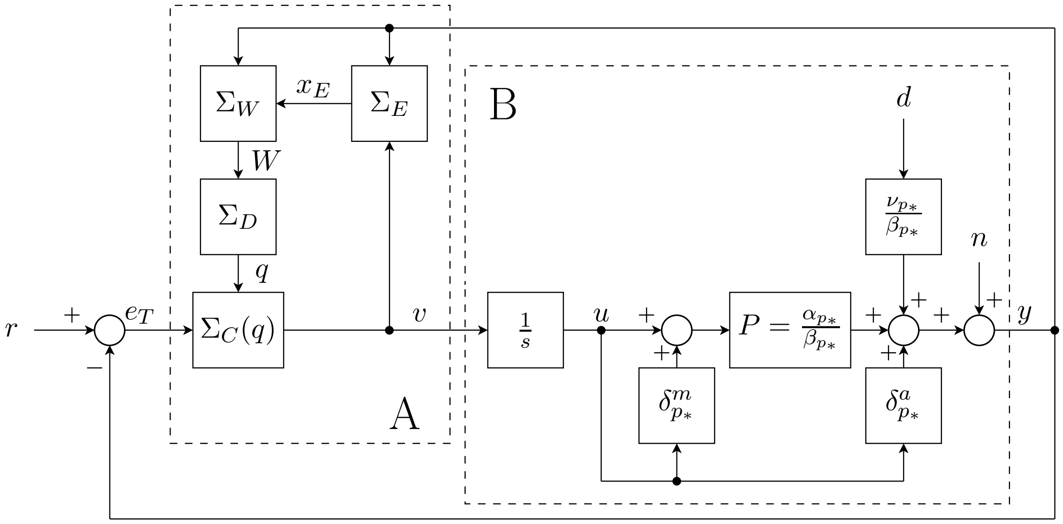

FIGURE 1.13: Closed loop system considered in Morse (1996, 1997)

be evaluated for all $p \in \mathcal{P}$. Therefore Σ_E is a state shared implementation of Figure 1.12. This reduces the computational complexity from evaluating i observer state and output equations to the complexity of evaluating one observer state equation and i output equations which is significant for a later implementation. However, computational trade-offs are expected, especially for large plant model sets.

The building block Σ_W returns an exponentially weighted matrix W which is constructed from y and x_E (details are omitted). Σ_D then determines from W and C_p the observer which performs best and lets q point to the corresponding plant model (after some suitable delay, or dwell time). In short, q points to the plant model which corresponding observer shows the smallest output error (measured in some weighted L_2 norm), where q is suitably delayed.

In Morse (1997) the algorithm from Morse (1996) is then shown to be stable in the presence of additive and multiplicative perturbations δ^a_{p*} and δ^m_{p*} of suitable size where d, n are non-zero bounded, piecewise-continuous disturbance inputs (Figure 1.13 (B)). In Morse (2004) the setup of Morse (1996) is revisited and explicit bounds on the gain from the disturbance d to the tracking error e_T are given, however with d entering the system before the plant P (without a corresponding disturbance weight), and where the output disturbance n is assumed to be zero.

Although these results are significant (especially since they are the first of their kind) they are unfortunately not able to preserve all of the potential features promised by the MMSAC concept as stated in Section 4.2:

- LTI SISO:
  All results are given for LTI SISO plants. A discussion on how suitable the given analysis is for a later generalisation to MIMO and non-linear systems is mostly

absent. It is noted that the problem of controller construction for a continuum of plant models becomes more complicated in the MIMO and non-linear case. Initial work has been performed by Mosca et al. (2001).

We note that Morse's analysis relies heavily on transfer function and linear state space equation notation that inherently ties it to the linear domain and makes a later generalisation at least cumbersome. Stability analysis aside, we have mentioned before that even the construction of observers can be problematic in the non-linear case, e.g. see Chang et al. (2001) where a great deal of attention goes to the construction and stability proof of the utilised non-linear state observer.

- No complete unstructured uncertainty model:
  In order to obtain a comprehensive robustness result it is essential to not only deal with additive and multiplicative uncertainty but also with uncertainties of an inverse multiplicative type. A common way to address the complete unstructured uncertainty problem in the linear domain is to employ a coprime perturbation model, discussed in Section 4, page 53, where it is required that we allow for two possible disturbance inputs placed symmetrically on both sides of the plant as in Figure 1.1. The disturbance model in Figure 1.1 is also utilised in the robustness theory of Georgiou and Smith (1990, 1997) hence if we are able to establish a finite bound on the gain from the disturbances to the internal signals we automatically have a powerful robustness analysis at hand, i.e. for non-linear systems see Georgiou and Smith (1997), Theorem 1.

  In contrast, the model in Figure 1.13 from Morse (1996, 1997) injects the disturbance d, n after the plant. To move d to the left beyond $\frac{1}{s}$ is impossible if the true plant does not have a 'natural' pole at zero and we were forced to artificially augment the plant input with an integrator in order to meet the given plant constraints since then the noise input d would act on the controller. These structural issues essentially prevent a straight forward generalisation of the results to allow for general unstructured uncertainties and a direct application of known robustness results.

- Pole at zero:
  Morse assumes that the true system possesses a pole at zero. This assumption is rather restrictive since it only holds for a small class of physical systems, e.g. a mass, spring, damper arrangement only has a pole at zero if the spring constant is zero (there is no spring). In order to apply Morse's theory we would therefore have to artificially augment systems with an integrator, which has other undesirable effects, as discussed above.

- Results only in a (weighted) L_2 setting:
  It is desirable to have gain stability results not only in L_2 but also in other relevant signal spaces, for example L_∞ is interesting in practice since it deals with possible offsets naturally present in any physical signal.

A different type of analysis deals with the improvement of performance in MMSAC systems while preserving stability. Narendra and Balakrishnan (1993) considers a number of parametrised adaptive controllers and utilises a switching scheme to select one of them for closed loop operation. In Narendra et al. (1995) and Narendra and Balakrishnan (1997) the authors extend Narendra and Balakrishnan (1993) and also consider fixed controllers and re-initialised adaptive controllers alongside the (free-running) adaptive controllers from Narendra and Balakrishnan (1993). They also perform extensive numerical simulation evaluating the performance of different combinations of the mentioned controller choices. No gain bounds or robustness margins are given where the authors merely note: "Since all the models used in the procedure [...] are either fixed or adaptive, one would expect the overall system to be robust under perturbations, if each model-controller pair is individually robust. This indeed turns out to be the case." Such robustness results would therefore at best inherit the limitations of the robustness theory for classical adaptive controllers and no explicit details were given.

For the present analysis of MMSAC algorithms we therefore conclude the following:

- Limitations by analysis:
  For all discussed algorithms there remains a vast gap between theory and practice, e.g. the wide gap between the class of systems MMSAC algorithms can be used for (see the list at the beginning of Section 4.2) or can be implemented for in practice and the class of systems that the analysis applies to. Also the assumption that the plant set is compact postulates a priori knowledge of a bound on the uncertainty, which makes the algorithm conservative (Chapter 6 shows how such limitations may be removed for the algorithm under consideration in this thesis). This is very unfortunate since non-conservativeness is thought to be one of the key benefits of adaptive control.

  Reducing complexity by limiting scope is the natural thing to do when approaching complex problems however the adoption of many structural assumptions into the analysis seems to have inhibited the generalisation effort over the last decade, i.e. the system classes considered by the authors remain virtually unaltered to date.

- Limited robustness results:
  We have argued that although the theory of Morse (1996, 1997, 2004) is suitable for showing stability in the presence of additive and multiplicative uncertainty, it fails to fully incorporate unstructured uncertainty since this is disallowed by the structure of the utilised disturbance model. The claim of robustness in Narendra and Balakrishnan (1997) can be considered problematic since it relies on traditional robust adaptive control results. For all other approaches robustness results do not exist at all.

  Since any control algorithm is subjected to input and output disturbances as well as unmodeled dynamics in practice, the corresponding robustness theory must include these. Furthermore explicit measures of the algorithm's performance, e.g. in the form of gain bounds, are very important since they are essential to conduct performance-orientated design (see below).

- No theoretically grounded design methodology:
Since in MMSAC the controller design procedure K is usually given, design relates almost exclusively to the question on how to choose an appropriate plant model set $\mathcal{P}_i$ controlling an uncertain plant P. For example if

$$P_p = \frac{1}{s-p}$$

where $1 \leq p \leq 10$ is an uncertain parameter, how should the finite plant set $\mathcal{P}_i$ be chosen?

Recall that in MMSAC there exists the basic requirement that every atomic plant model controller pair $[P_p, C_{K(p)}],\ p \in \mathcal{P}_i$ is gain stable. Furthermore we must ensure that the plant models are distributed such that over the whole uncertainty set of P there always exists a controller $C_{K(p)},\ p \in \mathcal{P}_i$ such that $[P_p, C_{K(p)}]$ is gain stable for all $1 \leq p \leq 10$. Otherwise we could find a value of p such that all controllers are destabilising. For example we might choose $\mathcal{P}_i = \{1, 5, 10\}$ where the plant models are given by $P_p,\ p \in \mathcal{P}_i$, however, we then have to ensure that all the intermediate parametrisations $a \in [1, 10] \setminus \{1, 5, 10\}$ can be stabilised by at least one of the corresponding controllers $C_{K(p)},\ p \in \mathcal{P}_i$.

This basic relationship between placement of plant models and atomic stability is exploited in Anderson et al. (2000). It is shown that to an uncertain plant P, incorporating a bounded real uncertainty, and the corresponding compact uncertainty set $\mathcal{P}$, a finite plant model set $\mathcal{P}_i \subset \mathcal{P}$ can be constructed such that there always exists a corresponding controller that stabilises P, i.e.

$$\exists n \in \mathbb{N},\ \mathcal{P}_i \subset \mathcal{P},\ |\mathcal{P}_i| < n,\ s.t.\ \forall p_* \in \mathcal{P}\ \exists p \in \mathcal{P}_i\ s.t.\ [P_{p_*}, C_{K(p)}] \text{ is gain stable.}$$

The argument can be constructed by noting that for each atomic plant-controller pair $[P_p, C_{K(p)}]$, which is required to be (gain) stable, by standard linear robust stability theory, there exists a robustness margin of radius $b_{P_p, C_{K(p)}}$ around each $P_p,\ p \in \mathcal{P}$. The union of these neighbourhoods with radii corresponding to the robustness margins then results in a cover. By compactness, this cover of $\mathcal{P}$ therefore has a finite sub-cover which determines the finite plant model set $\mathcal{P}_i$. This leads to the desired result.

In Fekri et al. (2006) the authors are explicitly interested in a performance-orientated design guideline for the plant model set $\mathcal{P}_i$ and the corresponding controller set. They utilise the performance of atomic plant-controller pairs $[P_p, C_{K(p)}]$ for that purpose. For a scalar uncertainty p the number and distribution of the plant models is then determined by an iterative process. Starting from the upper bound of the uncertain parameter p_u, $p \leq p_u$ the performance of the first controller design C_1 (where the controller is constructed for each new uncertainty interval by mixed μ synthesis techniques) is evaluated for the atomic closed loop $[P_p, C_{K(p)}]$ over the uncertainty interval $\alpha_1 \leq p \leq p_u$ for decreasing α_1. Since the performance of the (fixed) controller under increasing uncertainty will naturally decrease (it is conservative) it will eventually cross some pre-defined lower performance bound A. This event defines α_1 and the first controller C_{p_1}. The procedure is then repeated for $\alpha_2 \leq p \leq \alpha_1$ with α_2 decreasing until the lower bound of the uncertainty p_l, $p_l \leq p$ it reached. This implicitly defines $\mathcal{P}_i$. Since the atomic performance of matching plant and controller pairs $[P_p, C_{K(p)}]$ is also a measure of atomic robustness this essentially relates the number of plant models to the atomic robust stability margins as in Anderson et al. (2000); however it also gives a design guideline for $\mathcal{P}_i$.

Note that for non-parametric uncertainties the design problem becomes more complex since we also have to consider the geometric distribution of the plant models.

We conclude that the relationship between the number of plant models and atomic robustness margins is in the one dimensional case, at least conceptually, well established however design on this level, although it produces answers, remains heuristic since we do not know how a particular construction of $\mathcal{P}_i$ will effect the global performance of the algorithm. It is obvious that we are missing a key constraint in the form of a global measure of performance in order to sensibly optimise $\mathcal{P}_i$.

This motivates Chapter 6 where we will give performance-orientated design guidelines for the EMMSAC algorithm.

We will now focus on the class of switching algorithms where the performance function is implemented by some optimal estimator.

4.5 Estimation-based Multiple Model Switched Adaptive Control

The idea of EMMSAC, i.e. to utilise optimal disturbance estimation for performance evaluation, is due to Fisher-Jeffes (2003) and Vinnicombe (2004). It forms the basis for what follows. To emphasise the optimality aspect of EMMSAC we will introduce the algorithm from a system identification point of view.

Let $\mathcal{P}$ be the class of systems under consideration. Consider Figure 1.14 where the signal $(u_2, y_2)^\top$ is an observed measurement of a dynamical system, i.e. from Figure 1.1, where for simplicity we have assumed that all signals are bounded. Let $\mathcal{M}_p$ denote the graph of P_p, $p \in \mathcal{P}$, that is the set of all allowable (or compatible) bounded input-output combinations $(u_1^p, y_1^p)^\top$ of P_p.

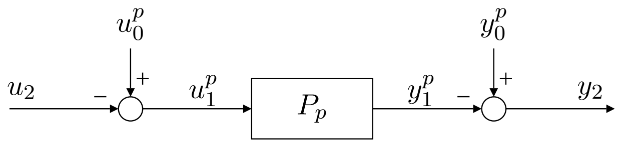

FIGURE 1.14: System identification from the observation $(u_2, y_2)^\top$

Now consider an optimal system identification algorithm that determines the plant $P_{q_f},\ q_f \in \mathcal{P}$ such that the error $(u_0^{q_f}, y_0^{q_f})^\top$ between the observed signals $(u_2, y_2)^\top$ and the signals $(u_1^{q_f}, y_1^{q_f})^\top$ in the graph of P_{q_f} is minimal at time $k \in \mathbb{N}$:

$$\begin{aligned} q_f(k) &= \operatorname*{argmin}_{p\in\mathcal{P}} \left(\min_{(u_1^p,y_1^p)^\top \in \mathscr{T}_k \mathcal{M}_p} \|\mathscr{T}_k(u_2,y_2)^\top + \mathscr{T}_k(u_1^p,y_1^p)^\top\| \right) \\ &= \operatorname*{argmin}_{p\in\mathcal{P}} \left(\min_{(u_1^p,y_1^p)^\top \in \mathscr{T}_k \mathcal{M}_p} \|\mathscr{T}_k(u_0^p,y_0^p)^\top\| \right) \end{aligned} \tag{1.5}$$

where $\mathscr{T}_k v$ represents the truncation of a signal v a time $k \in \mathbb{N}$. The inner minimisation can also be thought of as a (metric) projection onto the graph $\mathscr{T}_k\mathcal{M}_p$ of P_p — as depicted in Figure 1.15 — hence it represents the distance between the plant model and the observation $\mathscr{T}_k(u_2, y_2)^\top$.

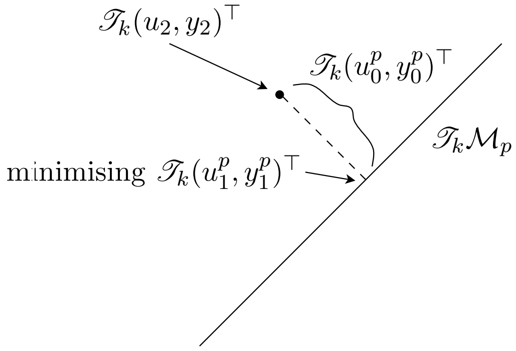

FIGURE 1.15: Projection onto the graph $\mathscr{T}_k\mathcal{M}_p$ of $P_p,\ p \in \mathcal{P}$ at time $k \in \mathbb{N}$

A possible multiple model control strategy employing this identification scheme would put a controller designed to stabilise the 'identified' plant $q_f(k)$ into closed loop. A concrete design procedure would then be as follows.

- For all $P_p,\ p \in \mathcal{P}$ construct a corresponding controller $C_{K(p)}$ such that the atomic closed loop $[P_p, C_{K(p)}]$ is stable.
- For all $P_p,\ p \in \mathcal{P}$ use the above procedure to construct the smallest 'disturbance estimate' $\mathscr{T}_k(u_0^q, y_2^q)^\top$ that is consistent with the observation $\mathscr{T}_k(u_2, y_2)^\top$ and the plant P_p (inner minimisation) up to time $k \in N$.

- Let $q_f(k) \in \mathcal{P},\ k \in \mathbb{N}$ point to the plant model which corresponding disturbance estimate is minimal for all $p \in \mathcal{P}$.
- Switch the controller $C_{K(q_f(k))}$ corresponding to the plant $P_{q_f(k)}$ into closed loop at time $k \in \mathbb{N}$.

System identification on its own is a large and very active area of research and an in depth analysis of existing algorithms would go beyond the scope of this thesis (Ljung (1999) gives a good overview). We regard identification algorithms not minimising $\|(u_0^p, y_0^p)^\top\|$ as in equation (1.5) as non-optimal. Various solutions to the non-optimal identification problem (see Ljung (1999) and the references therein) and optimal identification problem (e.g. see Markovsky et al. (2005) for a recent approach via structured least-squares) are known, however the inclusion of disturbances into the analysis usually poses a major complication. Also concrete algorithms may face the problem of local minima if $\mathcal{P}$ is non-convex and are usually limited to off-line application.

A straightforward way to simplify the identification problem is to consider only a discrete, finite plant model set $\mathcal{P}_i \subset \mathcal{P}$ since this would reduce the outer minimisation problem in equation (1.5) to the simple comparison of finitely many scalars. The discrete identification problem would then read

$$q_f(k) = \operatorname*{argmin}_{p \in \mathcal{P}_i} \left(\min_{(u_1^p, y_1^p)^\top \in \mathscr{T}_k \mathcal{M}_p} \|\mathscr{T}_k (u_0, y_0)^\top\| \right). \tag{1.6}$$

Observe that discrete identification can be an approximate of the identification over the whole of $\mathcal{P}$ as in equation (1.5), i.e. if $\mathcal{P}$ represents a continuum.

In Chapter 3 we will show that in L_2, l_2 the size of the disturbance estimate can be determined from the residuals of a Kalman filtering process which allows a direct on-line implementation of the disturbance estimator and underlines the close relationship to MMAE and MMAC (in this special case). Furthermore, since the analysis presented in this thesis requires disturbance estimates to be optimal only over some finite interval $[k-\sigma, k],\ \sigma, k \in \mathbb{N}$ (also see French and Trenn (2005)), and since the inner minimisation problem is usually convex, e.g. in the linear case, there exist simple (matrix optimisation) techniques of bounded computational complexity to compute finite horizon disturbance estimates in a general $l_r,\ 1 \leq r \leq \infty$ norm setting (see Chapter 3).

The switching function in equation (1.6) forms the heart of the EMMSAC algorithm considered in this thesis.

As noted, the idea to utilise optimal disturbance estimation for performance evaluation is due to Fisher-Jeffes (2003) and Vinnicombe (2004). It was the key that allowed for a simplified, axiomatic analysis in French and Trenn (2005), it opened the door to further generalisations in Buchstaller and French (2007, 2008) and also made this thesis possible. Also the clever treatment of disturbances in the analysis and the fundamental properties of estimators deduced in French and Trenn (2005) live on in the present result and guide the developments to date. Last but not least the knowledge that finite bounds on the gain from the external disturbances to the internal signals promises robustness (Georgiou and Smith (1997)) defines the overall setting and objective of the analysis.

In Vinnicombe (2004) the author was able to establish initial bounds on the gain from the external disturbances to the internal signals for a plant model set consisting of only two plant models $\frac{1}{z}, -\frac{1}{z}$ (the simultaneous stabilisation problem). Fisher-Jeffes (2003) was able to show that such gain bounds can be established by either a version of Lyapunov stability theory adapted to account for switching or linear matrix inequalities (LMIs) for any two plant models. Furthermore he shows that the problem of determining optimal disturbance estimates is equivalent to the problem of calculating the scaled residuals of a Kalman filter. Unfortunately the path to a more general result is rather unclear since the reasoning is specifically tied to the analysis of two plant models. A novel way of treating disturbances in French and Trenn (2005), i.e. to utilised input and output disturbances as a central part of the argument instead of considering them an unwanted nuisance, then opened up the algorithm to a simplified and axiomatic analysis. This change of perspective allowed the authors to first state four general assumption on the disturbance estimator and then to establish $l_r,\ 1 \leq r \leq \infty$ gain bounds for the class of dead-beat stabilisable plants[1] based on these abstract assumptions, thus divorcing the problem of (optimal) disturbance estimation and the robust stability analysis of the algorithm. The authors also introduce a finite horizon disturbance estimator which is only optimal over a finite interval however it is shown to meet the (estimator) assumptions and is therefore applicable.

Finally we emphasise the key differences and similarities between traditional observer based MMSAC control algorithms in the sense of Morse et. al., and the introduced EMMSAC control algorithm. Observe that:

- The disturbance estimate in EMMSAC replaces the observer error in MMSAC as the performance signal.
- Observers, similarly to estimators, give some notion of distance from the observed signal $(u_2, y_2)^\top$ to the plant it has been constructed for, however this distance is in general not minimal (optimal) in the given sense.
- In l_2 the Kalman filter may be utilised for optimal disturbances estimation. Observe that the Kalman filter estimator is also an observer or has observer structure. Hence in the special case where Kalman filters are utilised for performance evaluation, EMMSAC and MMSAC algorithms coincide in the performance function.

[1] The class of dead-beat stabilisable plant is the class of plants where to every member P there exists a dead-beat controller C such that for the closed loop $[P, C]$ and zero disturbances the output y_2 is forced to zero in one time step.

To conclude the introduction we set the contributions of our work into perspective against previous results.

5 Contributions of this thesis

We start by discussing what we consider the most important contribution of this thesis: the axiomatic treatment of the problem in the theory. All results are the fruits of this abstraction effort.

Axiomatic treatment in theory

- Robustness:
  By Georgiou and Smith (1997) finite bounds on the gain from the disturbances $(u_0, y_0)^\top$ to the internal signals $(u_2, y_2)^\top$ translate into explicit robustness guarantees. This fact motivates the overall setting of the analysis: to show for an algorithm that such bounds exist. In Chapter 5 this is done explicitly for the EMMSAC algorithm. This way of showing robustness is rather different to the one for example in Morse (1996, 1997), where the author proves error convergence for the algorithm in the presence of additive and multiplicative uncertainty and output disturbances.

  The gain bound approach to robustness has the advantage that it essentially cleans the analysis of any uncertainty related objects, in fact robustness can be completely neglected at first, one merely has to show that such a gain bounds exists and robustness follows.

- Estimators:
  In Chapter 3 we establish abstract assumptions on the disturbance estimator on which the subsequent analysis will rest. This axiomatic treatment of the relevant estimator properties initiated by French and Trenn (2005) has the advantage that unlike other multiple model adaptive algorithms to date — which are tied to one specific performance evaluating element, i.e. the so-called multi estimator in MMSAC (Morse (1996, 1997)) or Kalman filters in MMAC or RMMAC — we are free to choose any estimator that fulfils the assumptions. In particular we will show that optimal (finite and infinite horizon) disturbance estimation algorithms fulfil these assumptions. The optimal infinite horizon estimators in an l_2 setting are closely related to the Kalman filter however different horizons length and signals spaces give rise to a variety of different estimators.

- Atomic plant-controllers pairs:
  Similar to the estimator assumption we will not explicitly give a controller design or even assume a certain representation of the plant models and corresponding controllers, i.e. state space matrices or transfer functions, but only require the atomic

loop interconnection $[P_p, C_{K(p)}]$, $p \in \mathcal{P}$ to satisfy two rather simplistic (linear) signal growth assumptions (see Chapter 4) on which the subsequent analysis will rest. We call these two assumptions the controller assumptions. This approach is fundamentally different to other theoretical treatments of this problem in the literature, i.e. in MMAC controllers are required to be state controllers, in MMSAC (Morse (1996, 1997)) we are free to chose the controllers however the notation is based on linear transfer matrix function notation. Also in Fisher-Jeffes (2003) the analysis is tied to state space notation where the controller design is fixed (H_∞). This axiomatic treatment leads to the greater generality of the algorithm as discussed next.

We began the discussion of multiple model switched adaptive algorithms in Section 4.2 by compiling a list of desirable features that an algorithm could possess. During the course of this thesis we will show that in fact all these features are preserved by the analysis. We will now discuss these and additional features of the algorithms, where we note that they almost exclusively follow directly from the axiomatic treatment of key elements as described above.

Generality of the algorithm

- Broad system class, full controller design freedom:
  In MMSAC (Morse (1996, 1997)) it is assumed that the true plant has a pole at zero and we noted that this poses problems in generalising the underlying disturbance model. Unfortunately, since a great deal of the analysis rests on this assumption, it is not straightforward to remove it. In French and Trenn (2005) the system class is limited to dead-beat stabilisable systems where the controllers are dead-beat. For classical adaptive controllers it is often imposed that the plants are minimum phase and the relative degree as well as the sign of the high frequency gain is known, i.e. see Narendra and Annaswamy (1989).

  We do not require any such assumptions for EMMSAC.

  We will only require that the controller design procedure $K : \mathcal{P} \to \mathcal{C}$ is such that any atomic closed loop pair $[P_p, C_{K(p)}]$, $p \in \mathcal{P}$ satisfies the controller assumptions; this can be achieved by **any** control design methodology. It will be shown in Chapter 4 that for linear systems this assumption simply relates to atomic closed loop pairs that are (gain) stable. However this assumption is also satisfied by non-linear atomic closed loop pairs which show linear growth. Although non-linear systems are not the focus of this thesis we note that this fact brings a fully non-linear treatment of the problem within reach.

- MIMO:
  The majority of adaptive algorithms in the literature are assumed to be operating on SISO plants. For classical adaptive controllers this arises from the structure

of the problem since for general MIMO systems it becomes rather difficult to construct an appropriate parameter update law. As mentioned before, multiple model type algorithms do not have these structural problems and MIMO integrates almost naturally into the scheme. The restriction to SISO for example in Morse (1996, 1997, 2004) and Narendra and Annaswamy (1989) therefore originates from simplicity of the analysis (however see Mosca et al. (2001) where a MIMO MMSAC design is considered).

Since many control applications are indeed MIMO problems, the creation of an algorithm and corresponding robustness result that allows for MIMO is considered an important issue in MMSAC. For example Fekri et al. (2006) acknowledge this fact and specifically design their Robust Multiple Model Adaptive Control (RMMAC) algorithm such that it can be applied to the MIMO case.

In the case of EMMSAC, the axiomatic nature of the approach allows MIMO almost by accident. Recall that from a plant model and controller point of view it is required that the atomic closed loop pairs $[P_p, C_{K(p)}]$, $p \in \mathcal{P}$ satisfy the controller assumptions. It turns out that it is irrelevant if they are MIMO or SISO since the assumption only deals with the size of signals. For the construction of the estimator, the optimisation problem simply becomes higher dimensioned, which is computationally more expensive but otherwise unproblematic (see the point 'Optimisation based performance evaluation' below).

- Non conservative:

  One key problem that sparked the investigation of adaptive control algorithms was the conservativeness of linear controllers; indeed basic continuously tuned adaptive controllers have the virtue of being universal. In order to ensure the stability of the algorithms in the presence of disturbances and unmodeled dynamics, various modifications (dead zones, projections, σ-modification, etc.) were considered which lead to the introduction of some conservativeness to the design. However see French (2008) where the author was able to establish that the underlying unmodified universal controllers are robust to unmodeled dynamics in the presence of sufficiently small disturbances. In MMSAC (Morse (1996, 1997)) the algorithm is limited to compact plant model sets, which translates into the condition that there must be a known bound on the uncertainty of the plant, in turn leading to conservativeness. In French and Trenn (2005) performance degrades for increasingly large uncertainties.

  The basic EMMSAC designs presented in this thesis are also conservative, however we will present a variant of the EMMSAC algorithm in Chapter 6 that maintains its performance invariant to the size of the uncertainty — it is universal.

- Continuous plant sets:

  Unlike French and Trenn (2005), where the established bound on the gain from the disturbances to the internal signals scales with the number of elements in the plant set $\mathcal{P}_i$ (hence the robustness guarantee is lost for large sets $\mathcal{P}_i$), we will show in Chapter 5 that the present gain bound is invariant to the number of plant models in $\mathcal{P}_i$. Instead it depends on the 'complexity' of $\mathcal{P}_i$ (see Chapter 6). Note that the analysis in MMSAC (Morse (1996, 1997)) is also invariant to the number of plant models within a plant model set $\mathcal{P}_i$. However see Hespanha et al. (2001) where for the same algorithm the established bound on the size of the state as well as the robustness margin scale with the number of elements in the plant model set. The authors then propose a modification to the switching logic to circumvent this issue.

- Optimisation based performance evaluation:
  Usually multiple model type algorithms are based on specific implementations of the performance evaluator, i.e. the multi estimator in MMSAC (Morse (1996, 1997)), or Kalman filters in MMAC and RMMAC. The present analysis rests on abstract estimator assumptions that can be satisfied by infinite and finite optimal disturbance estimators (see Chapter 3). Since in particular the finite horizon disturbance estimation problem is a standard convex optimisation problem with many possible solutions, i.e. in l_2 via the pseudo inverse or in l_∞ via linear programming, the analysis is applicable to a variety of algorithms. Furthermore the Kalman filter provides a finite dimensional realisation of the infinite horizon l_2 optimal estimator.

  Even in the non-linear domain, under appropriate convexity assumptions, the finite horizon optimisation problem remains computationally tractable.

- Any $l_r,\ 1 \leq r \leq \infty$ norm:
  Since the entire analysis is based on the gain relationships between signals, or parts of signals, it can be conducted in any $l_r,\ 1 \leq r \leq \infty$ norm. In contrast, algorithms that are based on Kalman filter state estimates such as MMAC or RMMAC only apply in the L_2, l_2 setting. The stability proofs for MMSAC (Morse (1996, 1997)) are also limited to L_2.

- Fully modularised:
  The reason why modularisation is very important is twofold. Firstly it simplifies the analysis since every sub-component can be analysed separately. For example we would like to argue about plant model set design, controller design, and construction of efficient estimators individually, since each component is complex enough in its own right. Secondly it allows the actual implementation to be modularised. This means that the individual sub-components can be constructed separately, where changes in one component does not require changes in a second component. For example in Fisher-Jeffes (2003) this is not the case and parts of the estimator are utilised to construct the controller, hence a later modification to either the controller or the estimator would imply a complete re-analysis and re-implementation of both of them. This also significantly hinders later generalisations.

On the other hand for example, if a design uses common states in the estimator (e.g. a Kalman filter) and the controller then an implementation can exploit this.

- Fully unstructured uncertainty model:
  In contrast to MMSAC Morse (1996, 1997) and classical adaptive, continuously tuned schemes, where only additive and multiplicative uncertainties are permitted by the corresponding stability and robustness analysis, we allow the plant to be perturbed by a fully unstructured uncertainty in the gap metric which is very important for a later implementation. This is a direct consequence of the robustness analysis in the style of Georgiou and Smith (1997).

- No (stochastic) assumptions on the disturbances:
  In many publications in control the analysis is simplified by imposing assumptions on the disturbances which are acting on the system. Standard assumptions are that the disturbances are produced by a stationary Gaussian processes, that they are white, sufficiently rich, Lipschitz differentiable, zero, etc. We will only require that they are bounded in an l_r, $1 \leq r \leq \infty$ norm. This is a further benefit of conducting a robustness analysis in the style of Georgiou and Smith (1997).

- Non-convex and simultaneous stabilisation control problems:
  That EMMSAC is applicable to such problems follows from its multiple model nature.

Finally we want to emphasise the importance of the following contribution: a theoretically grounded design methodology. From an implementation perspective, any body of theory can only be of value if it eventually, directly or indirectly, finds its way into a practical application. Design guidelines that lead to a solution to a given (control) problem, under utilisation of the available design freedom in the algorithm with respect to performance and uncertainty description, are an essential tool that allows this transition.

Design grounded in theory

As discussed in Section 4.4, attempts have been made in Anderson et al. (2000) and Fekri et al. (2006) to establish initial design guidelines for the plant model sets of multiple model algorithms. In both cases, plant model sets are constructed on the basis of atomic robust stability margins rather than any global measure of performance; hence from a global performance point of view these constructions are heuristics. The authors draw attention to this and explicitly ask the following questions:

1. How to divide the initial large parameter uncertainty set into N smaller subsets?
2. How to determine the 'size' or 'boundary' of each parameter subsets?
3. How large should N be? Presumably the 'larger' the N, the 'better' the performance of the adaptive system should be.

Additionally we may ask:

4. How to prevent a conservative design?

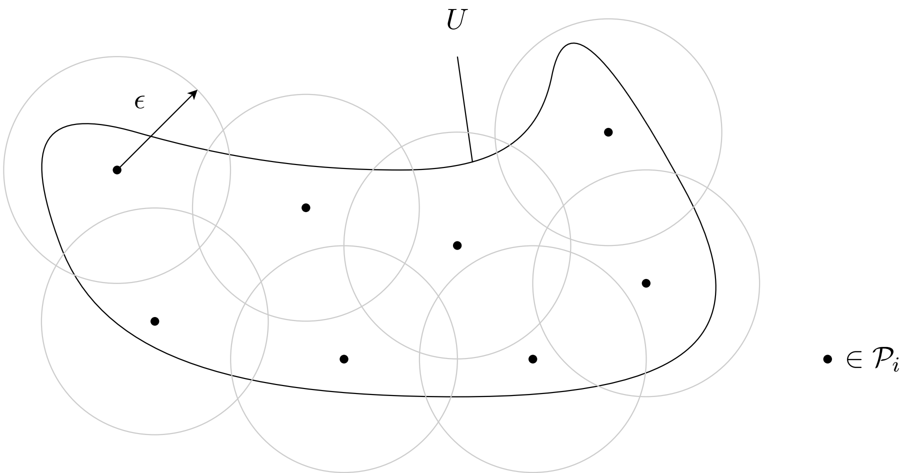

FIGURE 1.16: Covering U by neighbourhoods of size ϵ around $p \in \mathcal{P}_i$

In this thesis the problem of 'dividing up' the uncertainty into neighbourhoods will present itself in the following way. Let the set $U \subset \mathcal{P}$ denote the uncertainty in the true plant $P = P_{p_*}$, i.e. $p_* \in U$. For the sake of the argument, take U to be a compact continuum in a finite dimensional parameter space. The techniques in this thesis now construct an infinite dimensional (unrealisable) multiple model adaptive controller based on a continuum of plant models $\mathcal{P}_i = U$ (and a continuum of corresponding estimators). This controller stabilises any $p_* \in U$ and provides a robustness margin of size ϵ, which depends on U.

However, since the amount of available computational resource is usually finite, we want $\mathcal{P}_i$ to be finite and need to discretise the controller. Since ϵ defines a global robust stability margin around each $p_* \in \mathcal{P}_i \subset U$, we can arrange $\mathcal{P}_i$ such that the neighbourhoods of robustness 'cover' U in order to ensure stability for all $p_* \in U$ — as depicted in Figure 1.16. The theory then ensures that the corresponding multiple model adaptive controller, based on a single atomic controller and estimator for each neighbourhood, also stabilises any $p_* \in U$. Such a controller is typically realisable.

The selection of $\mathcal{P}_i$ and the determination of ϵ thus address the first two questions of Anderson et al. (2000) and Fekri et al. (2006). It remains to determine how the number of plant models influences the performance, and how to prevent a conservative design.

To investigate these question we will utilise two fundamental scaling geometries. Let the parameter bound $l > 0,\ l \in \mathbb{R}$ and the parameter discretisation step $m > 0,\ m \in \mathbb{R}$ define the set

$$\mathcal{P}_{l,m} = \left\{(i, 1, 1) \in \mathbb{R}^3 \mid i = \pm am,\ a \in \mathbb{N},\ |i| \leq l\right\},$$

where $\mathcal{P}_{l,m}$ parametrises the plant model $P_p,\ p \in \mathcal{P}_{l,m}$ given by

$$P_{(a,b,c)} : x_p(k+1) = ax_p(k) + bu_1^p(k),\ y_1^p(k) = cx_p(k),\ x_p(-k) = 0,\ \forall k \in \mathbb{N}.$$

Consider Figure 1.17. We now ask the following scaling questions and note that scaling is performed off-line — we are interested how the algorithm behaves for different scaling scenarios and use fixed plant model sets for the argument.

How does the algorithm perform in the case of

- Refinement scaling:
  The number of (unique) plant models is increasingly large however the plant model set is bounded, i.e. fix $l < \infty$ where $m > 0$. This leads to a dense plant model set.
- Expansion scaling:
  The number of plant models is increasingly large however the distance between them is kept constant, i.e. fix $m > 0$ where $l > 0$. This leads to a large, sparse plant model set.

These scaling scenarios are motivated by the two possibilities that: either the uncertainty U is bounded and we might want to have a large number of plant models that are close in the hope to increase the performance (on the other hand the increased number of candidate plants might degrade e.g. the transient performance), or U is overly large and we will have to introduce a large number of distinct plant models in order to provide a stabilising controller (this may lead to conservatism).

Since we will establish a global measure of performance and robustness $\epsilon = b_{P,C} = \hat{\gamma}^{-1}$ in Chapter 5 that we can optimise for and which reflects the geometric trade-offs in choosing all key variables, we will be able to give explicit answers to these scaling questions. A technique that plays an important role in this respect is that operations on the plant model set may be performed on-line, which we denote dynamic EMMSAC. Analogously, an EMMSAC algorithm based on a constant plant model set is said to be static. ; The results established in Chapter 6 show the following:

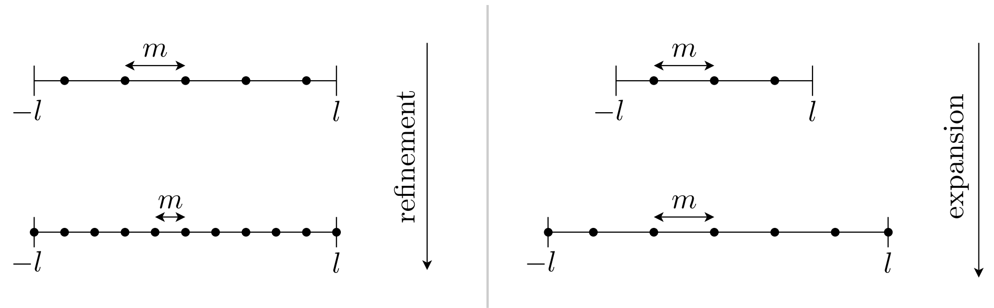

FIGURE 1.17: Refinement and expansion scaling by $\mathcal{P}_{l,m}$

- Above a certain critical refinement level an EMMSAC design becomes stabilising and a common bound $\hat{\gamma}$ for the true gain γ is given for all higher levels of refinement, i.e. performance does not diverge under refinement scaling. This opens the door to on-line refinement schemes that start off with a plant model set refined beyond the critical stabilising refinement level, and then introduces plant models corresponding to regions where the algorithm is expecting the true plant, based on the observation of closed loop signals.
- The actual closed loop gain γ for static EMMSAC is conservative in the expansion geometry, i.e. $\gamma \to \infty$ as $l \to \infty$. To address this issue we introduce a dynamic EMMSAC scheme that follows the strategy to expand the plant model set on-line until the performance is satisfactory. This technique allows the construction of a constant gain (function) bound that is invariant to the level of uncertainty. Hence the algorithm is universal.

This addresses the third and fourth question, and gives the insight required to provide a systematic approach to design addressing the first to questions.

6 Chapter Organisation

In Chapter 2 we will introduce necessary notation and give a brief introduction to signals and systems, uncertainty descriptions and (modern) robust stability analysis. This chapter is not intended to give an exhaustive study of such topics but only to supply the necessary machinery for the arguments that follow.

Chapter 3 formally introduces disturbance estimation. Two exemplar constructions of disturbance estimators are presented, followed by the introduction of five axiomatic requirements on the estimator. All that follows from there on will rest on these estimator assumptions, not on particular estimator realisations. All exemplar estimators are shown to meet the estimator assumptions. The equivalence between infinite horizon optimal estimation in l_2 and the Kalman filter is established. Continuity properties of the estimators are discussed.

Chapter 4 describes the actual EMMSAC algorithm with all involved sub-components. It starts by establishing two controller assumptions which are then shown to be met by all minimal MIMO LTI plants and controllers. All that follows from there on will rest on these controller assumption, not on particular controller constructions. The plant-generating operator G is introduced and its role in dynamic EMMSAC is discussed with the help of examples.

Chapter 5 is technical and establishes the main result: a bound $\hat{\gamma}$ on the gain from the external disturbances w_0 to the internal signals w_2. The Chapter introduces the device of covers of the uncertainty set in order to express the gain bound in terms of the complexity of the candidate plant set (described by the size of the cover set) and hence achieves gain bounds which are independent of the size of the candidate plant set. The global gain bound is established for both algorithms based on continuums of plants and for sampled (and realisable) versions.

Chapter 6 relates the cover constructions of Chapter 5 to the concept of metric entropy and then asks two fundamental scaling questions: the behaviour of the gain bound if there is a refinement in parameter space, or an expansion. With the established gain bound in Chapter 5, which is invariant to the number of elements in the plant model set, the next main result is established, which, under some continuity assumptions, shows that the global gain bound is invariant to refinement scaling. Expansion scaling is investigated and it is shown for a fundamental example that the actual closed loop gain scales badly in this case. A dynamic EMMSAC extension is introduced that establishes a gain (function) bound which is invariant to expansion scaling. It is then investigated when dynamic EMMSAC constructions promise tighter gain bounds than static EMMSAC constructions and vice versa. The chapter is concluded with an example where it is shown qualitatively how to conduct performance-orientated design both for the static and dynamic version of EMMSAC.

In Chapter 7 conclusions are drawn and future directions of research are indicated.

Chapter 2

Preliminaries

In this chapter, we will establish the notation used in the remainder of the thesis and review the underlying mathematical and system-theoretic framework. We will first introduce the notion of a signal and an operator (which acts on a signal). Then we will discuss various important properties of closed loop systems comprising of an interconnection between two operators, such as well-posedness, stability and robustness and their relation to uncertainty modelling in the gap metric.

1 Norms and signals

In order to study physical systems analytically, physical variables — such as speed, current or pressure — have to be expressed in a systematic kind of way. These variables can be considered maps from time to value which we call a signal.

A signal can now be defined in discrete time, e.g. when a temperature is measured/sampled every $T \in \mathbb{R}^+$ seconds, or in continuous time when measured continuously without interruption. This naturally leads to the signal space

$$\mathcal{S} := map(\mathbb{T}, \mathbb{R}^h)$$

where $h \in \mathbb{N}$ and $\mathbb{T}$ can be the set of real numbers $\mathbb{R}$, the set of positive real numbers $\mathbb{R}^+$, the set of integers $\mathbb{Z}$ or the set of natural numbers $\mathbb{N}$.

Sometimes we may record values only over a finite window of time. The corresponding signal is then only defined on a subset $[a, b],\ a \leq b,\ a, b \in \mathbb{T}$ of $\mathbb{T}$ where

$$[a, b] = \{x \in \mathbb{T} \mid a \leq x \leq b\}.$$

In this case we write $\mathcal{S}|_{[a,b]} := map([a, b], \mathbb{R}^h)$ hence $\mathcal{S}|_{[a,b]}$ is the set of maps that are defined only on the interval $[a, b]$, $a, b \in \mathbb{T}$.

Let $\mathbb{T} \in \{\mathbb{N}, \mathbb{Z}\}$ and

$$\begin{aligned} x &= (x_a, x_{a+1}, \cdots, x_b) \in \mathcal{S}|_{[a,b]},\ a \le b \\ y &= (y_c, y_{c+1}, \cdots, y_d) \in \mathcal{S}|_{[c,d]},\ c \le d \\ z &= (z_e, z_{e+1}, \cdots, z_f) \in \mathcal{S}|_{[e,f]},\ e \le f. \end{aligned}$$

The concatenation of signals is then defined as

$$\begin{aligned} cat(x,y) &:= (x_a, x_{a+1}, \cdots x_b, y_c, y_{c+1}, \cdots y_d) \\ cat(x,y,z) &:= cat(cat(x,y),z) \in \mathcal{S}|_{[g,h]}. \end{aligned}$$

For notational simplicity we often write (x,y,z) for $cat(x,y,z)$.

We also consider signals that are defined over the whole horizon, however we are only interested in their initial portion. For that purpose introduce the truncation operator $\mathscr{T}_t : \mathcal{S} \cup_{b\in\mathbb{T}} \mathcal{S}|_{[0,b]} \to \mathcal{S},\ t \in \mathbb{T}$ defined by:

$$(\mathscr{T}_t v)(\tau) = \begin{cases} v(\tau) & \text{if} \quad 0 \le \tau \le t,\ t \in \mathbb{T} \\ 0 & \text{otherwise} \end{cases}.$$

This operator returns a signal that equals $v \in \mathcal{S} \cup_{b\in\mathbb{T}} \mathcal{S}|_{[0,b]}$ up to time $t \in \mathbb{T}$ and is zero everywhere else.

An important property of a signal is its 'size' where we will have to define a suitable measure to make explicit what we mean by size. For that purpose we equip the signal space with a norm $\|\cdot\| : \mathcal{S} \cup_{a\le b} \mathcal{S}|_{[a,b]} \to \mathbb{R}^+ \cup \{\infty\}$.

Definition 2.1. $\|\cdot\| : \mathcal{S} \cup_{a\le b} \mathcal{S}|_{[a,b]} \to \mathbb{R}^+ \cup \{\infty\}$ *is said to be a norm if for all* $v, w \in \mathcal{S}$ *and* $v, w \in \mathcal{S}|_{[a,b]},\ a \le b$*:*

- $v = 0 \Leftrightarrow \|v\| = 0$: positivity,
- $\|av\| = |a|\|v\|,\ a \in \mathbb{R}$: homogenity,
- $\|v + w\| \le \|v\| + \|w\|$: triangle inequality.

Important examples of norms are L_r and l_r, $1 \le r \le \infty$, since they are able to express many physically relevant properties of a signal, e.g. the energy or its largest value. They are defined as follows: for $a \in \mathcal{S}$ where $\mathbb{T} = \mathbb{N}, \mathbb{Z}$ define

$$\begin{aligned} \|a\|_r &= \left(\sum_{i\in\mathbb{T}} |a(i)|^r \right)^{1/r},\ 1 \le r < \infty \\ \|a\|_\infty &= \sup_{i\in\mathbb{T}} |a(i)| \end{aligned}$$

and where $\mathbb{T} = \mathbb{R}, \mathbb{R}^+$ define

$$\begin{aligned} \|a\|_r &= \left(\int_{\mathbb{T}} |a(t)|^r dt \right)^{1/r}, \ 1 \le r < \infty \\ \|a\|_\infty &= \operatorname*{esssup}_{t \in \mathbb{T}} |a(t)|. \end{aligned}$$

If $a \in \cup_{a \le b} \mathcal{S}|_{[a,b]}$, then the sums and suprema are only taken over the relevant interval $[a, b]$. Note that we will often write $\|\cdot\|$ for $\|\cdot\|_r$ if the statement holds for any $1 \le r \le \infty$.

Although our overall goal is to control a system such that all signals are bounded (in norm), we cannot assume signals to be bounded a priori and have to account for the possibly that signals are indeed unbounded (in norm). To be able to refer to such bounded and unbounded signals, define the two corresponding signal spaces $\mathcal{V}$ and $\mathcal{V}_e$: For $\mathcal{V} \subset \mathcal{S}$ let

$$\mathcal{V} := \{a \in \mathcal{S} \mid a(-t) = 0, \ \forall t \in \mathbb{T}; \ \|a\| < \infty\}$$

and note that $\mathcal{V}$ is a normed vector space including only norm bounded signals. In this thesis the signal spaces under consideration will usually be $\mathcal{V} = L_r$ for $\mathbb{T} = \mathbb{R}$ and $\mathcal{V} = l_r$ for $\mathbb{T} = \mathbb{Z}$.

Since $\mathcal{V}$ does not contain signals $v \in \mathcal{S}$ such that $\|v\| = \infty$, i.e. $\|v\| = \infty \ \Rightarrow \ v \notin \mathcal{V}$, we extend the signal space $\mathcal{V}$ by signals that are allowed to grow unboundedly in norm over an infinite horizon, i.e. $\|\mathscr{T}_t v\| \to \infty$ for $t \to \infty$. However, we require that for any finite $t < \infty$, $\|\mathscr{T}_t v\|$ is bounded. Consequently, define the extended space $\mathcal{V}_e$, $\mathcal{V} \subset \mathcal{V}_e \subset \mathcal{S}$ by

$$\mathcal{V}_e \ := \ \{v \in \mathcal{S} \mid \forall t \in \mathbb{T} : \mathscr{T}_t v \in \mathcal{V}\} \, .$$

In for example French (2008) and French et al. (2006), a further signal space called the ambient space $\mathcal{V}_a$ is introduced to account for the possibility of a finite escape time, i.e. $\exists t < \infty$ such that $\|\mathscr{T}_t v\| = \infty$. However, in this thesis we will restrict our attention to systems where this cannot occur, e.g. switched linear systems, and therefore all signals can be measured by a finite norm over a finite interval. Observe that then in the particular cases where $\mathcal{V} = L_r, l_r, \ 1 \le r \le \infty$:

$$\mathcal{V}_e = \mathcal{S}, \ 1 \le r \le \infty.$$

If $v \in \mathcal{V}$ then v is said to be bounded, and if $v \in \mathcal{V}_e \setminus \mathcal{V}$ then v is said to be unbounded.

2 Operators and the frequency domain

We now introduce the notion of operators which act on signals. Define for $m \in \mathbb{N}$, the dimension of the input space, and $o \in \mathbb{N}$, the dimension of the output space, the input and output signal spaces

$$\mathcal{U} := \underbrace{\mathcal{V} \times \cdots \times \mathcal{V}}_{m} = \mathcal{V}^m, \ \mathcal{Y} := \underbrace{\mathcal{V} \times \cdots \times \mathcal{V}}_{o} = \mathcal{V}^o.$$

Define $\mathcal{U}_e, \mathcal{Y}_e$ accordingly.

In general, an operator is an object that maps some input signal $u \in \mathcal{U}_e$ to an output signal $y \in \mathcal{Y}_e$. For example $H : \mathcal{U}_e \to \mathcal{Y}_e$ might represent the input/output relationship corresponding to a plant $y_1 = Hu_1$. An important property is the signal amplification or gain attached to an operator.

A reasonable definition of such a gain is given by the induced operator norm

$$\|H\| := \sup_{u \in \mathcal{V}_e, \ t \in \mathbb{T}, \ \|\mathscr{T}_t u\| \neq 0} \frac{\|\mathscr{T}_t H \mathscr{T}_t u\|}{\|\mathscr{T}_t u\|}$$

which measures the maximum achievable input-output amplification of the input/output operator H, where the size of the input and output is measured in the corresponding signal norm.

An important property of an operator is causality:

Definition 2.2. *An operator H is said to be causal if:*

$$\mathscr{T}_t H \mathscr{T}_t v = \mathscr{T}_t H v, \ \forall t \in \mathbb{T}, \ v \in \mathcal{S}.$$

Causality ensures that the output of the operator H up to time $t \in \mathbb{T}$ cannot depend on the values of the input after $t \in \mathbb{T}$. Note that all physical systems are causal; a non-causal operator cannot be physically implemented since the computation of the current output relies on future input values.

The notation up to this point has purely been developed in the time domain. Since in the literature LTI systems are usually analysed in the frequency domain and the corresponding transfer function notation is expected to be more familiar to the reader, we will continue to present examples in this language when appropriate.

In continuous time ($\mathbb{T} = \mathbb{R}$) a signal $v \in \mathcal{S}$ in the time domain is related to a signal $\tilde{v}$ in the frequency domain via the Laplace transform

$$\tilde{v}(s) = \int_0^\infty e^{-st} v(t) dt$$

and in discrete time ($\mathbb{T} = \mathbb{Z}$) a signal $v \in \mathcal{S}$ in the time domain is related to a signal $\tilde{v}$ in the frequency domain via the Z-transform

$$\tilde{v}(z) = \sum_{n=0}^{\infty} v(n) z^{-n}.$$

Time domain and frequency domain LTI operators are then related via the transforms of their impulse and frequency responses.

We now quote an important result relating time and frequency domain. Observe that the L_2 and l_2 norms measure the energy of a signal. Since we are concerned with the stability analysis of dynamical systems one could argue that a good stability requirement is that if a system is fed with an input signal of bounded energy, it responds with an output signal of bounded energy. Hence we would require that input and output have a finite $\|\cdot\|_2$ norm.

Denote the transfer function matrix $\tilde{H}$ the frequency domain representation of some linear operator H. Then:

Theorem 2.3. *Let H be a linear time invariant operator. If $H : L_2 \to L_2$ then:*

$$\|\tilde{H}(j\mathbb{R})\|_\infty = \sup_{0\leq\omega\leq\infty} \bar{\sigma}(\tilde{H}(j\omega)) = \|H\|_2 = \sup_{u\in L_2,\ u\neq 0} \frac{\|Hu\|_2}{\|u\|_2}, \tag{2.1}$$

and if $H : l_2 \to l_2$ then:

$$\|\tilde{H}(\partial\mathsf{D})\|_\infty = \sup_{s\in\mathbb{C}:|s|=1} \bar{\sigma}(\tilde{H}(s)) = \|H\|_2 = \sup_{u\in l_2,\ u\neq 0} \frac{\|Hu\|_2}{\|u\|_2} \tag{2.2}$$

where $\bar{\sigma}(\tilde{H})$ denotes the maximum singular value[1] *of $\tilde{H}$ and $\mathsf{D} = \{s \in \mathbb{C} \mid |s| < 1\}$.*

Definition 2.4. *We let $\mathcal{H}_\infty$ denote the space of all functions that are analytic and bounded:*

- *in the open right-half plane $\mathbb{C}^+$ in continuous time, with norm* (2.1).
- *outside the unit disk $s \in \mathbb{C}, |s| > 1$ in discrete time, with norm* (2.2).

To distinguish between the continuous and discrete cases we write $\mathcal{H}_\infty = \mathcal{H}(j\mathbb{R})$ and $\mathcal{H}_\infty(\partial\mathsf{D})$ for the two respective cases.

3 Closed loop system, well-posedness and stability

Given a plant

$$P : \mathcal{U}_e \to \mathcal{Y}_e \tag{2.3}$$

satisfying

$$P(0) = 0 \tag{2.4}$$

[1]The maximum singular value $\bar{\sigma}(\tilde{H})$ of $\tilde{H}$ is given by $\bar{\sigma}(\tilde{H}) = \sqrt{\bar{\lambda}(\tilde{H}^*\tilde{H})}$ where $\bar{\lambda}$ returns the largest eigenvalue and $\tilde{H}^*$ is the conjugate transpose of $\tilde{H}$. The conjugate transpose X^* of a matrix $X = [x_{ab}] \in \mathbb{C}^{p\times q}$ is defined as $[\bar{x}_{ba}]$ where $\bar{x}_{ba} = r - qi$ if $x_{ba} = r + qi$.

and a controller

$$C : \mathcal{Y}_e \to \mathcal{U}_e \tag{2.5}$$

satisfying

$$C(0) = 0 \tag{2.6}$$

the closed loop system $[P, C]$ under consideration in Figure 2.1 is defined via the following set of system equations:

$$y_1 = Pu_1 \tag{2.7}$$

$$u_0 = u_1 + u_2 \tag{2.8}$$

$$y_0 = y_1 + y_2 \tag{2.9}$$

and

$$u_2 = Cy_2. \tag{2.10}$$

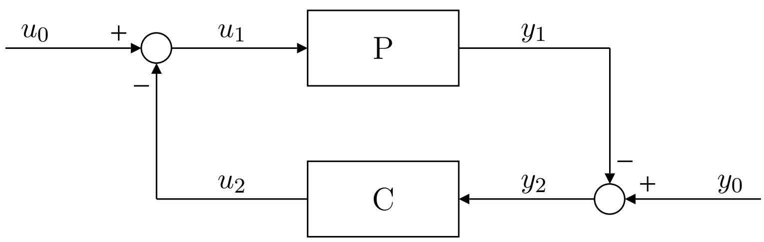

FIGURE 2.1: Closed loop $[P, C]$

For notational convenience we often write $\mathcal{W} = \mathcal{U} \times \mathcal{Y}$ and $\mathcal{W}_e = \mathcal{U}_e \times \mathcal{Y}_e$ where $w_0 = (u_0, y_0)^\top \in \mathcal{W}$ represents the input and output disturbances acting on the plant P, $w_1 = (u_1, y_1)^\top \in \mathcal{W}_e$ represents the plant's input and output and $w_2 = (u_2, y_2)^\top \in \mathcal{W}_e$ represents the observed signal or observation.

Our main concern in control theory is to study the stability, robustness and performance of such closed loop systems $[P, C]$. However, this is only feasible if the closed loop system satisfies essential properties that allow its analysis. In particular we require it to be well-posed:

Definition 2.5. *A closed loop system* $[P, C]$ *given by equations* (2.7)–(2.10) *is said to be well-posed if for all* $w_0 \in \mathcal{W}$ *there exists a unique solution* $(w_1, w_2) \in \mathcal{W}_e \times \mathcal{W}_e$.

We now verify this property for linear switched systems, which we are mainly concerned with in this thesis.

Definition 2.6. *Let $m, o, n \geq 1$. Let $w_0 \in map(\mathbb{T}, \mathbb{R}^m)$, $(w_1, w_2) \in map(\mathbb{T}, \mathbb{R}^o)$. A system $w_0 \mapsto (w_1, w_2)$ is said to be a linear switched system if there exists a decomposition $\mathbb{T} = \cup_{i\in\mathbb{N}}[t_i, t_{i+1}),\ t_i < t_{i+1},\ t_0 = 0$ such that: for all $i \in \mathbb{N}$, there exists $x_0 \in \mathbb{R}^n, A_i \in \mathbb{R}^{n\times n}, B_i \in \mathbb{R}^{n\times m}, C_i \in \mathbb{R}^{o\times n}, D_i \in \mathbb{R}^{o\times m}$ such that the equations*

$$x(t+1) = A_i x(t) + B_i w_0(t),\ x(t_i) = V(w_0, w_1, w_2)(t_i) \tag{2.11}$$

$$(w_1, w_2)(t) = C_i x(t) + D_i w_0(t) \tag{2.12}$$

have a unique solution $x(t),\ t \in [t_i, t_{i+1})$ where $V : \mathcal{W} \times \mathcal{W}_e \times \mathcal{W}_e \to \mathbb{R}^n$ is bounded and causal.

Lemma 2.7. *Linear switched systems are well-posed.*

Proof By induction on $i \in \mathbb{N}$, we assume w_1, w_2 are uniquely defined up to time t_i. Since V is bounded and causal, $x(t_i) \in \mathbb{R}^n$ is defined and for all $t \in [t_i, t_{i+1}],\ i \in \mathbb{N}$ equations (2.11),(2.12) describe a LTI system with some initial condition, which is known to be well-posed. Hence for bounded input signals $w_0 \in \mathcal{W}$ there exists a unique solution $(w_1, w_2) \in \mathcal{W}_e \times \mathcal{W}_e$ up to time t_{i+1}. The base step $i = 0$ holds trivially here, hence the linear switched system is well-posed as required. □

However, note that for non-linear systems well-posedness is not implicit and we have to take further measures to ensure it. Well-posedness is important since the potential non-existence of solutions, which arises e.g. if a system has a finite escape time, would require a rather different analysis over small windows of time where the system is ensured to have a solution. Furthermore, the non-uniqueness of solutions would be problematic since then the analysis would have to account for all (possibly infinitely many) of them.

Given a closed loop system $[P, C]$ which is structured as in Figure 2.1, a good measure of stability and performance is the amplification or the gain from the external disturbances w_0 to the internal signals w_1, w_2.

The following notation and results follow from Georgiou and Smith (1990, 1997). Let

$$H_{P,C} : \mathcal{W} \to \mathcal{W}_e \times \mathcal{W}_e : w_0 \mapsto (w_1, w_2)$$

denote the closed loop operator mapping the external disturbances $w_0 \in \mathcal{W}$ to the unique internal closed loop signals $w_1, w_2 \in \mathcal{W}_e$. Observe that the closed-loop operator $H_{P,C}$ can be decomposed into the operator $\Pi_{P//C}$ (which is the map from the disturbances $w_0 \in \mathcal{W}$ to the plant signals $w_1 \in \mathcal{W}_e$) and the operator $\Pi_{C//P}$ (which is the map from the disturbances $w_0 \in \mathcal{W}$ to the controller signals $w_2 \in \mathcal{W}_e$), i.e.

$$\begin{aligned} \Pi_{P//C} &: \mathcal{W} \to \mathcal{W}_e : w_0 \mapsto w_1, \\ \Pi_{C//P} &: \mathcal{W} \to \mathcal{W}_e : w_0 \mapsto w_2 \end{aligned}$$

where

$$H_{P,C} = (\Pi_{C//P}, \Pi_{P//C}).$$

We now define the notion of gain stability via the gain of $\Pi_{P//C}$ in the following way:

Definition 2.8. *Let the closed loop system $[P, C]$, defined by equations (2.7)–(2.10), be well-posed. $[P, C]$ is said to be gain stable if there exists a $M > 0$ such that:*

$$\sup_{w_0 \in \mathcal{W},\ w_0 \neq 0} \frac{\|\Pi_{P//C} w_0\|}{\|w_0\|} = \|\Pi_{P//C}\| < M < \infty.$$

Note that since for all $w_0 \in \mathcal{W}$ we have:

$$\Pi_{P//C} w_0 + \Pi_{C//P} w_0 = w_1 + w_2 = w_0,$$

it follows that

$$\Pi_{P//C} + \Pi_{C//P} = I.$$

Hence gain stability of $\Pi_{P//C}$ also ensures gain stability for $\Pi_{C//P}$ and $H_{P,C}$. We will therefore refer to $\Pi_{P//C}, \Pi_{C//P}$ and $H_{P,C}$ as the closed loop operator.

Sometimes this measure of stability is too strong, i.e. in a general (non-linear) setting the signal amplification from $(u_0, y_0)^\top$ to $(u_2, y_2)^\top$ might not be a linear gain. For that purpose we define the gain function $\gamma : \mathbb{R}^+ \to \mathbb{R}^+$ by

$$\gamma(r) = \sup\{\|\Pi_{P//C} w_0\| : \|w_0\| \leq r\},$$

and hence measure the maximum size of the internal signal $\|w_2\| = \|\Pi_{P//C} w_0\|$, given a disturbance w_0 of size $\|w_0\| \leq r \in \mathbb{R}$.

Definition 2.9. *Let the closed loop system $[P, C]$, defined by equations (2.7)–(2.10), be well-posed. $[P, C]$ is said to be gain function stable if for all $r > 0$ there exist $M_r > 0$ such that:*

$$\gamma(r) < M_r < \infty.$$

This definition is rather useful since especially universal adaptive control schemes do not appear to be gain stable — however, they can be gain function stable. For example, French (2008) has shown that continuously tuned adaptive systems are gain function stable. We will later show that the universal variants of the algorithms considered in this thesis will also show gain function stability.

Nominal stability, however, is not enough to ensure that a control algorithm works well in practice. We must show that it remains stable even if there is a certain amount of uncertainty in the plant, e.g. unmodeled dynamics. Stability in the presence of disturbances and uncertainty leads to the notion of robustness which we will discuss next.

4 Uncertainty and Robustness

The analysis of systems is preferentially performed on simplified nominal models (derived either empirically or analytically) rather than on complete mathematical descriptions of the underlying physical system. This is due to the following facts:

1. Complete knowledge of a physical system is unrealistic:
   In particular it is impossible to model the high frequency dynamics of physical systems accurately. To see this consider the classical mass-spring-damper arrange-

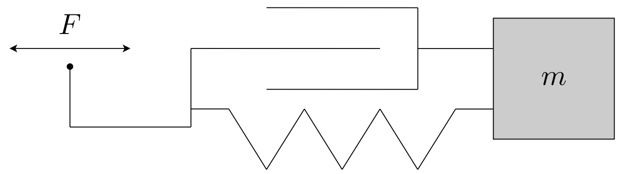

FIGURE 2.2: Mass spring damper arrangement with force F and mass m

   ment depicted in Figure 2.2, where an oscillating force F is acting. At low frequencies the behaviour of the system is dominated by the ideal equations of motion we all got to know in high school physics since sub-components of the arrangement with higher natural frequencies have little influence. However, at higher frequencies, for example close to the natural frequency of the material of the spring wire, the dynamic behaviour will be dominated by the complex dynamical behaviour of the spring wire itself. This makes the high frequency part incredibly difficult to model. Taking this example even further, it would in theory require a modeling effort on a sub-atomic level and beyond to obtain a completely accurate model of the system. In practice this high frequency part can fortunately be neglected for appropriately designed control systems.

2. Less detail promises simplicity:
   Even if detailed knowledge of a physical system is available, usually only a small subset of this information is required to design an appropriate controller. Hence in practice we intentionally neglect (dynamical) components that are irrelevant to the control objective in order to simplify analysis and design.

After obtaining an appropriate nominal model P for the physical plant P_{p_1} we then, by the above discussion, perform the stability analysis on P, i.e. we show that for some controller the closed loop $[P, C]$ is stable. However, we then need to show that C also stabilises the true plant P_{p_1} since this is the physical system we will have to control in practice. For that purpose we introduce so-called uncertainty models that quantify the mismatch between the physical plant P_{p_1} and the model P. The overall goal of a later robustness analysis is then to quantify how much uncertainty or mismatch the closed loop system $[P, C]$ can tolerate without becoming unstable or the performance degrading too far.

We differentiate between structured and unstructured uncertainty models. For example consider the plant given by the real rational transfer function

$$P_{p_1} = \frac{1}{s+a}$$

where $a_{min} \leq a \leq a_{max}$ is an uncertain parameter. We can therefore say that P_{p_1} lies in the uncertainty set

$$\left\{ \frac{1}{s+a} : a_{min} \leq a \leq a_{max} \right\},$$

which is structured since the possible uncertainty is dictated by the structure of $\frac{1}{s+a}$.

Although the simplicity of structured uncertainty models has virtue, they are inherently unable to express (dynamical) uncertainty outside of the defining structure. This is unfortunate since by the above argument we should always account for a certain amount of unstructured dynamics.

Consider the following example inspired by Doyle et al. (1990). Let

$$P_{p_1} = e^{-\tau s} \frac{1}{s-1}$$

where $0 \leq \tau \leq 0.01$ is an uncertain time delay. A silo filled by a short conveyor belt could have such a transfer function where the input is the flow into the silo and the output is the volume of material inside the silo. Since the properties of finite dimensional LTI systems are very well understood and the corresponding theory is much simpler than the infinite dimensional counterpart, we would usually like to simplify the plant P_{p_1} to a finite dimensional P and work with P instead. One possibility is to simplify P_{p_1} to $P = \frac{1}{s-1}$ and hence neglect the small time delay $e^{-\tau s}$. To describe the mismatch between P_{p_1} and its approximation P we now employ a multiplicative uncertainty model in the following way:

Let Δ_m be a stable transfer function. The multiplicative uncertainty set is then given by

$$\{(1 + \Delta_m)P : \ \|\Delta_m\| < r\}.$$

If we rearrange $P_{p_1} = (1 + \Delta_m)P$ to

$$\left\| \frac{P_{p_1}}{P} - 1 \right\| = \|\Delta_m\| < r$$

we can see that this uncertainty set describes a disk with centre 1, radius r in the complex plane. For the concrete example of $P_{p_1} = e^{-\tau s}\frac{1}{s-1}$ and $P = \frac{1}{s-1}$ we therefore have

$$\left\| \frac{P_{p_1}}{P} - 1 \right\| = \left\| e^{-\tau s} - 1 \right\| < r.$$

Since $e^{-\tau s}$ describes the unit circle in the complex plane we have that for sufficiently large r, $P_{p_1} \in \{(1+\Delta_m)P : \ \|\Delta_m\| < r\}$.

Now consider the transfer function

$$P_{p_1} = \frac{1001}{(s-1)(s+1000)}.$$

The filling process of a silo could have such a transfer function where the input is the flow into the silo and the output is the volume of material inside the silo; however the contents are setting over time which reduces the volume. As before we would like to simplify P_{p_1} to $P = \frac{1}{s-1}$ and neglect the dynamics due to the setting process. We now describe the mismatch between P_{p_1} and P with an additive uncertainty model:

Let Δ_a be a stable transfer function. The additive uncertainty set is then given by

$$\{P + \Delta_a : \ \|\Delta_a\| < r\}.$$

Since we can write P_{p_1} as

$$P_{p_1} = \frac{1001}{(s-1)(s+1000)} = \frac{(s+1000)-(s-1)}{(s-1)(s+1000)} = \frac{1}{s-1} - \frac{1}{s+1000},$$

we can see that for sufficiently large r, $P_{p_1} \in \{P + \Delta_m : \ \|\Delta_a\| < r\}$.

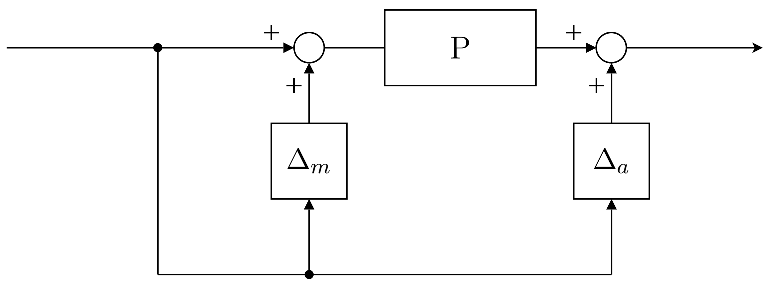

FIGURE 2.3: Additive and multiplicative uncertainty model

The combined additive and multiplicative uncertainty model is now depicted in Figure 2.3 where the corresponding uncertainty set is given by

$$\{(1+\Delta_m)P + \Delta_a : \ \|\Delta_m\| < r_1, \|\Delta_a\| < r_2\}, \ r_1, r_2 > 0.$$

The full model of the silo incorporating both the delay and the drying dynamics is given by

$$P_{p_1} = e^{-\tau s}\frac{1001}{(s-1)(s+1000)}.$$

We conclude that for sufficiently large $r_1, r_2 > 0$:

$$P_{p_1} \in \{(1+\Delta_m)P + \Delta_a : \ \|\Delta_m\| < r_1, \|\Delta_a\| < r_2\}, \ r_1, r_2 > 0.$$

As discussed in the introduction, classical robust adaptive control (Ioannou and Sun (1996)) as well as multiple model adaptive control in the sense of Morse et. all. is confined to this class of additive and multiplicative uncertainties. However, there are a number of important uncertainty scenarios which this uncertainty model is unable to represent such as neglected high frequency dynamics, low frequency parameter errors and especially neglected right half poles. They give rise to uncertainty models such as inverse additive, inverse multiplicative and others. See Doyle et al. (1982) for a comprehensive discussion of uncertainty scenarios and appropriate uncertainty models. One can then construct a linear robustness theory around such uncertainty models and impose constraints on the different Δs to ensure robust stability and performance of the closed loop system, e.g. see Doyle et al. (1990) and the references therein. Albeit being a viable strategy, employing a number of such uncertainty models and dealing with them in the analysis turns out to be rather cumbersome.

Zames and El-Sakkary (1980) introduced a more generic measure of uncertainty, the gap metric, which essentially fuses all the above uncertainty scenarios and represents the global uncertainty by a single scalar. We will now show how the linear gap metric is constructed.

Let X^* be the conjugate transpose to a matrix X. Let $\mathcal{R}$ denote the space of all real rational[2] transfer matrix functions where we write $\mathcal{RH}_\infty = \mathcal{R} \cap \mathcal{H}_\infty$. Let $P \in \mathcal{R}$. A normalised right coprime factorisation (NRCF), (N, M) for P satisfies:

$$P = NM^{-1}$$

and

$$M^*M + N^*N = 1 \qquad \text{(Bezout identity)}$$

where

$$M^*, M, N^*, N \in \mathcal{RH}_\infty.$$

With (M_i, N_i) being a NRCF for $P_{p_i},\ i \in \{1,2\}$, hence $P_{p_i} = \frac{N_i}{M_i},\ i \in \{1,2\}$, the linear directed gap distance

$$\vec{\delta} : \mathcal{R} \times \mathcal{R} \to \mathbb{R}^+$$

is given by

$$\vec{\delta}(p_1, p_2) \ = \ \inf_{(\Delta N, \Delta M) \in \mathcal{RH}_\infty} \left\{ \left\| \begin{matrix} \Delta N \\ \Delta M \end{matrix} \right\|_{\mathcal{H}_\infty} \ \middle| \ \begin{pmatrix} \Delta N \\ \Delta M \end{pmatrix} \in \mathcal{RH}_\infty, P_{p_2} = \frac{N_1 + \Delta N}{M_1 + \Delta M} \right\}. \tag{2.13}$$

Note that in general

$$\vec{\delta}(p_1, p_2) \neq \vec{\delta}(p_2, p_1),$$

[2] We say that a transfer function is real rational if it can be written as a ratio of polynomials in $s \in \mathbb{C}$ with real coefficients.

hence $\vec{\delta}$ is not a metric. We therefore let

$$\delta(p_1, p_2) = \max\{\vec{\delta}(p_1, p_2), \vec{\delta}(p_2, p_1)\}$$

then $\delta(p_1, p_2) = \delta(p_2, p_1)$ and δ is a metric (see El-Sakkary (1985)). The linear gap metric measures the size of the smallest coprime perturbation that is required to perturb P_{p_1} to P_{p_2} — as depicted in Figure 2.4. It can be argued (see Vinnicombe (2000)) that this measure incorporates all the features captured by the standard unstructured uncertainty models (additive/multiplicative/inverse multiplicative, etc.).

A further advantage of the gap metric is that "perturbations which are small in the gap are precisely those which give small closed-loop errors" (see the introduction of Georgiou and Smith (1997)). Hence the gap metric allows the comparison of stable and unstable plants where for example $\delta(\frac{1}{s-\epsilon}, \frac{1}{s+\epsilon})$ is small if ϵ is small. In contrast, additive and multiplicative uncertainty models do not allow such a comparison since in that case Δ_m or Δ_a would have to be unstable.

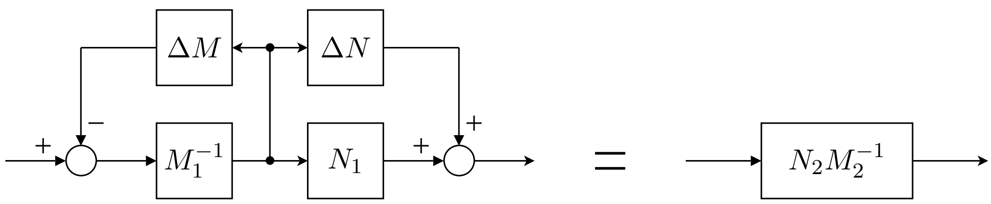

FIGURE 2.4: Coprime perturbation model

It is important to note that by Georgiou (1988) the L_2 directed linear gap in continuous time is equivalently given by:

$$\vec{\delta}(p_1, p_2) = \inf_{Q \in \mathcal{H}_\infty(j\mathbb{R})} \left\| \begin{pmatrix} M_1 \\ N_1 \end{pmatrix} - \begin{pmatrix} M_2 \\ N_2 \end{pmatrix} Q \right\|_\infty .$$

By Cantoni and Glover (1998) the l_2 directed linear gap in discrete time can also be expressed as:

$$\vec{\delta}(p_1, p_2) = \inf_{Q \in \mathcal{H}_\infty(\partial \mathsf{D})} \left\| \begin{pmatrix} M_1 \\ N_1 \end{pmatrix} - \begin{pmatrix} M_2 \\ N_2 \end{pmatrix} Q \right\|_\infty .$$

Hence the calculation of the gap reduces to a standard $\mathcal{H}_\infty$ optimisation problem and is therefore computationally tractable.

These key observations now allows, with the definition of $\|\Pi_{P//C}\|$ or the (gain) effect of disturbances on the internal signals, the construction of a major robustness result.

Let $\mathcal{V}$ be L_2 or l_2 and

$$b_{P,C} = \begin{cases} \frac{1}{\|\Pi_{P//C}\|_2} & \text{if } \|\Pi_{P//C}\| > 0 \\ 0 & \text{otherwise} \end{cases}$$

denote the robust stability margin of $[P, C]$ and note that for $\|\Pi_{P//C}\| \to \infty$, $b_{P,C} \to 0$.

Theorem 2.10. *Let $P_{p_1}, P_{p_2}, C \in \mathcal{R}$ and assume the closed loop $[P_{p_i}, C]$, $i = \{1, 2\}$ to be well-posed. Let the closed loop $[P_{p_1}, C]$ be gain stable. If*

$$\vec{\delta}(p_1, p_2) < \frac{1}{\|\Pi_{P_{p_1}//C}\|_2} = b_{P_{p_1},C}$$

then the closed loop system $[P_{p_2}, C]$ is gain stable. □

Proof The proof can be found in Georgiou and Smith (1990) which is based on Zames and El-Sakkary (1980). □

The above result is the basis for the analysis performed in this thesis. In particular it shows that robustness can be established purely by considering the nominal system: one has to show that $\|\Pi_{P//C}\|$ is finite and robustness follows. Theorem 2.10 is only valid in the linear domain and since we are dealing with switched linear systems in this thesis, which are inherently non-linear, this result does not apply directly. We would also like to perform analysis in other signal spaces than L_2, l_2.

However, Theorem 2.10 can be generalised via the following construction:

Definition 2.11. *The graph $\mathcal{M}_p$ of $P_p, p \in \mathcal{P}$ is defined by:*

$$\mathcal{M}_p = \left\{ v \in \mathcal{W} \;\middle|\; \begin{array}{c} \exists (u_1^p, y_1^p)^\top \in \mathcal{W} \text{ s.t. } P_p u_1^p = y_1^p, \\ v = (u_1^p, y_1^p)^\top \end{array} \right\} \subset \mathcal{W}.$$

A signal $w \in \mathcal{W}$ is said to be in the graph of P_p if $w \in \mathcal{M}_p$.

Note that the graph $\mathcal{M}_p$ is the collection of bounded pairs $(u_1^p, y_1^p)^\top \in \mathcal{W}$ compatible with the plant $P_p,\ p \in \mathcal{P}$.

Define the possibly empty set of maps between the graphs of P_{p_1} and $P_{p_2}, p_1, p_2 \in \mathcal{P}$

$$\mathcal{O}_{p_1,p_2} := \{\Phi : \mathcal{M}_{p_1} \to \mathcal{M}_{p_2} \mid \Phi \text{ is causal, bijective, and } \Phi(0) = 0\}.$$

Now define the non-linear directed gap by

$$\vec{\delta}(p_1, p_2) := \begin{cases} \inf_{\Phi \in \mathcal{O}_{p_1,p_2}} \sup_{x \in \mathcal{M}_{p_1} \backslash 0,\ k>0} \left(\frac{\|\mathcal{T}_k (\Phi - I)|_{\mathcal{M}_{p_1}} \mathcal{T}_k x\|}{\|\mathcal{T}_k x\|} \right) & \text{if } \mathcal{O}_{p_1,p_2} \neq \emptyset \\ \infty & \text{if } \mathcal{O}_{p_1,p_2} = \emptyset \end{cases}$$

As before, we symmetrise this relation to give the non-linear gap metric

$$\delta : \mathcal{P} \times \mathcal{P} \to [0, \infty]$$

with

$$\delta(p_1, p_2) = \max\{\vec{\delta}(p_1, p_2), \vec{\delta}(p_2, p_1)\}.$$

Note that for $\delta(p_1, p_2) < 1,\ p_1, p_2 \in \mathcal{P}$ this definition can be shown to equal the definition in equation (2.13) for linear systems in L_2, l_2, e.g. see the Appendix of Georgiou and Smith (1997).

Theorem 2.12. *Let* $\mathcal{U} = \mathcal{Y} = l_r,\ 1 \leq r \leq \infty$. *Let* $P_{p_1}, P_{p_2} \in map(\mathcal{U}_e, \mathcal{Y}_e),\ C \in map(\mathcal{Y}_e, \mathcal{U}_e)$ *and assume the closed loop* $[P_{p_i}, C],\ i = \{1, 2\}$ *to be well-posed. Let the closed loop* $[P_{p_1}, C]$ *be gain stable. If*

$$\vec{\delta}(p_1, p_2) < \frac{1}{\|\Pi_{P_{p_1}//C}\|} = b_{P_{p_1},C}$$

then the closed loop system $[P_{p_2}, C]$ *is gain stable and*

$$\|\Pi_{P_{p_2}//C}\| \leq \|\Pi_{P_{p_1}//C}\| \frac{1 + \vec{\delta}(p_1, p_2)}{1 - \|\Pi_{P_{p_1}//C}\| \vec{\delta}(p_1, p_2)}.$$

Proof The proof can be found in Georgiou and Smith (1997). □

5 Finite horizon analysis

Since we are concerned with a switched system we will have to deal with signals that are defined only on a finite intervals of time between switches. This motivates the following finite horizon treatment of signals and operators.

The restriction operator $\mathscr{R}_{\sigma,t} : \mathcal{S} \to \mathbb{R}^{h(\sigma+1)}$ has the purpose to extract only a finite window of length $\sigma \geq 0$ of a signal, i.e. for $\sigma, t \in \mathbb{T}$ define

$$\mathscr{R}_{\sigma,t} v := (v(t - \sigma), \cdots, v(t)),\ v \in map(\mathbb{T}, \mathbb{R}^h).$$

Hence $\mathscr{R}_{\sigma,t} v$ returns a signal that equals $v \in \mathcal{S}$ over some finite interval of length σ and is undefined everywhere else.

Although we intend to present the analysis in this thesis in an ‘operators act on signals’ kind of way, and we will do so wherever possible, in some cases this is impractical and we adopt the following alternative notation: For $0 \leq a \leq b,\ a, b \in \mathbb{T}$ let

$$\begin{aligned} [a, b] &:= \{x \in \mathbb{T} \mid a \leq x \leq b\} \\ [a, b) &:= \{x \in \mathbb{T} \mid a \leq x < b\} \end{aligned}$$

noting that $[a, a] := \{a\}$.

Let the size of the given intervals $|\cdot|$ be defined by

$$\begin{aligned} |[a,b]| &:= b-a+1 \\ |[a,b)| &:= b-a. \end{aligned}$$

For a signal $v \in \mathcal{S}$ we then define the restriction of v over the interval $I = [c,d]$ by

$$v|_I := (v(c), \cdots, v(d))$$

where $c \leq d,\ c,d \in \mathbb{T}$, and similarly for $I = [c,d)$.

Note that for $v \in \mathcal{S},\ a \leq b$ we have $v|_{[a,b]} = \mathscr{R}_{b-a,b} v$ hence $\|v|_{[a,b]}\| = \|\mathscr{R}_{b-a,b} v\|$. Also $\|v|_{[0,b]}\| = \|\mathscr{T}_b v\|$, however $v|_{[0,b]} \neq \mathscr{T}_b v$ because the domains differ.

6 Projections and disturbance estimation

The problem of disturbance estimation, i.e. to find the smallest disturbances that are compatible with a plant $P_p,\ p \in \mathcal{P}$ and the observation w_2, can be understood as a (metric) projection onto a particular (linear) vector space (also see Figure 1.15).

Definition 2.13. *Let $\emptyset \neq Y \subset X$ be finite dimensional normed vector spaces. $\Pi_Y : X \to Y$ is said to be a projection if for all $x \in X$*

$$\Pi_Y x \in \{n \in Y \mid \|n - x\| \leq \|m - x\|,\ \forall m \in Y\}.$$

Definition 2.14. *A subset X of a normed vector space Y is said to be open if for all $x \in X$ there exists $\epsilon > 0$ such that $B(x,\epsilon) \subset X$ where $B(x,\epsilon) = \{y \in Y \mid \|x - y\| < \epsilon\}$ defines an open neighbourhood of radius ϵ around x. A set X is said to be closed if the complement $X^c = Y \setminus X$ is open.*

Definition 2.15. *A vector space X is said to be convex if $x_1, x_2 \in X$ implies that $(1-t)x_1 + tx_2 \subseteq X$ for all $0 \leq t \leq 1$.*

Lemma 2.16. *Suppose $\emptyset \neq Y \subset X$, Y is closed and convex and X is a linear subspace of $l_r,\ 1 < r < \infty$. Then there exists a unique projection $\Pi_Y : X \to Y$.*

Proof A sufficient condition for uniqueness is that the norm is strictly convex (see Boyd and Vandenberghe (2004), Chapter 8.1). The $L_r, l_r,\ 1 < r < \infty$ norm can be shown to have this property. □

Note that if $\Pi_Y : X \to Y$ is unique we can write

$$\Pi_Y x = \operatorname*{argmin}_{n \in Y} \|n - x\|.$$

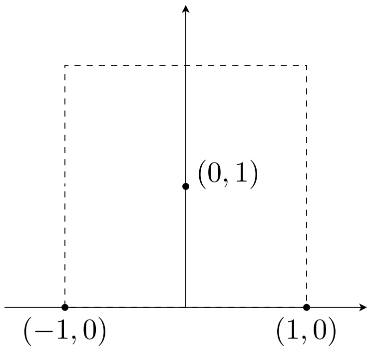

FIGURE 2.5: Unit ball around $(0, 1)$ in L_∞

As the restrictions in Lemma 2.16 suggest, such best approximations are not necessarily unique. For example if

$$Y = \{(\lambda, 0) \mid \lambda \in \mathbb{R}\} \in \mathbb{R}^2,\ X = \mathbb{R}^2$$

with the L_∞ norm, then the point $(0, 1)$ is a distance of 1 away from all the points $\{(\lambda, 0) \mid -1 < \lambda < 1\}$ since the unit ball around $(0, 1)$ forms a square — as depicted in Figure 2.5. A similar argument holds for L_1 where the unit ball is a tilted square.

However, we can guarantee the existence of projections in any $L_r, l_r,\ 1 \le r \le \infty$ norm.

Lemma 2.17. *Suppose $\emptyset \neq Y \subset X,\ Y$ is closed and convex and X is a linear subspace of $L_r, l_r,\ 1 \le r \le \infty$. Then there exists a projection $\Pi_Y : X \to Y$.*

Proof See for example Boyd and Vandenberghe (2004), Chapter 8.1.1 as required. □

The disturbance estimation algorithms considered in this thesis will utilise the distance $\|x - \Pi_Y x\|$ rather than Π_Y itself, hence the existence without uniqueness is sufficient for our purposes.

Chapter 3

Disturbance estimation

In this Chapter we introduce the disturbance estimator as motivated in the introduction. After discussing the basic estimator structure and some key examples of so-called finite and infinite horizon estimators, we will state five abstract estimator assumptions on which the subsequent analysis will rest. In addition to generality, the strong advantage of this axiomatic approach is that we separate the problem of realising (efficient) disturbance estimation from the problem of robust stability analysis of the closed loop system. Our study of disturbance estimation algorithms will not be exhaustive, however the finite and infinite horizon estimator will be shown to satisfy the estimator assumptions. We will also illustrate how these algorithms can be implemented in practice, including the relation to the Kalman filter.

1 The disturbance estimation principle

The purpose of the disturbance estimator is at each time step to assign a positive scalar to each candidate plant, termed the residual, which has the interpretation of being a measure of the size of the disturbance signals $w_0^p = (u_0^p, y_0^p)^\top$ required to 'explain' the observation $w_2 = (u_2, y_2)^\top$ in a manner consistent with the candidate plant P_p — as depicted in Figure 3.1.

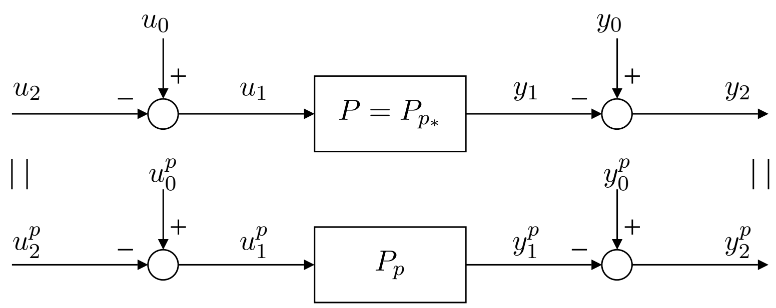

FIGURE 3.1: Disturbances and consistency with the observation

Let $\mathcal{P}$ be a set, parametrising a collection of plant operators

$$P_p : \mathcal{U}_e \to \mathcal{Y}_e : u_1^p \mapsto y_1^p,\ p \in \mathcal{P}. \tag{3.1}$$

For example, in the case of linear systems, we let $\mathcal{P} = \mathcal{P}_{LTI}$ where

$$\mathcal{P}_{LTI} :=$$
$$\left\{ (A,B,C,D) \in \cup_{n\geq 1}\mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} \times \mathbb{R}^{o\times n} \times \mathbb{R}^{o\times m} \;\middle|\; \begin{array}{ll} (A,B) & \text{is controllable} \\ (A,C) & \text{is observable} \end{array} \right\} \tag{3.2}$$

and

$$P_p : \mathcal{U}_e \to \mathcal{Y}_e,\ p = (A_p, B_p, C_p, D_p) \in \mathcal{P}_{LTI} \tag{3.3}$$

is defined by

$$x_p(k+1) = A_p x_p(k) + B_p u_1^p(k) \tag{3.4}$$
$$y_1^p(k) = C_p x_p(k) + D_p u_1^p(k) \tag{3.5}$$
$$x_p(-k) = 0,\ k \in \mathbb{N}. \tag{3.6}$$

Note that since $x_p(-k) = 0$ for all $k \in \mathbb{N}$ it follows that $y_1^p(-k) = P_p(u_1^p)(-k) = 0$ for all $k \in \mathbb{N}$.

The residual operator is then of the form

$$X : \mathcal{W}_e \to map(\mathbb{N}, map(\mathcal{P}, \mathbb{R}^+)) : w_2 \mapsto [k \to (p \mapsto r_p[k])] \tag{3.7}$$

where $r_p[k]$ is said to be the residual of a plant $P_p,\ p \in \mathcal{P}$ at time $k \in \mathbb{N}$.

In Section 2 we will impose considerable structure on the operator X, in particular that it factorises in the form $X = NE$, where the assumptions on N, E ensure that X can be given the interpretation above. Note that the implementation of the EMMSAC algorithm requires a realisation of the operator $X = NE$ and it is only the analysis that requires the factorisation into the operators N, E. We now introduce two key classes of disturbance estimators. Both classes are based on measuring the sizes of minimal consistent disturbances for which we introduce the following notation:

Consider disturbances $(u_0^p, y_0^p)^\top$ that are consistent with a plant model P_p and the observation $(u_2, y_2)^\top$; we are interested in disturbances $(u_0^p, y_0^p)^\top$ such that, given the plant model $P_p,\ p \in \mathcal{P}$ and given an observation signal $(u_2, y_2)^\top$, we have $(u_2^p, y_2^p)^\top = (u_2, y_2)^\top$ over the interval $[a,b],\ a \leq b,\ a, b \in \mathbb{Z}$ (also see Figure 3.1). Typically, the observations

$(u_2, y_2)^\top$ are generated from a 'true' plant P_{p_*} (where the true plant P_{p_*} is not necessarily equal to the plant model P_p) with 'observed' signal $(u_2, y_2)^\top$ and non-observed 'true' disturbances $(u_0, y_0)^\top$.

We therefore require that the equations

$$y_1^p = P_p u_1^p \tag{3.8}$$
$$u_0^p = u_1^p + u_2 \tag{3.9}$$
$$y_0^p = y_1^p + y_2 \tag{3.10}$$

are satisfied over the interval $[a, b]$:

Definition 3.1. *Let $a \leq b,\ a, b \in \mathbb{Z}$. The set of weakly consistent disturbance signals $\mathcal{N}_p^{[a,b]}(w_2)$ to a plant $P_p,\ p \in \mathcal{P}$ and the observation $w_2 = (u_2, y_2)^\top$ is defined by:*

$$\mathcal{N}_p^{[a,b]}(w_2) := \left\{ v \in \mathcal{W}|_{[a,b]} \;\middle|\; \begin{array}{c} \exists (u_0^p, y_0^p)^\top \in \mathcal{W}_e \ s.t. \\ \mathscr{R}_{b-a,b} P_p \left(u_0^p - u_2\right) = \mathscr{R}_{b-a,b}(y_0^p - y_2), \\ v = (\mathscr{R}_{b-a,b} u_0^p, \mathscr{R}_{b-a,b} y_0^p) \end{array} \right\} \subset \mathcal{W}|_{[a,b]}.$$

Hence $\mathcal{N}_p^{[a,b]}(w_2)$ denotes the set of all disturbances $(u_0^p|_{[a,b]}, y_0^p|_{[a,b]})^\top$ compatible with observation $(u_2|_{[a,b]}, y_2|_{[a,b]})^\top$ and equations (3.8)–(3.10) for $p \in \mathcal{P}$.

For the remainder of this chapter we assume $\mathcal{N}_p^{[a,b]}(w_2)$ is closed and convex for all $a \leq b \in \mathbb{T},\ w_2 \in \mathcal{W}_e$, noting that if P_p is linear, then this holds.

The following classes of optimal disturbance estimators are from French and Trenn (2005).

1.1 Estimator A: The infinite horizon estimator

Let $k \in \mathbb{N}$ and $w_2 \in \mathcal{W}_e$. To a plant model $P_p,\ p \in \mathcal{P}$, let the residual operator X_A be given by:

$$X_A(w_2)(k)(p) = r_p^A[k] = \inf\{r \geq 0 \mid r = \|v_0\|,\ v_0 \in \mathcal{N}_p^{[0,k]}(w_2)\}, \tag{3.11}$$

where $\mathcal{N}_p^{[0,k]}(w_2)$ is the set of all disturbance signals consistent with the observation w_2 as well as the plant P_p over the interval $[0, k]$.

Observe that a direct implementation of X_A is not feasible since the computational complexity of the optimisation problem grows with $k \in \mathbb{N}$. However, in the l_2 setting the residuals $r_p^A[k],\ p \in \mathcal{P}$ can be determined from the residuals in a Kalman filter bank (see Fisher-Jeffes (2003) and Section 4). This makes the computation of $r_p^A[k]$ feasible as the Kalman filter algorithm is recursive — the computational complexity is invariant to $k \in \mathbb{N}$ and only depends on the order of the corresponding plant model $p \in \mathcal{P}$.

1.2 Estimator B: The finite horizon estimator

Let $k, \lambda \in \mathbb{N}$ and let $w_2 \in \mathcal{W}_e$. To a plant model $P_p,\ p \in \mathcal{P}$ let the residual operator X_B be given by:

$$X_B(w_2)(k)(p) = r_p^B[k] = \left\| r_p^B[k-1], i_p[k] \right\| \tag{3.12}$$

where

$$i_p[k] = \inf\{r \geq 0 \mid r = \|v_0\|, v_0 \in \mathcal{N}_p^{[k-\lambda,k]}(w_2)\} \tag{3.13}$$

and $\mathcal{N}_p^{[k-\lambda,k]}(w_2)$ is the set of all disturbance signals consistent with the observation w_2 and the plant P_p over the interval $[k-\lambda, k]$.

Note that the formulation of X_B is recursive by construction, therefore the computational complexity does not depend on $k \in \mathbb{N}$ but only on the complexity of the involved optimisation, i.e. the computation of $i_p[k]$ for all $k \in \mathbb{N}$, which is of bounded complexity. The norm in (3.12) and (3.13) can be taken to be $l_r,\ 1 \leq r \leq \infty$, giving rise to different optimisations. In Section 5 we will show that such standard (matrix) optimisation problems can be solved by many possible implementations, i.e. in the linear case via computing a suitable pseudo inverse in l_2 or via linear programming in l_∞ .

2 Estimator structure

We have indicated that there will be a requirement that the residual operator X factorises to $X = NE$. This factorisation is necessary since we will argue about the estimator's internal properties, such as consistency and structure of disturbance estimates, that cannot be inferred from the residual only.

For $k \in \mathbb{N},\ p \in \mathcal{P}$ define the estimation operator

$$E : \mathcal{W}_e \to map(\mathbb{N}, map(\mathcal{P}, map(\mathbb{N}, \mathbb{R}^h))) \tag{3.14}$$

by

$$w_2 \mapsto \left[k \mapsto (p \mapsto d_p[k])\right] \tag{3.15}$$

where $d_p[k]$ represents the time series of the disturbance estimates at time $k \in \mathbb{N}$ corresponding to a plant $p \in \mathcal{P}$ denoted by

$$d_p[k] : \mathbb{N} \to map(\mathbb{N}, \mathbb{R}^h)$$

and

$$d_p[k] = (d_p[k](0), d_p[k](1), \ldots, d_p[k](k), 0, \cdots)$$

where $h \in \mathbb{N} \cup \{\infty\}$ depends on the plant.

Note that this estimate will not be recursive in general, i.e.

$$\mathscr{T}_k d_p[l] \neq \mathscr{T}_k d_p[k],\ l > k.$$

Since we are interested in the size of the disturbance estimates, we define the norm operator

$$N : map(\mathbb{N}, map(\mathcal{P}, map(\mathbb{N}, \mathbb{R}^h))) \to map(\mathbb{N}, map(\mathcal{P}, \mathbb{R}^+)) \tag{3.16}$$

by

$$\big[k \mapsto (p \mapsto d_p[k])\big] \mapsto \big[k \mapsto (p \mapsto \|d_p[k]\| = r_p[k])\big] \tag{3.17}$$

where we recall that X is needed for the algorithm and the factorisation NE is only for analytical purposes.

We will now revisit estimator A and B and investigate their internal structure by giving an explicit formulation of the estimation operator E.

2.1 Estimator A: The infinite horizon estimator

To a plant model $P_p,\ p \in \mathcal{P},\ k \in \mathbb{N}$ let estimator A with $h = \infty$ in equation (3.14) be given by:

$$\begin{aligned} E_A(w_2)(k)(p) &= d_p^A[k] \in map(\mathbb{N}, \mathbb{R}^h) && (3.18)\\ d_p^A[k] &= \mathscr{T}_k \operatorname*{argmin}_{w_0 \in \mathcal{N}_p^{[0,k]}(w_2)} \|w_0\| && (3.19) \end{aligned}$$

if there exists a unique minimum, or any $d_p^A[k]$ satisfying

$$d_p^A[k] \in \{w_0 \in \mathcal{N}_p^{[0,k]}(w_2) |\ \|w_0\| = \inf\{r \geq 0 \mid r = \|v_0\|,\ v_0|_{[0,k]} \in \mathcal{N}_p^{[0,k]}(w_2)\} \tag{3.20}$$

if the minimum is not unique, where we recall from Definition 3.1 that $\mathcal{N}_p^{[0,k]}(w_2)$ is the set of all disturbance signals consistent with the observation w_2 as well as the plant P_p over the interval $[0, k]$.

Observe that $d_p^A[k]$ is structured as in Figure 3.2, i.e. at every time instance $i,\ 0 \leq i \leq k$ the disturbance estimate $d_p[k](i)$ consists of a single element $(u_0^p(i), y_0^p(i))^\top$.

To see that X_A does indeed factorise to N and E_A consider the following lemma:

Lemma 3.2. *Let X_A be defined as in equation* (3.11)*, E_A be defined as in equations* (3.18)–(3.20) *and N be defined as in equations* (3.16),(3.17)*. Then $X_A = NE_A$.*

Proof Observe that for all $w_2 \in \mathcal{W}_e$,

$$NE_A(w_2)(k)(p) = \|d_p[k]\| = r_p[k] = X_A(w_2)(k)(p),\ \forall k \in \mathbb{N},\ \forall p \in \mathcal{P}.$$

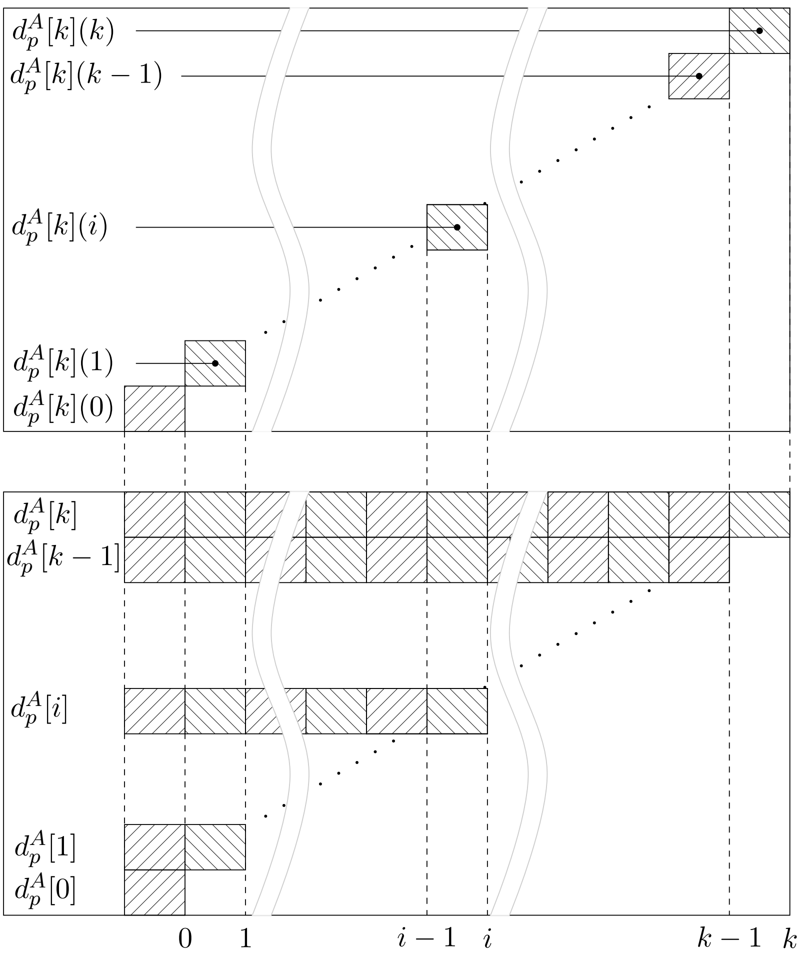

FIGURE 3.2: Structure of $d_p^A[k]$

Hence $X_A = NE_A$ as required. □

2.2 Estimator B: The finite horizon estimator

The second example is motivated by the fact that the EMMSAC algorithm only requires disturbance estimates that are consistent over suitable finite intervals of length $j \in \mathbb{N},\ 0 \leq j \leq \lambda$, where $\lambda \in \mathbb{N}$ is fixed. This allows for the construction of a finite horizon estimator as follows.

Let $k, \lambda, i \in \mathbb{N},\ 0 \leq i \leq k$. To a plant model $P_p,\ p \in \mathcal{P}$ let estimator B with $h = (m+o)(\lambda+1)$ in equation (3.14) be given by:

$$E_B(w_2)(k)(p) = d_p^B[k] \in map(\mathbb{N}, \mathbb{R}^h) \tag{3.21}$$

$$d_p^B[k](i) = \operatorname*{argmin}_{w_0 \in \mathcal{N}_p^{[i-\lambda,i]}(w_2)} \|w_0\|, \tag{3.22}$$

if there exists a unique minimum, or any $d_p^B[k](i)$ satisfying

$$d_p^B[k](i) \in \mathcal{N}_p^{[i-\lambda,i]}(w_2) \mid \|w_0\| = \inf\{r \geq 0 \mid r = \|v_0\|,\ v_0 \in \mathcal{N}_p^{[i-\lambda,i]}(w_2)\} \tag{3.23}$$

if the minimum is not unique, where $\mathcal{N}_p^{[i-\lambda,i]}(w_2)$ is the set of all disturbance signals consistent with the observation w_2 and the plant P_p over the interval $[i-\lambda, i]$.

Observe that $d_p^B[k](i) = d_p^B[i](i)$ for $0 \leq i \leq k$ and that $d_p^B[k]$ is structured as in Figure 3.3, i.e. at every time instance $i,\ 0 \leq i \leq k$ the disturbance estimate $d_p[k](i)$ consists of a 'slice' of disturbance estimates with length λ.

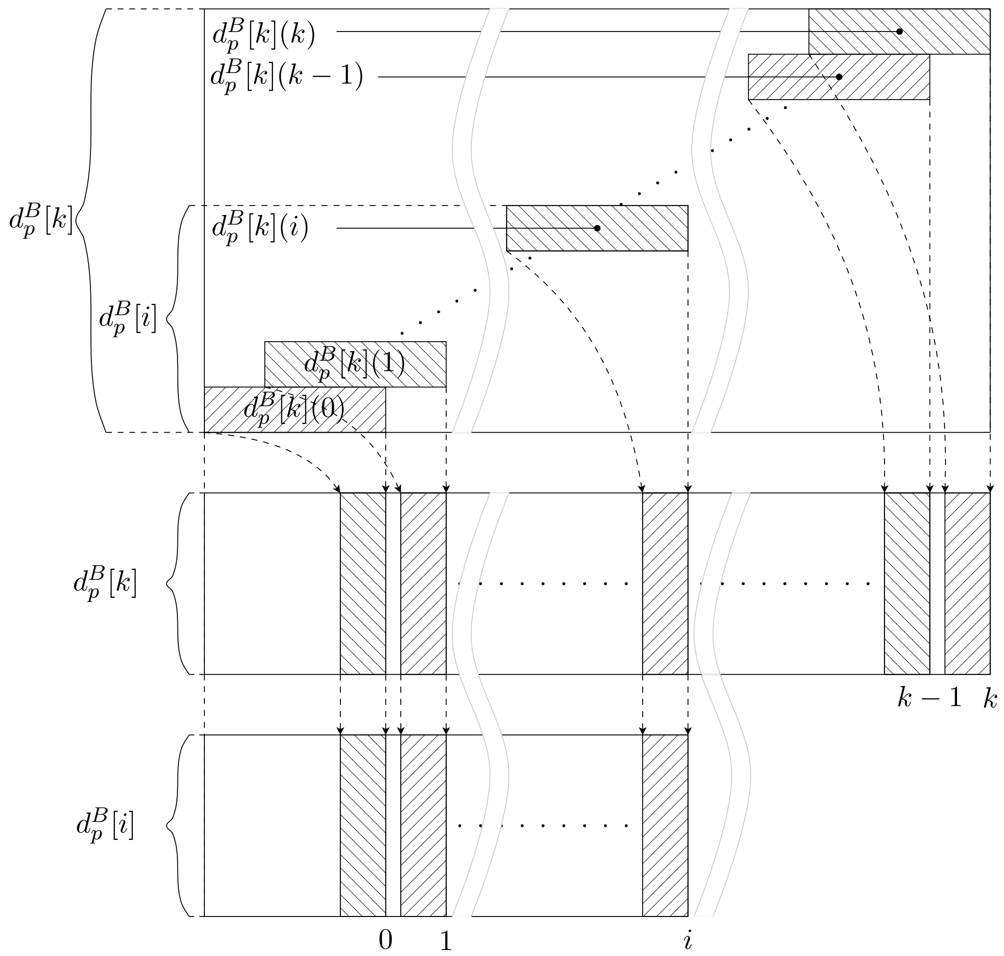

FIGURE 3.3: Structure of $d_p^B[k]$

To see that X_B does indeed factorise to N and E_B consider the following lemma:

Lemma 3.3. *Let X_B be defined as in equations* (3.12),(3.13)*, E_B be defined as in equations* (3.21)–(3.23) *and N be defined as in equations* (3.16),(3.17)*. Then $X_B = NE_B$.*

Proof Note that

$$\|a, b\|_r = \big\| \|a\|_r, \|b\|_r \big\|_r,\ a, b \in l_r,\ 1 \leq r \leq \infty.$$

Since $d_p^B[k](i) = d_p^B[i](i)$ for $0 \leq i \leq k$, we have for all $w_2 \in \mathcal{W}_2$ that:

$$\begin{aligned} NE_B(w_2)(k)(p) &= \|d_p[k]\| \\ &= \|d_p[k](0), d_p[k](1), \cdots, d_p[k-1](k), d_p[k](k)\| \\ &= \|d_p[k-1](0), d_p[k-1](1), \cdots, d_p[k-1](k-1), d_p[k](k)\| \\ &= \|\|d_p[k-1]\|, \|d_p[k](k)\|\| \end{aligned}$$

where

$$d_p[k](k) \in \{w_0 \in \mathcal{N}_p^{[k-\lambda,k]}(w_2) | \ \|w_0\| = \inf\{r \geq 0 \mid r = \|v_0\|, \ v_0 \in \mathcal{N}_p^{[k-\lambda,k]}(w_2)\}.$$

Since

$$i_p[k] = \|d_p[k](k)\| = \inf\{r \geq 0 \mid r = \|v_0\|, v_0 \in \mathcal{N}_p^{[k-\lambda,k]}(w_2)\},$$

we arrive with $r_p[k] = \|d_p[k]\|$ at

$$NE_B(w_2)(k)(p) = r_p[k] = \left\| r_p[k-1], i_p[k] \right\| = X_B(w_2)(k)(p).$$

Hence $X_B = NE_B$ as required. □

3 The estimator axioms

Instead of working with estimator A or B directly, we now state 5 abstract estimator assumptions that any estimator is required to satisfy and on which the subsequent analysis will rest. The purpose of such an axiomatic treatment, as discussed in the introduction, is to separate the problem of conducting (robustness) analysis from the problem of (efficient) disturbance estimation; any particular construction or implementation of an estimator is allowed as long as it satisfies the following estimator assumptions.

Assumption 3.4. *Let $\lambda \in \mathbb{R}$ be given.*

1. *(Causality): E is causal.*

2. *(Minimality): There exists a $\mu > 0$ such that for all $k \geq 0$, for $p \in \mathcal{P}$ and for all $(w_0, w_1, w_2) \in \mathcal{W} \times \mathcal{W}_e \times \mathcal{W}_e$ satisfying equations* (2.7)–(2.9) *for $P = P_p$,*

$$NE(w_2)(k)(p) = \|E(w_2)(k)(p)\| = \|d_p[k]\| \leq \mu \|\mathscr{T}_k w_0\|.$$

3. *(Weak consistency): Let $0 \leq j \leq \lambda$. For all $p \in \mathcal{P}$ there exist maps*

$$\Phi_j : map(\mathbb{N}, \mathbb{R}^h) \to \mathbb{R}^{m(j+1)} \times \mathbb{R}^{o(j+1)},$$

such that for all $(w_0, w_1, w_2) \in \mathcal{W} \times \mathcal{W}_e \times \mathcal{W}_e$ *satisfying equations* (2.7)–(2.9) *for* $P = P_p$ *and for all* $k \in \mathbb{N}$,

$$\Phi_j E(w_2)(k)(p) = \Phi_j d_p[k] \in \mathcal{N}_p^{[k-j,k]}(w_2)$$

and

$$\|\Phi_j E(w_2)(k)(p)\| = \|\Phi_j d_p[k]\| \leq \|\mathscr{R}_{j,k} d_p[k]\| = \|\mathscr{R}_{j,k} E(w_2)(k)(p)\|.$$

4. *(Monotonicity): For all* $p \in \mathcal{P}$, *for all* $k, l \in \mathbb{N}$ *with* $0 \leq k \leq l$ *and for all* $(w_0, w_1, w_2) \in \mathcal{W} \times \mathcal{W}_e \times \mathcal{W}_e$ *satisfying equations* (2.7)–(2.9) *for* $P = P_p$, *there holds*
$$\|E(w_2)(k)(p)\| = \|d_p[k]\| \leq \|\mathscr{T}_k d_p[l]\| = \|\mathscr{T}_k E(w_2)(l)(p)\|.$$

5. *(Continuity): There exists a* $c : \mathbb{Z} \to \mathbb{R}$, $\|c\| < \infty$ *and a function* $\chi : \mathcal{P} \times \mathcal{P} \to \mathbb{R}^+ \cup \{\infty\}$, $\chi(p,p) = 0$, $\forall p \in \mathcal{P}$, *such that for all* $p_1, p_2 \in \mathcal{P}$ *and* $w_2 \in \mathcal{W}_e$, *there holds*

$$\begin{aligned} \big| \|\Phi_j E(w_2)(k)(p_1)\| - \|\Phi_j E(w_2)(k)(p_2)\| \big| &= \big| \|\Phi_j d_{p_1}[k]\| - \|\Phi_j d_{p_2}[k]\| \big| \\ &\leq \chi(p_1,p_2) \|\Upsilon_k w_2\|, \ 0 \leq j \leq \lambda, \ k \in \mathbb{N} \end{aligned}$$

where Φ_j *is as in Assumption 3 and*

$$(\Upsilon_k w_2)(i) = \begin{cases} c(k-i) w_2(i) & \text{if} \quad i \leq k \\ 0 & \text{else} \end{cases}.$$

Although these assumptions may appear rather technical, they have very intuitive interpretations:

1. A later implementation requires causality so we impose it from the start. Note that N is always causal and therefore it suffices to assume that E is causal.

2. We denote an optimal infinite horizon disturbance estimator to a plant P_p, an estimator which returns the smallest disturbances consistent with the observation w_2 and P_p over the interval $[0,k]$, $k \in \mathbb{N}$. However, to require optimality over the interval $[0,k]$, $k \in \mathbb{N}$ per se is too strong in a sense that no other estimator than the infinite horizon estimator will satisfy it.

   Instead, we impose a milder assumption: we require that the disturbance estimate $\|E_p(w_2)(k)(p)\|$, corresponding to an estimator constructed to a plant model P_p which equals the true plant P, is smaller (up to a constant $\mu > 0$) than the true disturbances $\|\mathscr{T}_k w_0\|$ at time $k \in \mathbb{N}$. Note that the true disturbances are always consistent with P and the observation w_2 since w_2 is constructed from them.

3. The purpose of the map $\Phi_j : map(\mathbb{N}, \mathbb{R}^h) \to \mathbb{R}^{m(j+1)} \times \mathbb{R}^{o(j+1)}$ is to extract the relevant parts from the disturbance estimate to allow a test for weak consistency.

4. This assumption is also inspired by the properties of the optimal infinite horizon estimator. We require that a disturbance estimate $d_p[k]$ is minimal at time $k \in \mathbb{N}$ in a sense that no future disturbance estimate $d_p[l],\ l > k$ truncated at k given by $\mathscr{T}_k d_p[l]$, can be smaller.

5. The last assumption implies that if two plants are close to each other, their disturbance estimates are also 'close', i.e. their difference is small in norm. χ can be thought of as a measure of distance between two plants, and in fact it will turn out later that for the estimators considered here, χ is related to the gap metric.

   The weight c describes the effect of w_2 on the disturbance estimates. Since we require $\|c\| < \infty$ this effect is required to be bounded. For example in $l_r,\ 1 \leq r < \infty$ we can say that the effect of w_2 needs to diminish over time since $\|c\|_r < \infty$ implies that c is summable, hence c needs to converge to zero. Therefore $\Upsilon_k w_2$ returns a weighted signal w_2 such that earlier $w_2(i),\ i \in \mathbb{N}$ have smaller weights. For l_∞ we merely require that the weights $c(i),\ i \in \mathbb{N}$ are finite.

We will now show that these assumptions are met by the given estimator constructions.

Lemma 3.5. *Estimator A fulfils assumptions 3.4(1–5).*

Proof Let $1 \leq r \leq \infty$. Let $\lambda = \infty$.

1. Causality: The disturbance estimate at time $k \in \mathbb{N}$ does not depend on future information $w_2|_{(k,\infty)}$ and is therefore causal.

2. Minimality: Observe that for any $(w_0, w_1, w_2) \in \mathcal{W} \times \mathcal{W}_e \times \mathcal{W}_e$ satisfying equations (2.7)–(2.9) for $P = P_p$ and for $k \in \mathbb{N}$ we have $\mathscr{T}_k w_0 \in \mathscr{T}_k \mathcal{N}_p^{[0,k]}(w_2)$. Hence

$$\|E_A(w_2)(k)(p)\| = \inf\{r \geq 0 \mid r = \|v_0\|, v_0 \in \mathcal{N}_p^{[0,k]}(w_2)\} \leq \|\mathscr{T}_k w_0\| \leq \|w_0\|$$

   and hence $\mu = 1$.

3. Weak consistency: Let $0 \leq j \leq \lambda,\ p \in \mathcal{P},\ w_2 \in \mathcal{W}_e$. Let Φ_j be defined by $\Phi_j x = \mathscr{R}_{j,k} x,\ x \in \mathcal{S}$, and therefore $\|\Phi_j E_A(w_2)(k)(p)\| = \|\mathscr{R}_{j,k} E_A(w_2)(k)(p)\|$. We then have

$$\Phi_j E_A(w_2)(k)(p) = \mathscr{R}_{j,k} E_A(w_2)(k)(p) \in \mathscr{R}_{j,k} \mathcal{N}_p^{[0,k]}(w_2) \subset \mathcal{N}_p^{[k-j,k]}(w_2).$$

4. Monotonicity: Let $p \in \mathcal{P}$, let $k \leq l,\ k, l \in \mathbb{N}$ and suppose $(w_0, w_1, w_2) \in \mathcal{W} \times \mathcal{W}_e \times \mathcal{W}_e$ satisfy equations (2.7)–(2.9) for $P = P_p$. Observe that $\mathscr{T}_k E_A(w_2)(l)(p) \in$

$\mathscr{T}_k\mathcal{N}_p^{[0,k]}(w_2)$. Since

$$\|E_A(w_2)(k)(p)\| = \inf\{r \geq 0 \mid r = \|v_0\|,\ v_0 \in \mathcal{N}_p^{[0,k]}(w_2)\}$$

it follows that $\|E_A(w_2)(k)(p)\| \leq \|\mathscr{T}_k E_A(w_2)(l)(p)\|$ as required.

5. Continuity: Let $p_1, p_2 \in \mathcal{P},\ k \in \mathbb{N},\ w_2 \in \mathcal{W}_2$. Then

$$\left|\|\Phi_j d^A_{p_1}[k]\| - \|\Phi_j d^A_{p_2}[k]\|\right| \leq \left|\|\Phi_j d^A_{p_1}[k] - \Phi_j d^A_{p_2}[k]\|\right| \quad \leq \quad \chi(p_1,p_2)\|\Upsilon_k w_2\|$$

where

$$\chi(p_1,p_2) = \begin{cases} 0 & \text{if } p_1 = p_2 \\ \infty & \text{if not} \end{cases}$$

for some Υ with $\|c\| < \infty$. □

Lemma 3.6. *Let $1 \leq r \leq \infty$ and $\lambda \in \mathbb{N}$. Estimator B fulfils Assumptions 3.4(1–5).*

Proof Let $k \in \mathbb{N}$.

1. Causality: E_B is invariant to $w_2|_{(k,\infty)}$.

2. Minimality: Observe that for any $(w_0, w_1, w_2) \in \mathcal{W} \times \mathcal{W}_e \times \mathcal{W}_e$ satisfying equations (2.7)–(2.9) for $P = P_p$ and for $k \in \mathbb{N}$ we have $\mathscr{R}_{\lambda,i} w_0 \in \mathcal{N}_p^{[i-\lambda,i]}(w_2),\ 0 \leq i \leq k$. Hence

$$\begin{aligned}\|d^B_p[k](i)\| &= \inf\{r \geq 0 \mid r = \|v_0\|, v_0 \in \mathcal{N}_p^{[i-\lambda,i]}(w_2)\} \\ &\leq \|\mathscr{R}_{\lambda,i} w_0\|,\ 0 \leq i \leq k,\ k \in \mathbb{N}.\end{aligned}$$

This leads to

$$\begin{aligned}\|E_B(w_2)(k)(p)\| &= \|d^B_p[k](0), d^B_p[k](1), \ldots, d^B_p[k](k)\| \\ &\leq \|\mathscr{R}_{\lambda,0}w_0,\ \mathscr{R}_{\lambda,1}w_0,\ \cdots,\ \mathscr{R}_{\lambda,k}w_0\| \\ &\leq \left\| \begin{matrix} w_0(-\lambda), & w_0(1-\lambda), & \cdots, & w_0(k-\lambda) \\ w_0(1-\lambda), & w_0(2-\lambda), & \cdots, & w_0(k+1-\lambda) \\ \vdots & \vdots & \vdots & \vdots \\ w_0(0), & w_0(1), & \cdots, & w_0(k) \end{matrix} \right\| \\ &= (\lambda+1)^{1/r}\|w_0\| \\ &= \mu\|w_0\|\end{aligned}$$

where the first inequality follows from the fact that $\|\|a\|, \|b\|\| = \|(a,b)\|$ holds in $l_r,\ 1 \leq r \leq \infty$.

3. Weak consistency: Let $0 \leq j \leq \lambda$, $p \in \mathcal{P}$. Let Φ_j be defined by $\Phi_j d_p^B[k] = \mathscr{R}_{j,\lambda} d_p^B[k](k)$. Since
$$\mathscr{R}_{j,\lambda} d_p^B[k](k) \subset \mathscr{R}_{j,k} d_p^B[k]$$
there holds
$$\|\Phi_j E_B(w_2)(k)(p)\| = \|\mathscr{R}_{j,\lambda} d_p^B[k](k)\| \leq \|\mathscr{R}_{j,k} d_p^B[k]\| = \|\mathscr{R}_{j,k} E_B(w_2)(k)(p)\|.$$
Also
$$\Phi_j d_p^B[k] = \mathscr{R}_{j,\lambda} d_p^B[k](k) \in \mathcal{N}_p^{[k-j,k]}(w_2).$$

4. Monotonicity: Let $p \in \mathcal{P}$, let $k \leq l,\ k, l \in \mathbb{N}$ and suppose $(w_0, w_1, w_2) \in \mathcal{W} \times \mathcal{W}_e \times \mathcal{W}_e$ satisfy equations (2.7)–(2.9) for $P = P_p$. Since $\mathscr{T}_k d_p^B[l] = d_p^B[k]$ it follows that
$$\|E_p^B(w_2)(k)(p)\| = \|\mathscr{T}_k E_p^B(w_2)(l)(p)\|.$$

5. Continuity: Note that in the finite horizon case a similar construction for χ as in the infinite horizon case is sufficient to satisfy Assumption 3.4(5). However, we will later seek to establish an additional continuity property of χ, hence give an alternative construction.

   Let $1 \leq j \leq \lambda$, $k \in \mathbb{N}$, $p \in \mathcal{P}$. From Assumption 3 let Φ_j be defined by
$$\Phi_j d_p^B[k] = \mathscr{R}_{j,\lambda} d_p^B[k](k).$$
   Define $\Pi_p^{[k-\lambda,k]} : \mathcal{W}_2|_{[k-\lambda,k]} \to \mathcal{W}_2|_{[k-\lambda,k]}$ by
$$\Pi_p^{[k-\lambda,k]} \mathscr{R}_{k-\lambda,k} w_2 = d_p^B[k](k).$$
   We therefore have
$$\Phi_j d_p^B[k] = \mathscr{R}_{j,\lambda} \Pi_p^{[k-\lambda,k]} \mathscr{R}_{\lambda,k} w_2.$$
   For $p_1, p_2 \in \mathcal{P}$ let
$$\chi_k(p_1,p_2) = \sup_{x \in \mathcal{W}_e|_{[k-\lambda,k]},\ \|x\| \neq 0} \frac{\left|\|\mathscr{R}_{j,\lambda}\Pi_{p_1}^{[k-\lambda,k]} x\| - \|\mathscr{R}_{j,\lambda}\Pi_{p_2}^{[k-\lambda,k]} x\|\right|}{\|x\|}$$
   hence
$$\begin{aligned}\left|\|\Phi_j d_{p_1}^B[k]\| - \|\Phi_j d_{p_2}^B[k]\|\right| &\leq \left|\|\mathscr{R}_{j,\lambda}\Pi_{p_1}^{[k-\lambda,k]}\mathscr{R}_{\lambda,k} w_2\| - \|\mathscr{R}_{j,\lambda}\Pi_{p_2}^{[k-\lambda,k]}\mathscr{R}_{\lambda,k} w_2\|\right| \\ &\leq \chi(p_1,p_2)\|\mathscr{R}_{\lambda,k} w_2\|.\end{aligned}$$

   It is important to observe that for $p \in \{p_1, p_2\}$, $\Pi_p^{[i-\lambda,i]} = \Pi_p^{[j-\lambda,j]}$ for all $i, j > \sigma,\ i, j \in \mathbb{N}$ since $\mathcal{N}_p^{[i-\lambda,i]}(w_2) = \mathcal{N}_p^{[j-\lambda,j]}(w_2)$ for all $i, j > \sigma,\ i, j \in \mathbb{N}$. However for $0 \leq i \leq \sigma$ the constraint set $\mathcal{N}_p^{[i-\lambda,i]}(w_2)$ includes the additional constraint that the initial condition of P_p is zero. Hence we can let

$$\chi(p_1, p_2) = \max_{k \geq 0} \chi_k(p_1, p_2).$$

Also note that it follows trivially that $\chi(p,p) = 0,\ p \in \mathcal{P}$.

Finally, let

$$c(i) = \begin{cases} 1 & \text{for} \quad 0 \leq i \leq \lambda \\ 0 & \text{else} \end{cases}$$

and

$$(\Upsilon_k w_2)(i) = \begin{cases} c(k-i) w_2(i) & \text{if} \quad i \leq k \\ 0 & \text{else} \end{cases}.$$

We then have

$$\begin{aligned} \|\mathscr{R}_{\lambda,k} w_2\| &= \|w_2(k), w_2(k-1), \cdots, w_2(k-\lambda)\| \\ &= \|c(0)w_2(k), c(1)w_2(k-1), \cdots, c(\lambda)w_2(k-\lambda)\| \\ &= \|\Upsilon_k w_2\| \end{aligned}$$

hence

$$\big| \|\Phi_j d^B_{p_1}[k]\| - \|\Phi_j d^B_{p_2}[k]\| \big| \leq \chi(p_1, p_2) \|\Upsilon_k w_2\|$$

as required. □

In Chapter 6, it will be important to require that $\chi : \mathcal{P} \times \mathcal{P} \to \mathbb{R}^+ \cup \{\infty\}$ is continuous on certain subsets.

Conjecture 3.7. *Let $1 < r < \infty$. Suppose $\Omega \subset \mathcal{P}_{LTI}$ is compact. For $p_1, p_2 \in \Omega$, let*

$$\chi(p_1, p_2) = \max_{k \geq 0} \sup_{x \in \mathcal{W}_e|_{[k-\lambda,k]},\ \|x\| \neq 0} \frac{\big| \|\mathscr{R}_{j,\lambda} \Pi^{[k-\lambda,k]}_{p_1} x\| - \|\mathscr{R}_{j,\lambda} \Pi^{[k-\lambda,k]}_{p_2} x\| \big|}{\|x\|}$$

Then $\chi|_\Omega$ is continuous.

Continuity of $\chi|_\Omega$ is expected to follow from the well-posedness of the underlying optimisation problem. However at present, this remains an open question.

Note that for $r = 2$, an alternative choice for χ is given by

$$\chi(p_1, p_2) = \max_{k \geq 0} \|\Pi^{[k-\lambda,k]}_{p_1} - \Pi^{[k-\lambda,k]}_{p_2}\|$$

since

$$\begin{aligned}\big|\|\Phi_j d^B_{p_1}[k]\| - \|\Phi_j d^B_{p_2}[k]\|\big| &\leq \|\Phi_j d^B_{p_1}[k] - \Phi_j d^B_{p_2}[k]\| \\ &= \|\mathscr{R}_{j,\lambda}\Pi^{[k-\lambda,k]}_{p_1}\mathscr{R}_{\lambda,k}w_2 - \mathscr{R}_{j,\lambda}\Pi^{[k-\lambda,k]}_{p_2}\mathscr{R}_{\lambda,k}w_2\| \\ &= \|\mathscr{R}_{j,\lambda}(\Pi^{[k-\lambda,k]}_{p_1} - \Pi^{[k-\lambda,k]}_{p_2})\mathscr{R}_{\lambda,k}w_2\| \\ &\leq \|(\Pi^{[k-\lambda,k]}_{p_1} - \Pi^{[k-\lambda,k]}_{p_2})\mathscr{R}_{\lambda,k}w_2\| \\ &\leq \|\Pi^{[k-\lambda,k]}_{p_1} - \Pi^{[k-\lambda,k]}_{p_2}\|\|\mathscr{R}_{\lambda,k}w_2\| \\ &\leq \chi(p_1,p_2)\|\mathscr{R}_{\lambda,k}w_2\|.\end{aligned}$$

χ is closely related to the gap metric. To see this note that since $\tilde{w}^p_0 = \Pi^{[0,\infty]}_p w_2$ is the unique minimizer in $\mathcal{N}^{[0,\infty]}_p(w_2)$, we have with $\hat{\Pi}^{[0,\infty]}_p = I - \Pi^{[0,\infty]}_p$ that

$$\hat{\Pi}^{[0,\infty]}_p w_2 = w_2 - \Pi^{[0,\infty]}_p w_2 = \tilde{w}^p_1 \in \mathcal{M}_p$$

and $\hat{\Pi}^{[0,\infty]}_p$ has the interpretation of a (unique) projection onto the graph $\mathcal{M}_p$ of the plant $p \in \mathcal{P}$.

Now recall from Georgiou and Smith (1990) that for linear plants and $\mathcal{V} = L_2, l_2$:

$$\delta(p_1,p_2) = \|\hat{\Pi}^{[0,\infty]}_{p_1} - \hat{\Pi}^{[0,\infty]}_{p_2}\| = \|\Pi_{p_1}{}^{[0,\infty]} - \Pi_{p_2}{}^{[0,\infty]}\|.$$

Hence χ is a version of the gap metric where only signals over finite intervals $[k-\lambda,k]$ are considered.

The use of a finite horizon in estimator B is penalised with a

$$\mu = (\lambda+1)^{1/r} > 1,\ \ 1 \leq r \leq \infty$$

in contrast to estimator A, where $\mu = 1$. However, the computational complexity of estimator B is invariant to k and only depends on the horizon length $\lambda \in \mathbb{N}$.

3.1 Continuity of χ for Estimator A

We now give an alternative formulation of χ in estimator A and show that χ does not allow a sensible interpretation as a distance since it may be unbounded.

Define

$$\Pi^{[0,k]}_p\mathscr{R}_{k,k}w_2 = d^A_p[k] = E_A(w_2)(p)(k).$$

Let

$$\hat{\Upsilon}_k v = (c^{-1}(k)v(0), c^{-1}(k-1)v(1),\ \cdots, c^{-1}(0)v(k)),\ \ v \in \mathcal{W}_e,$$

hence $\hat{\Upsilon}_k\Upsilon_k w_2 = \mathscr{R}_{k,k}w_2$. Similarly to the construction for estimator B, with $p_1, p_2 \in \mathcal{P}$ let

$$\chi_k(p_1,p_2) = \sup_{x\in\mathcal{W}_e|_{[0,k]},\ \|x\|\neq 0} \frac{\big|\|\mathscr{R}_{j,k}\Pi_{p_1}^{[0,k]}\hat{\Upsilon}_k x\| - \|\mathscr{R}_{j,k}\Pi_{p_2}^{[0,k]}\hat{\Upsilon}_k x\|\big|}{\|x\|}.$$

Since

$$\Phi_j d_p^A[k] = \mathscr{R}_{j,k} d_p^A[k] = \mathscr{R}_{j,k}\Pi_p^{[0,k]}\mathscr{R}_{k,k}w_2$$

we have that

$$\begin{aligned}
\big|\|\Phi_j d_{p_1}^A[k]\| - \|d_{p_2}^A[k]\|\big| &= \big|\|\mathscr{R}_{j,k}\Phi_j\Pi_{p_1}^{[0,k]}\mathscr{R}_{k,k}w_2\| - \|\mathscr{R}_{j,k}\Pi_{p_2}^{[0,k]}\mathscr{R}_{k,k}w_2\|\big| \\
&= \big|\|\mathscr{R}_{j,k}\Pi_{p_1}^{[0,k]}\hat{\Upsilon}_k\Upsilon_k w_2\| - \|\mathscr{R}_{j,k}\Pi_{p_2}^{[0,k]}\hat{\Upsilon}_k\Upsilon_k w_2\|\big| \\
&\leq \chi_k(p_1,p_2)\|\Upsilon_k w_2\|.
\end{aligned}$$

For $r=\infty$, we can now let $c=1$ since then $\|c\|_\infty = 1 < \infty$ and $(\hat{\Upsilon}_k v_2)(i) = (\Upsilon_k v_2)(i) = v(i),\ v\in\mathcal{W}_e,\ 0\leq i\leq k$ where

$$\chi_k(p_1,p_2) = \sup_{x\in\mathcal{W}_e|_{[0,k]},\ \|x\|\neq 0} \frac{\big|\|\mathscr{R}_{j,k}\Pi_{p_1}^{[0,k]}x\| - \|\mathscr{R}_{j,k}\Pi_{p_2}^{[0,k]}x\|\big|}{\|x\|}.$$

Hence we arrive with $\chi(p_1,p_2) = \sup_{k\geq 0}\chi_k(p_1,p_2)$ at

$$\big|\|\Phi_j d_{p_1}^A[k]\| - \|\Phi_j d_{p_2}^A[k]\|\big| \leq \chi(p_1,p_2)\|\Upsilon_k w_2\|$$

which we expect to have a similar continuity property as in the finite horizon case.

However for $1\leq r<\infty$, since $\|c\|_r < 0,\ 1\leq r<\infty$ implies that $c(k)\to 0$ as $k\to\infty$, we have that $c(k)^{-1}\to\infty$ as $k\to\infty$. However $\chi_k(p_1,p_2)$ is given by

$$\begin{aligned}
\chi_k(p_1,p_2) &= \sup_{x\in\mathcal{W}_e|_{[0,k]},\ \|x\|\neq 0} \frac{\big|\|\mathscr{R}_{j,k}\Pi_{p_1}^{[0,k]}\hat{\Upsilon}_k x\| - \|\mathscr{R}_{j,k}\Pi_{p_2}^{[0,k]}\hat{\Upsilon}_k x\|\big|}{\|x\|} \\
&\leq \max\{\|\mathscr{R}_{j,k}\Phi_j\Pi_{p_1}^{[0,k]}\hat{\Upsilon}_k\|, \|\mathscr{R}_{j,k}\Phi_j\Pi_{p_2}^{[0,k]}\hat{\Upsilon}_k\|\}
\end{aligned}$$

and the given upper bound scales with k. Hence χ_k may indeed be unbounded.

To develop alternative formulations of Assumption 3.4(5) such that a continuity property can be satisfied by estimator A, remains an open question.

We will now introduce the Kalman filter and show its relevance for disturbance estimation.

4 The Kalman filter as a disturbance estimator

The question that historically motivated the Kalman filter was as follows. Considering

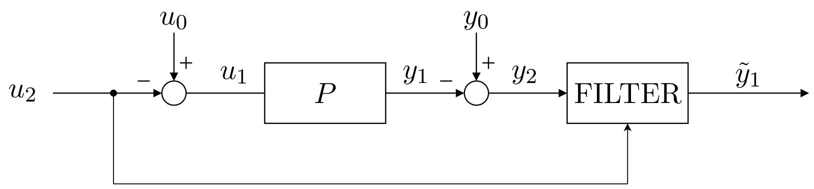

FIGURE 3.4: A common filtering problem: reconstruct $\tilde{y}$ from y

Figure 3.4, assume that the plant P is driven by a signal u_2, for example a force, that is corrupted by an unknown input disturbance signal u_0. Furthermore assume that the plant output y_1 is corrupted by an unknown disturbance signal y_0, resulting in the signal y_2. The objective is now to reconstruct or predict the undisturbed output signal y_1 or an estimate $\tilde{y}_1$ of it from y_2 and u_2 — hence to 'filter' away the effects of y_0 and u_0. Applications of major historic importance are the tracking of ballistic missiles or airplanes with radar. In both cases, the basic dynamical properties of the tracked objects, as well as the input (e.g. thrust), are known. However the (position) data from radar or other positioning systems is often noisy due to reflections, weather conditions etc.

An early approach to such a filtering problem was given by Wiener (1949) where his Wiener filter is constructed such that the expected value of the squared output error $e = y_2 - \tilde{y}_1$ is minimised. Due to the computational complexity of the filter implementation that grows with time, the Wiener filter was of limited use to on-line applications such as tracking.

A decade later Kalman (1960) (in discrete time) and Kalman and Bucy (1961) (in continuous time) gave an efficient, recursive solution to the problem, overcoming these limitations, which is known as the famous Kalman (Bucy) filter algorithm.

Define

$$\bar{\mathcal{P}}_{LTI} := \left\{ (A,B,C) \in \cup_{n\geq 1}\mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} \times \mathbb{R}^{o\times n} \;\middle|\; \begin{array}{ll} (A,B) & \text{is controllable} \\ (A,C) & \text{is observable} \end{array} \right\}. \tag{3.24}$$

Let P_{p,x_0^p} be defined by

$$P_{p,x_0^p} : \mathcal{U}_e \to \mathcal{Y}_e : u_1^p \mapsto y_1^p,\ p = (A_p, B_p, C_p) \in \bar{\mathcal{P}}_{LTI}$$

where

$$
\begin{aligned}
x_p(k+1) &= A_p x_p(k) + B_p u_1^p(k) \\
y_1^p(k) &= C_p x_p(k) \\
x_p(0) &= x_0^p, \ k \in \mathbb{N}.
\end{aligned}
$$

This definition is similar to the one in equations (3.4)–(3.6), however with a possible non-zero initial condition x_0^p.

The discrete-time Kalman filter equations, corresponding to P_{p,x_0^p} and written in the notational style of Willems (2004, 2006) for comparability, are as follows. Let:

- $(w, v)^\top = (u_0^p, y_0^p)^\top$
- $(u, y)^\top = (u_2, y_2)^\top$
- $(F, G, B, H) = (A_p, B_p, -B_p, C_p)$

and $T \geq 0$, $\Sigma : \mathbb{N} \mapsto \mathbb{R}^{n \times n}$, $\hat{x} : [0, T] \mapsto \mathbb{R}^n$ be given by

$$\hat{x}(k+1/2) = \hat{x}(k) + \Sigma(k)H^\top[H\Sigma(k)H^\top + I]^{-1}[y(k) - H\hat{x}(k)] \tag{3.25}$$

$$\Sigma(k+1/2) = \Sigma(k) - \Sigma(k)H^\top[H\Sigma(k)H^\top + I]^{-1}H\Sigma(k) \tag{3.26}$$

$$\hat{x}(k+1) = F\hat{x}(k+1/2) + Bu(k) \tag{3.27}$$

$$\Sigma(k+1) = F\Sigma(k+1/2)F^\top + GG^\top \tag{3.28}$$

$$\tilde{y}_1(k) = H\hat{x}(k) \tag{3.29}$$

where the initial conditions are specified by $\Sigma(0), \hat{x}(0)$. The idea is that $\hat{x}$ is an estimate of x_p and $\tilde{y}_1$ (as in Figure 3.4) is given by equation (3.29).

Traditionally the Kalman filter is analysed in a stochastic setting where, analogously to the Wiener filter, it can be shown to minimise the expected value of the squared estimation error (e.g. see Kalman (1960); Maybeck (1979); Stengel (1986); Welch and Bishop (2001)) — in fact the Kalman filter is known to asymptotically reduce to the Wiener filter. However, the Kalman filter also allows an interpretation as a deterministic least-squares filter (see Mortenson (1968); Hijab (1980); Fleming (1997); McEneaney (1998)). In particular Swerling (1971), Sontag (1998) and Willems (2004) analyse the continuous-time Kalman filter in the deterministic domain and are able to show explicitly that the Kalman filtering problem is equivalent to the deterministic least-squares filtering problem. Fisher-Jeffes (2003) utilised dynamic programming to deterministically show the connection between the Kalman filter and the least-squares filter in discrete time.

The intention of the following argument is to establish an alternative, simple and complete proof showing the equivalence of discrete-time Kalman filtering and least-squares filtering in the deterministic setting.

Define

$$\mathcal{Z}_p^{[a,b]}(w_2) := \left\{ v \in \mathbb{R}^{m(T+1)} \times \mathbb{R}^{o(T+1)} \times \mathbb{R}^n \;\middle|\; \begin{array}{c} \exists (u_0^p, y_0^p, x_0^p)^\top \in \mathcal{U}_e \times \mathcal{Y}_e \times \mathbb{R}^n \; s.t. \\ \mathscr{R}_{b-a,b} P_{p,x_0^p}(u_0^p - u_2) = \mathscr{R}_{b-a,b}(y_0^p - y_2), \\ v = \left(\mathscr{R}_{b-a,b} u_0^p, \mathscr{R}_{b-a,b} y_0^p, x_p^0\right)^\top \end{array} \right\},$$

which is the set of initial conditions x_0^p and disturbance signals u_0^p, y_0^p that are compatible with a plant P_{p,x_0^p} and the observation u_2, y_2 over the interval $[a,b],\ a \leq b$.

Let

$$(\tilde{u}_0^p, \tilde{y}_0^p, \tilde{x}_0^p) = \underset{u_0^p, y_0^p, x_0^p \in \mathcal{Z}_p^{[0,k]}(w_2)}{\operatorname{argmin}} \left(\|x_0^p\|^2_{\Sigma^{-1}(0)} + \|u_0^p\|_2^2 + \|y_0^p\|_2^2 \right),\ k \in \mathbb{N} \tag{3.30}$$

be the smallest (in a least-squares sense) such disturbances and initial condition over the interval $[0,k],\ k \in \mathbb{N}$.

Definition 3.8. *A causal operator $F : \mathcal{W}_e \to \mathcal{Y}_e : (u_2, y_2) \mapsto \tilde{y}_1$ that constructs a signal $\tilde{y}_1$ for the plant $P = P_{p,\tilde{x}_0^p}$ and observation $(u_2, y_2)^\top \in \mathcal{W}_e$ such that*

$$\tilde{y}_1(k) = P_{p,\tilde{x}_0^p}(\tilde{u}_0^p - u_2)(k),\ k \in \mathbb{N}$$

where $(\tilde{u}_0^p, \tilde{y}_0^p, \tilde{x}_0^p)$ are as in equation (3.30), is called a deterministic least-squares filter.

The least-squares filter therefore reconstructs the output signal $\tilde{y}_1(k)$ at time $k \in \mathbb{N}$, which forms the prediction of $y_1(k)$, by driving a plant $P = P_{p,x_0^p}$ with initial condition $x_0^p = \tilde{x}_0^p$ and disturbances $(\tilde{u}_0^p, \tilde{y}_0^p)$; that are the smallest least-squares solutions, consistent with the observation $(u_2, y_2)^\top \in \mathcal{W}_e$ and P_{p,x_0^p} up to time $k \in \mathbb{N}$.

As a notion of the output error between the observation y_2 and the estimation of the Kalman filter, define the (scaled) residual $r : \mathbb{N} \to \mathbb{R}^+$ by

$$r(T) = \left[\sum_{k=0}^{T} \|y_2(k) - \tilde{y}_1(k)\|^2_{[H\Sigma(k)H^\top + I]^{-1}} \right]^{1/2},\ T \geq 0.$$

Note that the inverse $[H\Sigma(k)H^\top + I]^{-1}$ exists since it can be shown that $\Sigma(k)$ is positive semi-definite for all $k \in \mathbb{N}$ provided $\Sigma(0) \geq 0$ (see Lemma A.1 in the Appendix).

We now claim the following:

Theorem 3.9. [Theorem A.6] *Let* $p = (A_p, B_p, C_p) \in \bar{\mathcal{P}}_{LTI}$ *and suppose* C_p *is full row rank. Let* $(F, G, B, H) = (A_p, B_p, -B_p, C_p)$*. The Kalman filter equations* (3.25)–(3.29) *with initial condition* $\hat{x}(0) = 0$ *and* $\Sigma(0) = \Sigma(0)^\top > 0$ *describe a deterministic least-squares filter:*

$$r^2(T) = \inf_{(u_0^p, y_0^p, x_0^p) \in \mathcal{Z}_p^{[0,T]}(w_2)} (\|x_0^p\|_{\Sigma^{-1}(0)} + \|u_0^p\|_2^2 + \|y_0^p\|_2^2).$$

Proof The proof of this result can be found in the Appendix. The overall strategy of the argument is based on Willems (2006) and Willems (2004). The statements of Lemma A.2, A.4, A.5 are from Willems (2004). The proof of Theorem 3.9[Theorem A.6] is based on Willems (2004). All other theorem/lemma statements and proofs are due to the present author. □

For the purpose of this thesis it is of interest that the Kalman filter computes the residual $r(T)$ recursively, which in turn is related to the computation of the smallest consistent disturbances, rather than the predictive abilities of the Kalman filter. That is, the Kalman filter may be used to compute the size of the least-squares disturbance estimates, i.e. the infinite horizon disturbance estimates in l_2 (estimator A). Observe that the infinite horizon disturbance estimates of estimator A are constructed such that they are compatible with the observation w_2 and the plant model $P_p = P_{p,x_0^p}$ for a zero initial condition $x_0^p = 0$. Hence, to be able to use the Kalman filter for disturbance estimation, we have to ensure that the residual computation relates to the least-squares filter initialised to zero. In Fisher-Jeffes (2003) this is assumed implicitly, but does not appear to be proved. Here we state the required property as a theorem:

Theorem 3.10. [Theorem A.8] *Let* $p = (A_p, B_p, C_p) \in \bar{\mathcal{P}}_{LTI}$ *and suppose* C_p *is full row rank. Let* $(F, G, B, H) = (A_p, B_p, -B_p, C_p)$*. The Kalman filter equations* (3.25)–(3.29) *with initial condition* $\hat{x}(0) = 0$ *and* $\Sigma(0) = \Sigma(0)^\top = 0$ *describe a deterministic least-squares filter initialised to zero:*

$$r^2(T) = \inf_{(u_0^p, y_0^p) \in \mathcal{N}_p^{[0,T]}(w_2)} (\|u_0^p\|_2^2 + \|y_0^p\|_2^2).$$

Proof The proof of is given in the Appendix. □

Now since

$$\begin{aligned} r^2(T) &= \inf_{(u_0^p, y_0^p) \in \mathcal{N}_p^{[0,T]}(w_2)} \left(\|u_0^p\|_2^2 + \|y_0^p\|_2^2\right) \\ &= \left(\inf\{r \geq 0 \mid r = \|v_0\|_2,\ v_0 \in \mathcal{N}_p^{[0,T]}(w_2)\}\right)^2 = (r_p[T])^2, \end{aligned}$$

the Kalman filter can be utilised to compute X_A in Section 1.1 in the l_2 setting. Note that the computation of the least-squares solution

$$(\tilde{u}_0^p, \tilde{y}_0^p) = \underset{(u_0^p, y_0^p) \in \mathcal{N}_p^{[0,T]}(w_2)}{\operatorname{argmin}} \left(\|u_0^p\|_2^2 + \|y_0^p\|_2^2 \right)$$

at time $T \in \mathbb{N}$ is not recursive, however the residual $r(T)$ of the Kalman filter is determined recursively, hence the computation is feasible.

An important implication of Theorem 3.10 is that we do not have to further show that the Kalman filter satisfies Assumptions 3.4 since by the equality to the least-squares filter (estimator A in l_2) these properties reflect back onto the Kalman filter; the Kalman filter implicitly utilises optimal (least squares) disturbance estimates that are consistent with the plant P_p and the observation w_2 to construct the estimate $\tilde{y}_1$.

5 Disturbance estimation by optimisation methods

With the Kalman filter we have already introduced the only known, realisable solution to the infinite horizon disturbance estimation problem (estimator A), which only applies in L_2, l_2. Limiting the focus to infinite horizon estimation would therefore, by the fact that practical implementations only exist in l_2, essentially reduce the application of the algorithm to l_2. Results in other l_r signal spaces $1 \leq r \leq \infty,\ r \neq 2$ would appear to be of theoretical (non-implementable) interest only.

By considering finite horizon estimation (Estimator B) we can overcome this limitation. The computation of the finite horizon disturbance estimates turns out to be much more approachable since it is a standard optimisation problem with many possible implementations.

To see this, recall that we have to solve the optimisation problem

$$\tilde{w}_0^p \in \inf \left\{ r \geq 0 \mid r = \|v_0\|,\ v_0 \in \mathcal{N}_p^{[b-a,b]}(w_2) \right\},\ a \leq b \in \mathbb{N}$$

for $P_p,\ p \in \mathcal{P}$.

Since P_p is assumed to be observable we can by Polderman and Willems (1997) find a matrix of suitable dimension $G_p : \mathbb{R}^{u \times v}$ such that for all $w_1^p \in \mathcal{M}_p^{[a,b]}$ we have $G_p w_1^p = 0$. Then G_p is called a kernel representation of P_p. Observe that with $w_0^p = w_1^p + \mathscr{R}_{b-a,a} w_2$ we have $G_p w_0^p = G_p \mathscr{R}_{b-a,a} w_2 = b$.

Equivalently we can therefore formulate the constraint optimisation problem in 'standard form' as follows:

$$\text{minimise } f(x) = \|x\|,\ \text{subject to the constraint } G_p x = b, \tag{3.31}$$

where $x = w_0^p$ and $b = G_p \mathscr{R}_{b-a,a} w_2$.

As discussed in Chapter 2 in terms of metric projections, such norm optimisation problems are unique in $l_r,\ 1 < r < \infty$ however not necessarily in $l_1,\ l_\infty$. However, by convexity, every solution to an $l_r,\ 1 \le r \le \infty$ norm optimisation problem is a global solution. For disturbance estimation this non-uniqueness does not matter since every solution will satisfy Assumptions 3.4 and the later analysis will merely require the size of the disturbance estimate, which is equal for all solutions.

In l_2, a solution to the optimisation problem can directly be calculated via the pseudo inverse (Moore-Penrose inverse) G_p^+ of G_p. Let $G_p = U\Sigma V^\top$ be the singular value decomposition of G_p and define the pseudo inverse $G_p^+ = V\Sigma^{-1}U^\top$. Then $x = G_p^+ b$ provides a unique solution to the optimisation problem (e.g. see Boyd and Vandenberghe (2004)) and we obtain $\tilde{w}_0^p = G_p^+ G_p \mathscr{R}_{b-a,a} w_2$ where $G_p^+ G_p$ is the (Euclidean) projection onto $\mathcal{N}_p^{[a,b]}(w_2)$.

In l_1 we can reformulate the optimisation problem in equation (3.31) to:

$$\text{minimise } y, \text{ subject to the constraints } y \ge x,\ y \ge -x,\ G_p x = b$$

where $x = w_0^p$ and $b = G_p \mathscr{R}_{b-a,a} w_2$. Equivalently we can write:

$$\text{minimise } c^\top z, \text{ subject to the constraints } Hz \ge 0,\ Jz = b$$

where $c^\top = \begin{bmatrix} 1 & 0 \end{bmatrix}$, $z = \begin{bmatrix} y \\ x \end{bmatrix}$, $H = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ and $J = \begin{bmatrix} 0 & G_p \end{bmatrix}$.

This can be solved by linear programming with algorithms such as the ellipsoid method, the interior point method or the simplex algorithm (e.g. see Schrijver (1998)). The disturbance estimate can then be computed from the minimising z by letting $\tilde{w}_0^p = \begin{bmatrix} 0 & 1 \end{bmatrix} z$.

In l_∞ we can solve the same linear programming problem:

$$\text{minimise } (c^\top z)(i), \text{ subject to the constraints } Hz \ge 0,\ Jz = b$$

for each $i \in [a, b]$ ($|b - a|$ times) and then take the maximum over the solutions.

In a general $l_r,\ 1 < r < \infty$ norm setting or in the non-linear domain, under appropriate convexity assumptions, there also exist efficient algorithms to solve the optimisation problem, e.g. by gradient descent, the Newton method or geometric programming (e.g. see Boyd and Vandenberghe (2004)).

This poses a major advantage over observer based switching algorithms since, as mentioned in the introduction, the construction of observers for a wide class of non-linear systems is difficult and unclear. In contrast, convex optimisation problems are well studied and many algorithms for various (non-linear) scenarios are readily available.

Chapter 4

Estimation-based Multiple Model Switched Adaptive Control

In this Chapter we will develop the structure of the EMMSAC algorithm. First we will introduce the notion of a controller design procedure K that assigns a corresponding controller to every plant model and formalise the requirement that a controller $C_{K(p)}$ to a plant P_p must be stabilising, i.e. the atomic closed loop $[P_p, C_{K(p)}]$, $\forall p \in \mathcal{P}$ needs to be gain stable. We then state two abstract controller assumptions on which the subsequent analysis will rest. The advantage of such an axiomatic approach is that it clears the analysis of any plant or controller structure (state space matrices, transfer functions, etc.); in fact it is irrelevant how plant and controllers are represented (they can be non-linear), as long as every atomic plant and controller pair fulfils the controller assumptions. This will very much benefit a later generalisation to a wider class of systems.

We will then define a switching signal q based on the estimator introduced in the previous chapter and define the switching controller C at time $k \in \mathbb{N}$ via the controller $C_{K(q(k))}$ corresponding to the plant $P_{q(k)}$.

1 Finite horizon behaviour of the atomic closed loop

A crucial design step for any multiple model type algorithm is to assign stabilising controllers to all plant models. We will do this via the controller design procedure given by a map

$$K : \mathcal{P} \to \mathcal{C}$$

where analog to $\mathcal{P}$ we let $\mathcal{C}$ be a set parametrising a collection of controller operators

$$u_2^c = C_c y_2^c \tag{4.1}$$

for $c \in \mathcal{C}$. For example in the case of linear systems we let $\mathcal{C} = \mathcal{C}_{LTI}$, where

$$\mathcal{C}_{LTI} := \left\{ (A,B,C,D) \in \cup_{n\geq 1}\mathbb{R}^{n\times n} \times \mathbb{R}^{n\times o} \times \mathbb{R}^{m\times n} \times \mathbb{R}^{m\times o} \;\middle|\; \begin{array}{ll} (A,B) & \text{is controllable} \\ (A,C) & \text{is observable} \end{array} \right\}, \tag{4.2}$$

and

$$C_c : \mathcal{Y}_e \to \mathcal{U}_e : y_2^c \mapsto u_2^c,\; c = (A_c, B_c, C_c, D_c) \in \mathcal{C}_{LTI} \tag{4.3}$$

is defined by

$$x_c(k+1) = A_c x_c(k) + B_c y_2^c(k) \tag{4.4}$$

$$u_2^c(k) = C_c x_c(k) + D_c y_2^c(k) \tag{4.5}$$

$$x_c(-k) = 0,\; k \in \mathbb{N}. \tag{4.6}$$

Note the clash of notation in equation (4.3) where C_c denotes both the controller operator and a state space matrix. This, however, is of no further consequence, since we will use either the operator or the state space description and not both at the same time, where the meaning is apparent from the context. Also note that since $x_c(-k) = 0$ for all $k \in \mathbb{N}$ it follows that $u_2^c(-k) = C_c(y_2^c)(-k) = 0$ for all $k \in \mathbb{N}$.

Let $\sigma(c),\; c \in \mathcal{C}$ denote the minimum length of the interval that the signal $(u_2^c, y_2^c)^\top$ needs to be observed to uniquely determine the initial condition of C_c, i.e.

$$\sigma(c) := \min \left\{ k \geq 0 \;:\; \forall l \geq 0,\quad \begin{array}{c} u_2^c = C_c y_2^c,\; \hat{u}_2^c = C_c \hat{y}_2^c, \\ (u_2^c, y_2^c)^\top|_{[l,l+k]} = (\hat{u}_2^c, \hat{y}_2^c)^\top|_{[l,l+k]}, \\ y_2^c = \hat{y}_2^c \Rightarrow u_2^c = \hat{u}_2^c \end{array} \right\}. \tag{4.7}$$

Similarly let $\sigma(p),\; p \in \mathcal{P}$ denote the minimum length of the interval that the signal $(u_1^p, y_1^p)^\top$ needs to be observed to uniquely determine the initial condition of P_p, i.e.

$$\sigma(p) := \min \left\{ k \geq 0 \;:\; \forall l \geq 0,\quad \begin{array}{c} y_1^p = P_p u_1^p,\; \hat{y}_1^p = P_p \hat{u}_1^p, \\ (u_1^p, y_1^p)^\top|_{[l,l+k]} = (\hat{u}_1^p, \hat{y}_1^p)^\top|_{[l,l+k]}, \\ u_1^p = \hat{u}_1^p \Rightarrow y_1^p = \hat{y}_1^p \end{array} \right\}. \tag{4.8}$$

For minimal MIMO LTI systems it can be shown that $\sigma(p) = n_p - 1$ where n_p is the dimension of $A_p \in \mathbb{R}^{n_p \times n_p},\; (A_p, \cdot, \cdot, \cdot) \in \mathcal{P}_{LTI}$ and $\sigma(c) = n_c - 1$ where n_c is the dimension of $A_c \in \mathbb{R}^{n_c \times n_c},\; (A_c, \cdot, \cdot, \cdot) \in \mathcal{C}_{LTI}$.

Instead of giving a particular controller design procedure $K : \mathcal{P} \to \mathcal{C}$ we will now state two general assumptions imposed upon the atomic closed loop systems $[P_p, C_c]$ and $[P_p, C_{K(p)}]$.

Assumption 4.1. *There exist functions*

$$\alpha, \beta : \mathcal{P} \times \mathcal{C} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$$

such that the following holds:

1. *(Linear growth of $[P_p, C_c]$): Let $p \in \mathcal{P},\ c \in \mathcal{C}$ and the closed loop system $[P_p, C_c]$ be well-posed. Let $l_1, l_1, l_2, l_3, l_4 \in \mathbb{N},\ l_1 < l_2 \leq l_3 < l_4$ and $I_1 = [l_1, l_2), I_2 = [l_2, l_3), I_3 = [l_3, l_4)$. Suppose $w_2, w_2^c, w_1^p \in \mathcal{W}_e,\ w_0^p \in \mathcal{W}$ satisfy equations* (3.8)–(3.10),(4.1) *on $I_1 \cup I_2 \cup I_3$. Suppose that either*

   $$w_2^c|_{I_1} = 0,\ w_2^c|_{I_2 \cup I_3} = w_2|_{I_2 \cup I_3}$$

   or

   $$w_2^c|_{I_1 \cup I_2 \cup I_3} = w_2|_{I_1 \cup I_2 \cup I_3}$$

   where

   $$|I_1| = l_2 - l_1 \geq \max\{\sigma(p), \sigma(c)\}. \tag{4.9}$$

   Then, in both cases:

   $$\|w_2|_{I_3}\| \leq \alpha(p, c, |I_2|, |I_3|)\|w_2|_{I_1}\| + \beta(p, c, |I_2|, |I_3|)\|w_0^p|_{I_1 \cup I_2 \cup I_3}\|. \tag{4.10}$$

2. *(Stability of $[P_p, C_{K(p)}]$): Let $p \in \mathcal{P}$ and $x \in \mathbb{N}$. Then*

   $$\alpha(p, K(p), a, x) \to 0 \text{ as } a \to \infty \tag{4.11}$$

   and α is monotonic in a.

Note that the monotonicity requirement in the second assumption follows without loss of generality since any function $\hat{\alpha}$ satisfying equation (4.11) can be dominated point-wise by a monotonic function α satisfying equation (4.11). In the special case of LTI systems we provide an explicit construction of a monotonic α satisfying equation (4.11) from a non-monotonic $\hat{\alpha}$ satisfying equation (4.11).

Assumptions 4.1 now allow the following interpretation: We expect to be able to bound future signals $\|w_2|_{I_3}\|$ by some (linear) function of the system's initial conditions, given by $\|w_2|_{I_1}\|$, and the system's input $w_0^p|_{I_1 \cup I_2 \cup I_3}$ for any well-posed closed loop system $[P_p, C_c]$. This is reflected by equation (4.10). However $\|w_2|_{I_1}\|$ only allows an interpretation as an initial condition if the interval I_1 is sufficiently long. This is reflected by equations (4.7)–(4.9). We will show below that the given assumptions hold for (stabilising) controller design procedures $K : \mathcal{P}_{LTI} \to \mathcal{C}_{LTI}$.

Note that the choice $w_2^c|_{I_1} = 0$ corresponds to an initialisation of the controller to zero at time l_2 and the choice $w_2^c|_{I_1} = w_2|_{I_1}$ corresponds to continued closed loop operation of the same controller. The need for such a construction will become apparent in Chapter 5 in the context of 'virtual' switching times where we do not actually switch from one controller to another but remain with the same controller, hence execute a virtual switch to the same controller.

We will now show that the given assumptions can be met by minimal MIMO LTI systems:

Definition 4.2. *$K : \mathcal{P} \to \mathcal{C}$ is said to be a stabilising design if $[P_p, C_{K(p)}]$ is gain stable for all $p \in \mathcal{P}$.*

Recall that

$$\mathcal{P}_{LTI} := \left\{ (A,B,C,D) \in \cup_{n\geq 1}\mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} \times \mathbb{R}^{o\times n} \times \mathbb{R}^{o\times m} \;\middle|\; \begin{array}{ll} (A,B) & \text{is controllable} \\ (A,C) & \text{is observable} \end{array} \right\}, \tag{3.2}$$

$$\mathcal{C}_{LTI} := \left\{ (A,B,C,D) \in \cup_{n\geq 1}\mathbb{R}^{n\times n} \times \mathbb{R}^{n\times o} \times \mathbb{R}^{m\times n} \times \mathbb{R}^{m\times o} \;\middle|\; \begin{array}{ll} (A,B) & \text{is controllable} \\ (A,C) & \text{is observable} \end{array} \right\} \tag{4.2}$$

and

$$\bar{\mathcal{P}}_{LTI} := \left\{ (A,B,C) \in \cup_{n\geq 1}\mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} \times \mathbb{R}^{o\times n} \;\middle|\; \begin{array}{ll} (A,B) & \text{is controllable} \\ (A,C) & \text{is observable} \end{array} \right\}. \tag{3.24}$$

Also define

$$\bar{\mathcal{C}}_{LTI} := \left\{ (A,B,C) \in \cup_{n\geq 1}\mathbb{R}^{n\times n} \times \mathbb{R}^{n\times o} \times \mathbb{R}^{m\times n} \;\middle|\; \begin{array}{ll} (A,B) & \text{is controllable} \\ (A,C) & \text{is observable} \end{array} \right\}. \tag{4.12}$$

Lemma 4.3. *Let $\mathcal{P}_{LTI}, \mathcal{C}_{LTI}$ be defined by equations (3.2),(4.2) and $\bar{\mathcal{P}}_{LTI}, \bar{\mathcal{C}}_{LTI}$ be defined by equations (3.24),(4.12). Let $K : \mathcal{P} \to \mathcal{C}$ where $(\mathcal{P}, \mathcal{C}) \in \{(\bar{\mathcal{P}}_{LTI}, \mathcal{C}_{LTI}), (\mathcal{P}_{LTI}, \bar{\mathcal{C}}_{LTI})\}$. Then Assumption 4.1(1) holds. Let $K : \mathcal{P} \to \mathcal{C}$ be a stabilising design. Then 4.1(2) holds.*

Proof Since

$$(\mathcal{P}, \mathcal{C}) \in \{(\bar{\mathcal{P}}_{LTI}, \mathcal{C}_{LTI}), (\mathcal{P}_{LTI}, \bar{\mathcal{C}}_{LTI})\},$$

we can let

$$p = (A_p, B_p, C_p, D_p) \in \mathcal{P}_{LTI},\ m, o, n_p = n \in \mathbb{N}$$

and

$$c = (A_c, B_c, C_c, D_c) \in \mathcal{C}_{LTI},\ m, o, n_c = n \in \mathbb{N}$$

where either D_p or D_c is zero.

Let the observability matrix $O_p \in \mathbb{R}^{on_p \times n_p}$ be given by

$$O_p = \begin{bmatrix} C_p \\ C_p A_p \\ \cdots \\ C_p A_p^{n_p-1} \end{bmatrix},$$

the controllability matrix $K_p \in \mathbb{R}^{n_p \times mn_p}$ be given by

$$K_p = \begin{bmatrix} A_p^{n_p-1} B_p & A_p^{n_p-2} B_p & \cdots & A_p B_p & B_p \end{bmatrix}$$

and the input-output matrix $T_p \in \mathbb{R}^{on_p \times mn_p}$ be given by

$$T_p = \begin{bmatrix} D_p & 0 & \cdots & 0 & 0 & 0 \\ C_p B_p & D_p & \cdots & 0 & 0 & 0 \\ C_p A_p B_p & C_p B_p & \ddots & 0 & 0 & 0 \\ C_p A_p^2 B_p & C_p A_p B_p & \cdots & D_p & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ C_p A_p^{n_p-2} B_p & C_p A_p^{n_p-3} B_p & \cdots & C_p A_p B_p & C_p B_p & D_p \end{bmatrix}.$$

Let $k \in \mathbb{N}$. With the controllability matrix K_p and by equations (3.4) we can write

$$x_p(k+1) = A_p^{n_p} x_p(k - n_p + 1) + K_p \cdot \begin{bmatrix} u_1^p(k - n_p + 1) \\ \vdots \\ u_1^p(k) \end{bmatrix}. \tag{4.13}$$

With the observability matrix O_p and equations (3.4),(3.5) we can also write

$$\begin{bmatrix} y_1^p(k+1) \\ \vdots \\ y_1^p(k+n_p) \end{bmatrix} = O_p x_p(k+1) + T_p \begin{bmatrix} u_1^p(k+1) \\ \vdots \\ u_1^p(k+n_p) \end{bmatrix}.$$

Note that since P_p is observable, $O_p \in \mathbb{R}^{on_p \times n_p}$ is rank n_p. Let O_p^+ denote the Moore-Penrose pseudoinverse of O_p. Since all columns of O_p are linearly independent, it follows that $O_p^+ O_p = I$ and hence we can rearrange to give

$$x_p(k-n_p+1) = O_p^+ \left(\begin{bmatrix} y_1^p(k-n_p+1) \\ \vdots \\ y_1^p(k) \end{bmatrix} - T_p \begin{bmatrix} u_1^p(k-n_p+1) \\ \vdots \\ u_1^p(k) \end{bmatrix} \right). \tag{4.14}$$

We can now see that in order to reconstruct the state x_p we have to observe the signals (u_1^p, y_1^p) for $n_p - 1$ time steps. Hence we have that $\sigma(p) = n_p - 1$. Analogously we have that $\sigma(c) = n_c - 1$.

Substituting equation (4.14) in equation (4.13) leads to

$$\begin{aligned} x_p(k+1) &= A_p^{n_p} O_p^+ \left(\begin{bmatrix} y_1^p(k-n_p+1) \\ \vdots \\ y_1^p(k) \end{bmatrix} - T_p \begin{bmatrix} u_1^p(k-n_p+1) \\ \vdots \\ u_1^p(k) \end{bmatrix} \right) \\ &\quad + K_p \begin{bmatrix} u_1^p(k-n_p+1) \\ \vdots \\ u_1^p(k) \end{bmatrix} \\ &= A_p^{n_p} O_p^+ \begin{bmatrix} y_1^p(k-n_p+1) \\ \vdots \\ y_1^p(k) \end{bmatrix} + (K_p - A_p^{n_p} O_p^+ T_p) \begin{bmatrix} u_1^p(k-n_p+1) \\ \vdots \\ u_1^p(k) \end{bmatrix} \end{aligned}$$

and therefore

$$|x_p(k+1)| \leq Y_p \left\| w_1^p|_{[k-n_p+1,k]} \right\| \leq Y_p \left\| w_2|_{[k-n_p+1,k]} \right\| + Y_p \left\| w_0^p|_{[k-n_p+1,k]} \right\| \tag{4.15}$$

where

$$Y_p = \left\| \begin{bmatrix} A_p^{n_p} O_p^+ & K_p - A_p^{n_p} O_p^+ T_p \end{bmatrix} \right\|.$$

Analogously for $c \in \mathcal{C}$ we have

$$|x_c(k+1)| \leq Y_c \left\| w_2^c|_{[k-n_c+1,k]} \right\| \tag{4.16}$$

where

$$Y_c = \left\| \begin{bmatrix} A_c^{n_c} O_c^+ & K_c - A_c^{n_c} O_c^+ T_c \end{bmatrix} \right\|.$$

For the closed loop $[P_p, C_c]$ we have by equations (3.4),(4.4) that

$$\begin{bmatrix} x_p(k+1) \\ x_c(k+1) \end{bmatrix} = \begin{bmatrix} A_p & 0 \\ 0 & A_c \end{bmatrix} \begin{bmatrix} x_p(k) \\ x_c(k) \end{bmatrix} + \begin{bmatrix} B_p & 0 \\ 0 & B_c \end{bmatrix} \begin{bmatrix} u_1^p(k) \\ y_2(k) \end{bmatrix}. \tag{4.17}$$

With equations (3.9),(3.10),

$$\begin{bmatrix} u_1^p(k) \\ y_2(k) \end{bmatrix} = \begin{bmatrix} u_0^p(k) \\ y_0^p(k) \end{bmatrix} - \begin{bmatrix} u_2(k) \\ y_1^p(k) \end{bmatrix}$$

and equations (3.5),(4.5),

$$\begin{bmatrix} u_2(k) \\ y_1^p(k) \end{bmatrix} = \begin{bmatrix} 0 & C_c \\ C_p & 0 \end{bmatrix} \begin{bmatrix} x_p(k) \\ x_c(k) \end{bmatrix} + \begin{bmatrix} 0 & D_c \\ D_p & 0 \end{bmatrix} \begin{bmatrix} u_1^p(k) \\ y_2(k) \end{bmatrix},$$

we have

$$\begin{bmatrix} u_1^p(k) \\ y_2(k) \end{bmatrix} = \begin{bmatrix} u_0^p(k) \\ y_0^p(k) \end{bmatrix} - \begin{bmatrix} 0 & C_c \\ C_p & 0 \end{bmatrix} \begin{bmatrix} x_p(k) \\ x_c(k) \end{bmatrix} - \begin{bmatrix} 0 & D_c \\ D_p & 0 \end{bmatrix} \begin{bmatrix} u_1^p(k) \\ y_2(k) \end{bmatrix}.$$

Furthermore by adding $\begin{bmatrix} 0 & D_c \\ D_p & 0 \end{bmatrix} \begin{bmatrix} u_1^p(k) \\ y_2(k) \end{bmatrix}$ we obtain

$$\begin{bmatrix} I & D_c \\ D_p & I \end{bmatrix} \begin{bmatrix} u_1^p(k) \\ y_2(k) \end{bmatrix} = \begin{bmatrix} u_0^p(k) \\ y_0^p(k) \end{bmatrix} - \begin{bmatrix} 0 & C_c \\ C_p & 0 \end{bmatrix} \begin{bmatrix} x_p(k) \\ x_c(k) \end{bmatrix}. \tag{4.18}$$

Observe that since by assumption either D_p or D_c is zero we have

$$\begin{bmatrix} I & D_c \\ D_p & I \end{bmatrix}^{-1} = \begin{bmatrix} I & -D_c \\ -D_p & I, \end{bmatrix}. \tag{4.19}$$

Multiplying inequality (4.18) with equation (4.19) from the left yields:

$$\begin{bmatrix} u_1^p(k) \\ y_2(k) \end{bmatrix} = \begin{bmatrix} I & -D_c \\ -D_p & I \end{bmatrix} \begin{bmatrix} u_0^p(k) \\ y_0^p(k) \end{bmatrix} - \begin{bmatrix} -D_cC_p & C_c \\ C_p & -D_pC_c \end{bmatrix} \begin{bmatrix} x_p(k) \\ x_c(k) \end{bmatrix}. \tag{4.20}$$

Substitution into equation (4.17) gives us

$$
\begin{aligned}
\underbrace{\begin{bmatrix} x_p(k+1) \\ x_c(k+1) \end{bmatrix}}_{x(k+1)} &= \begin{bmatrix} A_p & 0 \\ 0 & A_c \end{bmatrix} \begin{bmatrix} x_p(k) \\ x_c(k) \end{bmatrix} + \begin{bmatrix} B_p & 0 \\ 0 & B_c \end{bmatrix} \\
&\quad \cdot \left(\begin{bmatrix} I & -D_c \\ -D_p & I \end{bmatrix} \begin{bmatrix} u_0^p(k) \\ y_0^p(k) \end{bmatrix} - \begin{bmatrix} -D_c C_p & C_c \\ C_p & -D_p C_c \end{bmatrix} \begin{bmatrix} x_p(k) \\ x_c(k) \end{bmatrix} \right) \\
&= \begin{bmatrix} A_p & 0 \\ 0 & A_c \end{bmatrix} \begin{bmatrix} x_p(k) \\ x_c(k) \end{bmatrix} + \begin{bmatrix} B_p & -B_p D_c \\ -B_c D_p & B_c \end{bmatrix} \begin{bmatrix} u_0^p(k) \\ y_0^p(k) \end{bmatrix} \\
&\quad - \begin{bmatrix} -B_p D_c C_p & B_p C_c \\ B_c C_p & -B_c D_p C_c \end{bmatrix} \begin{bmatrix} x_p(k) \\ x_c(k) \end{bmatrix} \\
&= \underbrace{\begin{bmatrix} A_p + B_p D_c C_p & -B_p C_c \\ -B_c C_p & A_c + B_c D_p C_c \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_p(k) \\ x_c(k) \end{bmatrix}}_{x(k)} \\
&\quad + \underbrace{\begin{bmatrix} B_p & -B_p D_c \\ -B_c D_p & B_c \end{bmatrix}}_{B} \underbrace{\begin{bmatrix} u_0^p(k) \\ y_0^p(k) \end{bmatrix}}_{w_0^p(k)} . \qquad (4.21)
\end{aligned}
$$

From equation (4.20) we have

$$
\begin{aligned}
\begin{bmatrix} u_2(k) \\ y_2(k) \end{bmatrix} &= \begin{bmatrix} -u_1^p(k) \\ y_2(k) \end{bmatrix} + \begin{bmatrix} u_0^p(k) \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} -I & D_c \\ -D_p & I \end{bmatrix} \begin{bmatrix} u_0^p(k) \\ y_0^p(k) \end{bmatrix} - \begin{bmatrix} D_c C_p & -C_c \\ C_p & -D_p C_c \end{bmatrix} \begin{bmatrix} x_p(k) \\ x_c(k) \end{bmatrix} + \begin{bmatrix} u_0^p(k) \\ 0 \end{bmatrix} \\
&= \underbrace{\begin{bmatrix} -D_c C_p & C_c \\ -C_p & D_p C_c \end{bmatrix}}_{C} \begin{bmatrix} x_p(k) \\ x_c(k) \end{bmatrix} + \underbrace{\begin{bmatrix} 0 & D_c \\ -D_p & I \end{bmatrix}}_{D} \begin{bmatrix} u_0^p(k) \\ y_0^p(k) \end{bmatrix}
\end{aligned}
$$

and therefore

$$
\begin{aligned}
x(k+1) &= Ax(k) + Bw_0^p(k) \\
w_2(k) &= Cx(k) + Dw_0^p(k),
\end{aligned}
$$

where with $\rho = n_p + n_c$, $A \in \mathbb{R}^{\rho \times \rho}, B \in \mathbb{R}^{\rho \times (m+o)},\ C \in \mathbb{R}^{(m+o) \times \rho},\ D \in \mathbb{R}^{(m+o) \times (m+o)}$.

Let $l_1, l_1, l_2, l_3, l_4 \in \mathbb{N}$, $l_1 < l_2 \leq l_3 < l_4$ and $I_1 = [l_1, l_2), I_2 = [l_2, l_3), I_3 = [l_3, l_4)$. We now initialise the controller either with $x_c(l_2) = 0$ or $x_c(l_2) \neq 0$.

For $w_2^c|_{I_1} = w_2|_{I_1}$ we have with

$$
|I_1| \geq \max\{\sigma(p), \sigma(c)\} = \max\{n_p - 1, n_c - 1\}
$$

that $x_p(l_2)$ and $x_c(l_2)$ are uniquely defined and therefore by equations (4.15),(4.16)

$$\begin{aligned}\|x(l_2)\| &= \|x_p(l_2), x_c(l_2)\| \\ &\leq Y_p \|w_2|_{I_1}\| + Y_p \|w_0^p|_{I_1}\| + Y_c \|w_2|_{I_1}\| \\ &\leq (Y_p + Y_c) \|w_2|_{I_1}\| + Y_p \|w_0^p|_{I_1}\| .\end{aligned}$$

Analogously for $w_2^c|_{I_1} = 0$ there follows from equation (4.16) that $x_c(l_2) = 0$ hence

$$\begin{aligned}\|x(l_2)\| &= \|x_p(l_2), x_c(l_2)\| \\ &\leq Y_p \|w_2|_{I_1}\| + Y_p \|w_0^p|_{I_1}\| .\end{aligned}$$

So in either case

$$\|x(l_2)\| \leq (Y_p + Y_c) \|w_2|_{I_1}\| + Y_p \|w_0^p|_{I_1}\| . \tag{4.22}$$

We also have with

$$\begin{aligned}K_a &= \begin{bmatrix} A^{a-1}B & A^{a-2}B & \cdots & AB & B \end{bmatrix}, \ a \geq 0 \\ O_a &= \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{a-1} \end{bmatrix}, \ a \geq 0 \\ T_a &= \begin{bmatrix} D & 0 & \cdots & 0 & 0 & 0 \\ CB & D & \cdots & 0 & 0 & 0 \\ CAB & CB & \ddots & 0 & 0 & 0 \\ CA^2B & CAB & \cdots & D & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ CA^{a-2}B & CA^{a-3}B & \cdots & CAB & CB & D \end{bmatrix}, \ a \geq 0\end{aligned}$$

that

$$x(l_3) = A^{|I_2|}x(l_2) + K_{|I_2|} \begin{bmatrix} w_0^p(l_2) \\ \vdots \\ w_0^p(l_3 - 1) \end{bmatrix},$$

$$\begin{bmatrix} w_2(l_3) \\ \vdots \\ w_2(l_4 - 1) \end{bmatrix} = O_{|I_3|}x(l_3) + T_{|I_3|} \begin{bmatrix} w_0^p(l_3) \\ \vdots \\ w_0^p(l_4 - 1) \end{bmatrix}.$$

Substitution leads to

$$\begin{aligned}\begin{bmatrix} w_2(l_3) \\ \vdots \\ w_2(l_4-1)\end{bmatrix} &= O_{|I_3|}x(l_3) + T_{|I_3|}\begin{bmatrix} w_0^p(l_2) \\ \vdots \\ w_0^p(l_3-1)\end{bmatrix} \\ &= O_{|I_3|}A^{|I_2|}x(l_2) + \begin{bmatrix} O_{|I_3|}K_{|I_2|} & T_{|I_3|}\end{bmatrix}\begin{bmatrix} w_0^p(l_2) \\ \vdots \\ w_0^p(l_4-1)\end{bmatrix}.\end{aligned}$$

Taking norms and substituting equation (4.22) leads to

$$\begin{aligned}\|w_2|_{I_3}\| &\leq \|O_{|I_3|}A^{|I_2|}\|\|x(l_2)\| + \left(\|O_{|I_3|}K_{|I_2|}\| + \|T_{|I_3|}\|\right)\|w_0^p|_{I_2\cup I_3}\| \\ &\leq \|O_{|I_3|}A^{|I_2|}\|(Y_p+Y_c)\|w_2|_{I_1}\| \\ &\quad +(\|O_{|I_3|}A^{|I_2|}\|Y_p + \|O_{|I_3|}K_{|I_2|}\| + \|T_{|I_3|}\|)\|w_0^p|_{I_1\cup I_2\cup I_3}\|.\end{aligned}$$

We therefore arrive at

$$\|w_2|_{I_3}\| \leq \hat{\alpha}(p,c,|I_2|,|I_3|)\|w_2|_{I_1}\| + \beta(p,c,|I_2|,|I_3|)\|w_0^p|_{I_1\cup I_2\cup I_3}\|$$

where

$$\begin{aligned}\hat{\alpha}(p,c,|I_2|,|I_3|) &= \|O_{|I_3|}A^{|I_2|}\|(Y_p+Y_c) \\ \beta(p,c,|I_2|,|I_3|) &= (\|O_{|I_3|}A^{|I_2|}\|Y_p + \|O_{|I_3|}K_{|I_2|}\| + \|T_{|I_3|}\|).\end{aligned}$$

Hence Assumption 4.1(1) holds.

If K is stabilising design, then $[P_p, C_{K(p)}]$, $p \in \mathcal{P}$ is stable. This implies that with A defined as in equation (4.21):

$$A = \begin{bmatrix} A_p & -B_pC_{K(p)} \\ -B_{K(p)}C_p & A_{K(p)} + B_{K(p)}D_pC_{K(p)}\end{bmatrix}$$

it follows that A is a stable matrix and

$$\|A^a\| \to 0 \; for \; a \to \infty.$$

Therefore

$$\hat{\alpha}(p, K(p), |I_2|, |I_3|) \to 0 \text{ as } |I_2| \to \infty.$$

Although $\hat{\alpha}$ does converge for large $|I_2|$ it is not monotonic in $|I_2|$ in general.

Let $t, x \in \mathbb{N}$. Since $\hat{\alpha}(p, K(p), t, x) \to 0$ for $t \to \infty$, for all $N > 0$ there exist times t_i such that $\hat{\alpha}(p, K(p), s, x) < 1/N$ for all $s \geq t_i$. Therefore for all $t \in [t_{i-1}, t_i]$ we let

$$\hat{\alpha}(p, K(p), t, x) \leq \alpha(p, K(p), t, x) = \max_{t_{i-1} \leq s \leq t_i} \hat{\alpha}(p, K(p), s, x),$$

hence Assumption 4.1(2) holds as required. □

We will now utilise the established definition of a (stabilising) design procedure K to define the switching controller C.

2 The switching algorithm

We noted in the introduction that for infinite horizon disturbance estimation the size of the disturbance estimate or the residual (given by $X_A(w_2)(k)(p)$) can be thought of as the distance between the observation $w_2 \in \mathcal{W}_e$ and the plant $P_p,\ p \in \mathcal{P}$. Alternatively $X_A(w_2)(k)(p)$ can be thought of as a measure of how likely the observation $w_2 \in \mathcal{W}_e$ is explained by the plant P_p.

The intuitive choice for the switching strategy is therefore to define the switching signal $q_f(k),\ k \in \mathbb{N}$ as a pointer to the plant $p \in \mathcal{P}$ which is closest to the observation $\mathscr{T}_k w_2$ in the sense that $X_A(w_2)(k)(p)$ is minimal at time $k \in \mathbb{N}$. Therefore $q_f(k)$ points to the plant whose corresponding estimator is able to explain the observation $\mathscr{T}_k w_2$ with minimum disturbance. Note that the size of the finite horizon disturbance estimate $X_B(w_2)(k)(p) = NE_B(w_2)(k)(p)$ from Chapter 3, Section 2.2 does not directly represent the distance between the observation and the plant, since the structure of $E_B(w_2)(k)(p)$ is different. However, $X_B(w_2)(k)(p)$ preserves a *notion* of distance that appears to be sufficient for the argument.

We will now formally introduce the estimation-based switching algorithm — as depicted in Figure 4.1 — where D is the 'delay' operator, G is the 'plant-generating operator' and M is the minimisation operator which has the purpose to return the plant model which corresponding residual is minimal.

Initially consider the 'free' causal switching operator

$$S_f : \mathcal{W}_e \to map(\mathbb{N}, \mathcal{P}^*) : w_2 \mapsto q_f \tag{4.23}$$

via the combination of the estimation operator E, the norm operator N, the minimisation operator M and the plant-generating operator G with

$$S_f = M(NE, G) = M(X, G), \tag{4.24}$$

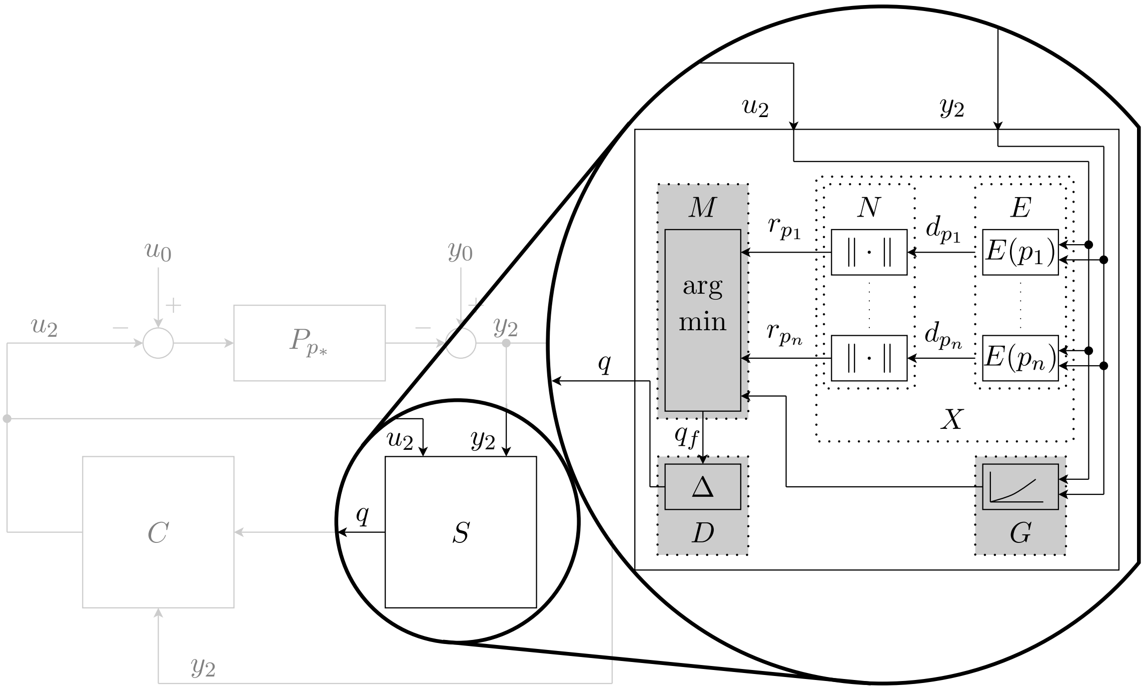

Figure 4.1: Magnified switching strategy S

where $N, E, X,\ X = NE$ have been defined in the previous section and G, M, D are defined below.

The plant-generating operator G is intended to specify the (candidate) plant set available to the algorithm at time $k \in \mathbb{N}$ and is defined as follows.

Let $\mathcal{P}^*$ be the powerset of $\mathcal{P}$. Let $\emptyset \neq \mathcal{P}_i \in \mathcal{P}^*,\ i \in \mathbb{N}$.

Definition 4.4. *A map $Q : \mathcal{W}_e \to map(\mathbb{N}, \mathcal{P}^*)$ is said to be a plant-generating operator if it is causal and satisfies*

$$Q(w_2)(0) = \mathcal{P}_1,\ Q(w_2)(k) = \mathcal{P}_{i(k)},\ k \in \mathbb{N}$$

for some $i : \mathbb{N} \to \mathbb{N}$ with $i(0) = 1$. Q is said to be finite if $\mathcal{P}_i$ is a finite set for all $i \in \mathbb{N}$, constant if $\mathcal{P}_i = \mathcal{P}_j,\ \forall i, j \in \mathbb{N}$ and compact if $\mathcal{P}_i$ is compact for all $i \in \mathbb{N}$.

Now let

$$G : \mathcal{W}_e \to map(\mathbb{N}, \mathcal{P}^*) \tag{4.25}$$

be a plant-generating operator, where we also define

$$\mathcal{P}^G := \bigcup_{w_2 \in \mathcal{W}_e} \bigcup_{k \in \mathbb{N}} G(w_2)(k) \subset \mathcal{P}.$$

$\mathcal{P}^G$ is the union of all plant model sets possibly represented by G. To improve readability we write $G(k) := G(w_2)(k),\ k \in \mathbb{N}$.

We often let G be constant however the algorithm then becomes conservative — as discussed in the introduction and later in Chapter 6. This motivates the time-varying, dynamic nature of G, as used in dynamic EMMSAC.

Definition 4.5. *An EMMSAC algorithm with an underlying plant-generating operator G that is:*

- *time-varying, i.e. there exist $i, j \in \mathbb{N}$ such that $G(i) \neq G(j)$, is said to be dynamic.*
- *constant, i.e. $G(i) = G(j)$ for all $i, j \in \mathbb{N}$, is said to be static.*

On the first pass of reading this document it is recommended to the reader to only consider the static EMMSAC case. A deeper discussion of G is conducted in the next section.

Since we intend to define the free switching signal q_f such that it points to the plant in the plant model set whose corresponding residual is minimal, we introduce the minimising operator M as follows. Let

$$M : (map(\mathbb{N}, map(\mathcal{P}, \mathbb{R}^+)), map(\mathbb{N}, \mathcal{P}^*)) \to map(\mathbb{N}, \mathcal{P}^*) \tag{4.26}$$

and

$$\left[k \mapsto (p \mapsto r_p[k]), k \mapsto G(k)\right] \mapsto \left[k \mapsto q_f(k)\right] \tag{4.27}$$

where

$$q_f(k) := \operatorname*{argmin}_{p \in G(k)} r_p[k], \ \forall k \in \mathbb{N}. \tag{4.28}$$

If there are multiple minimising residuals, an arbitrary ordering on $G(k)$ is imposed a priori, i.e. $G(k) = \{p_1, p_2, \cdots, p_n\}$, and $\operatorname{argmin}_{p \in G(k)} r_p[k]$ is defined to return the parameter $p_i \in G(k)$ with the smallest index i such that $r_{p_i}[k]$ is minimal.

Equation (4.28) also includes the implicit assumption that a minimiser exists. In the scenario considered in this thesis, whereby G is finite or G is compact and $p \mapsto r_p[k]$ is continuous, this holds.

The undelayed, 'free' switching signal $q_f(k)$ at time $k \in \mathbb{N}$ therefore is a direct function of the size of the residuals $r_p[k]$, $p \in G(k)$. Due to disturbances acting on the system, i.e. $w_0 \neq 0$, the switching signal q_f might not settle but switch rapidly between members of G. Since we would like to utilise the switching signal for controller selection this is undesirable as it can lead to instability.

For example consider a switched linear system given by $x_p(k+1) = A_p x(k) + B_p u(k)$ where $(A_p, B_p) \in \{(A_{p_1}, B_{p_1}), (A_{p_2}, B_{p_2})\}$ and $(A_{p_1}, B_{p_1}, \cdot, \cdot), (A_{p_2}, B_{p_2}, \cdot, \cdot) \in \mathcal{P}_{LTI}$ are of compatible dimension. It can be shown that there exist stable A_{p_1}, A_{p_2} and a sufficiently fast periodic switching sequence between (A_{p_1}, B_{p_1}) and (A_{p_2}, B_{p_2}) such that the system is unstable (see Liberzon (2003)).

The purpose of D is now to delay the free switching signal q_f long enough to prevent instability effects caused by rapid switching and to ensure the overall convergence of the closed loop signals. For that purpose we will associate a minimum delay $\Delta(p)$ to every plant $P_p,\ p \in \mathcal{P}$ which must elapse before another switch is permitted. We will encode this information into the 'transition delay' function

$$\Delta : \mathcal{P} \to \mathbb{N}. \tag{4.29}$$

This leads to the following definition of the delay operator D. Define

$$D : map(\mathbb{N}, \mathcal{P}) \to map(\mathbb{N}, \mathcal{P}) \tag{4.30}$$

by

$$[k \mapsto q_f(k)] \mapsto [k \mapsto q(k)] \tag{4.31}$$

where $q(k)$ is defined recursively:

$$q(k) := \begin{cases} q_f(k) & \text{if} \quad k - k_s(k) \geq \Delta(q(k_s(k))) \\ q(k_s(k)) & \text{else} \end{cases} \tag{4.32}$$

and where $k_s : \mathbb{N} \to \mathbb{N}$ is given by

$$k_s(k) := \max\{i \in \mathbb{N} \mid 0 \leq i \leq k,\ q(i) \neq q(i-1)\}. \tag{4.33}$$

Note that $k_s(k)$ returns the last time up to time $k \in \mathbb{N}$ where the algorithm switches from one plant to another. Also note that D is causal.

We are now in the position to define the global switching operator

$$\begin{aligned} S : \mathcal{W}_e &\to map(\mathbb{N}, \mathcal{P}^*) : w_2 \mapsto q \\ S &= DM(NE, G) = DM(X, G) \end{aligned}$$

as depicted in Figure 4.1 where we note that $S = DS_f$ and S_f is the free switching operator as given in equations (4.23),(4.24).

Let a controller design procedure $K : \mathcal{P} \to \mathcal{C}$ be given. The switching controller C is then defined via the switching signal q and the controller design procedure K in the following way: At every time instance $k \in \mathbb{N}$ the atomic controller, defined by $C_{K(q(k))}$, is put into closed loop — as depicted in Figure 4.2. However, since we allow the atomic controllers to have memory, we also have to define an initial condition at the switching time. We therefore let

$$C : \mathcal{Y}_e \to \mathcal{U}_e : y_2 \mapsto u_2 \tag{4.34}$$

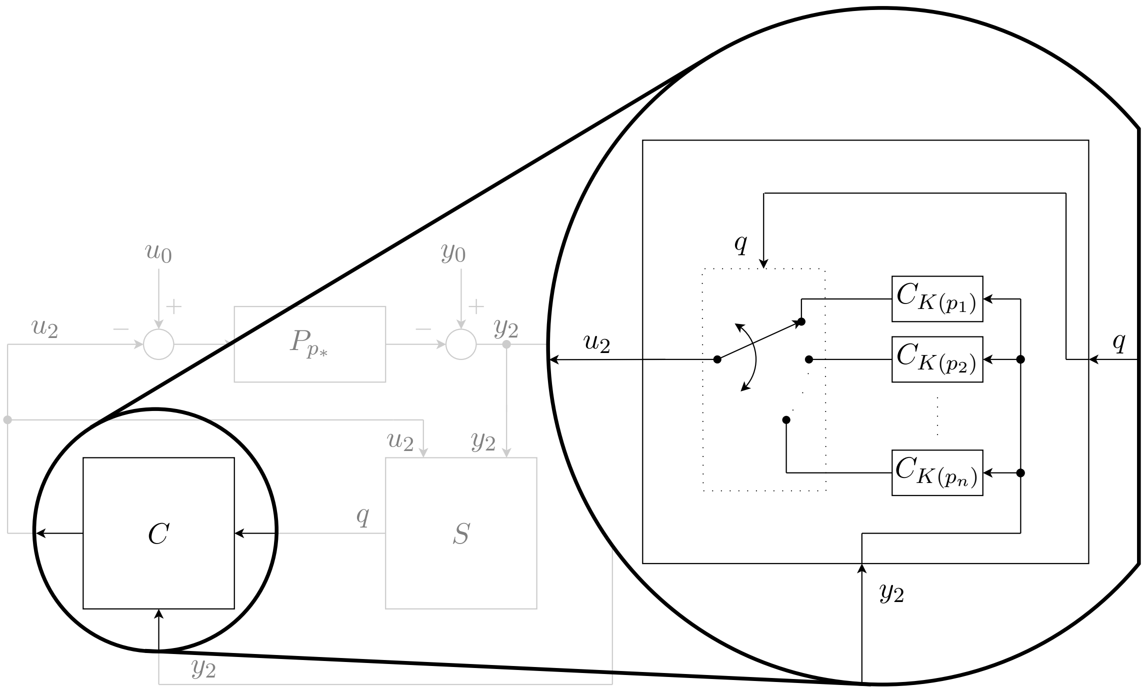

Figure 4.2: Magnified switching controller C

for all $k \in \mathbb{N}$ be defined by

$$u_2(k) = C_{K(q(k))}(y_2 - \mathscr{T}_{k_s(k)-1}y_2)(k), \tag{4.35}$$

where we recall that $k_s(k)$ is the last time $i \in \mathbb{N},\ i \leq k$ s.t. $q(i) \neq q(i-1)$. Equation (4.35) ensures a zero initial condition for the atomic controller $C_{K(q(k))}$ when it is switched into closed loop. Note that if E satisfies Assumption 3.4(1) (causality) and G is causal, then S is causal.

We therefore arrive at Figure 4.3 where all involved sub systems have been defined.

3 The plant-generating operator G

To shed some light on the role of the plant-generating operator G and to emphasise the ample design freedom we enjoy in EMMSAC, we will now briefly discuss a selection of algorithms for the construction of G and note that some of these ideas will be followed up in Chapter 6.

- Static EMMSAC:
  The standard approach in multiple model control, e.g. in Morse (1996, 1997), is to choose a constant plant model set and a corresponding controller set from which the algorithm may select controllers. Although the simplicity of this approach has

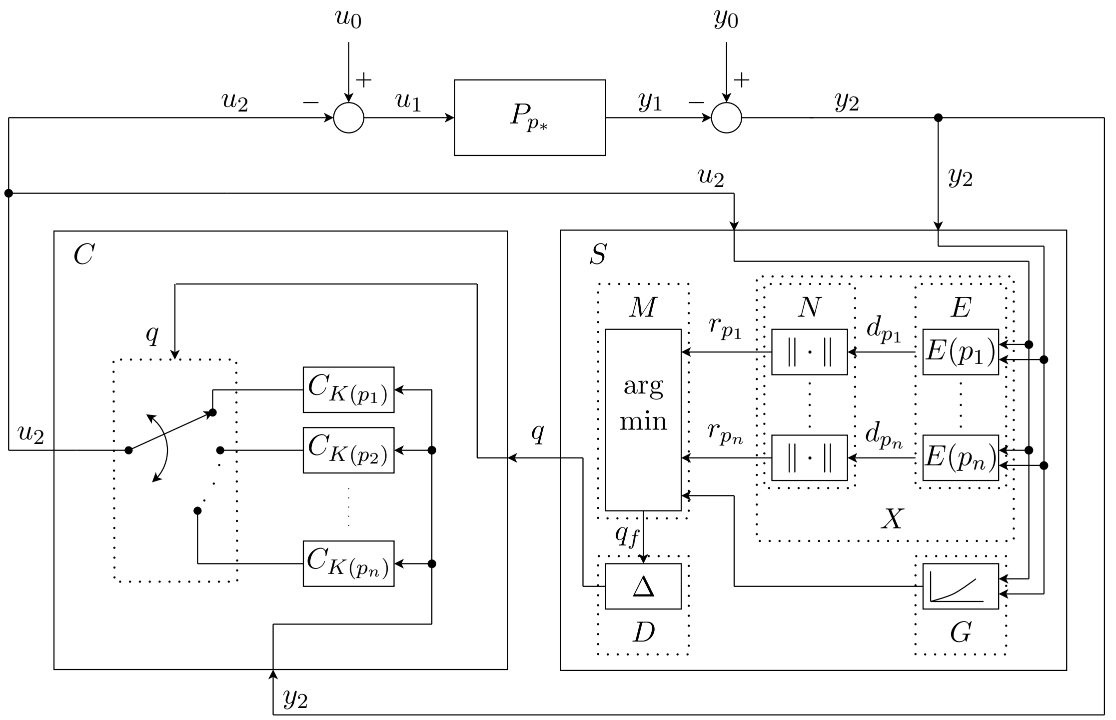

FIGURE 4.3: EMMSAC in detail

its virtue, the resulting algorithm is conservative as shown in Chapter 6, Theorem 6.6.

- Dynamic EMMSAC - expansion of G:

  The complication that the plant-generating operator G is allowed to be time-varying may initially appear unnecessarily complex to the reader. However, it is the device that brings unheard of freedom to the design process and by which we can make the algorithm universal. The mechanism which we will exploit in Theorem 6.6 — to show that the algorithm is conservative — is that the algorithm can be confused by appropriate choices of disturbances (of arbitrarily small size) to switch to the atomic controller in the controller set with the highest gain. Note that we have to ensure that at least one of the plant models in a constant plant model set G is close to the true plant $P = P_{p_*}$, so that its corresponding controller can have a stabilising effect on P. An increasingly large uncertainty in the plant will therefore necessarily lead to an increasingly large constant plant model set G and thus to an increasingly large corresponding controller set (typically incorporating controllers with increasingly high gains). This is enough to show that for constant G, the performance of the algorithm degrades with an increasingly large level of uncertainty in the true plant P — the algorithm is conservative.

  A remedy to this problem is to define a time-varying plant-generating operator G that specifies a plant model set that is initially small but expanded over time, i.e. the algorithm is initially only allowed to choose from a small number of plant models. This essentially eliminates the possibility that the algorithm switches to the worst case controller right away. We will then define a performance-orientated rule on how to expand G. In particular, G is expanded as long as performance is degrading, and expansion of G is stopped if the performance is converging. The intuitive idea behind such a choice of G is that the plant model set is expanded until a controller that stabilises the true plant $P = P_{p*}$ enters G. The hope is that if performance converges then G converges. And indeed, we will show in Chapter 6 that dynamic expansion of G is the device by which the algorithm achieves gain function bounds invariant to the level of uncertainty — that it is universal.

- Dynamic EMMSAC - refinement of G:
  Although a well chosen constant, coarse plant model set G may be sufficient to provide a stabilising controller for the true plant $P = P_{p*}$ with a bounded uncertainty, we expect the performance to diminish for an increasing distance between the true plant P and the closest plant model $P_p,\ p \in G$, since the corresponding controller $C_{K(p)}$ may only be mildly stabilising for P.

  By choosing a constant, dense plant model set G we expect the overall performance to be better, however this approach requires the implementation of a larger number of estimators where most of them will never produce a residual that is minimal; hence such a construction is usually conservative from a implementational point of view (see Chapter 6, Section 8). Also an overly dense plant model set G may lead to 'oversampling' effects, analogously to the 'over-fitting' of functions with too many control points. Although such effect may increase the actual closed loop gain, they do not do so unboundedly (see Chapter 5, Theorem 5.14 which establishes an upper bound on the closed loop gain; also if G is a compact continuum.)

  One possibility to address these problems is to have a time-varying G, initially containing only few plant models which form a coarse grid over the uncertainty set of P (ensuring stability) and then to refine G over time. A brute force method of doing so would be to introduce more and more plant models distributed uniformly over the uncertainty set of P.

  Usually there is probabilistic information available about the uncertainty problem in the sense that some uncertainties are more likely than others. For example if a manufacturing process is producing items of a mass m, we usually expect the uncertainty to form, for example, a Gaussian probability distribution around m, and therefore it is more likely that an item has a mass close to m. We could then utilise this information to refine the plant model set and add plant models in an increasingly large neighbourhood around m in order to increase the expected value of the average performance.

- Dynamic EMMSAC - advanced algorithms:
  The above methods to introduce new plant models to G are rather basic. More complex dynamic refinement schemes could include a local search for the smallest disturbance estimate — as depicted in Figure 4.4. For example, assume that

G consists initially of a coarse grid of plant models (Figure 4.4 (A)) which is sufficiently dense for the EMMSAC algorithm to be stabilising. We could then interpolate the position of a potentially better fitting plant model from the n smallest disturbance estimates and the corresponding plant models and implement its corresponding estimator on-line (Figure 4.4 (B)). This process can then be repeated (Figure 4.4 (C)) until, for example, the performance is satisfactory or the algorithm is only introducing plant models that are very close to each other and we can expect to have reached a minimum.

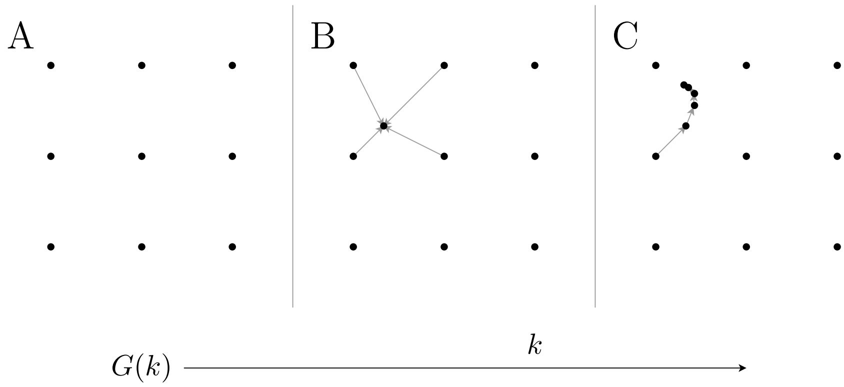

FIGURE 4.4: Local search via interpolation

An alternative approach would be to compute a local gradient from the n smallest disturbance estimates and then consecutively add plant models along this gradient to G (gradient descent). This approach has the advantage that it can project outside the initial plant model set hence is also potentially suitable to orchestrate the enlargement of G.

Essentially any performance-driven search scheme which provides a stabilising $G(k)$ for all $k \in \mathbb{N}$ can be incorporated in EMMSAC.

Such schemes are very interesting in this context since the problem of local minima is not an issue here: the search is only conducted locally where the global algorithm can switch to any plant model in the plant model set G. Hence the desired multiple model control properties, such as simultaneous stabilisation and control of non-convex sets, are preserved.

A scheme that is loosely related to the dynamic refinement of G is the one of Narendra et al. (1995) and Narendra and Balakrishnan (1997), where the authors utilise multiple model switched adaptive control as the global framework, but implement some atomic controllers as classical adaptive controllers to improve performance. The underlying idea is similar: the improvement of the expected performance by supplying an appropriate controller for the region of the uncertainty where the algorithm is expecting the true plant. This is achieved by the tuning behaviour of the atomic adaptive controllers. In contrast to our approach, where the multiple model scheme is used for local refinement, Narendra et. al. introduce additional complexity with the adaptive tuning scheme. Also, as mentioned in the introduction, such an approach is problematic since we will necessarily find ourselves confronted with the usual robust stability and structural limitations of the classical adaptive controller.

4 EMMSAC in practice

This section is intended to give the reader an idea of how to 'run' the algorithm in practice and what implementational complexity is to be expected.

Assume that we have decided upon some (possibly time-varying) construction of a plant-generating operator G (see Chapter 6 for a performance-orientated guideline on how to make that choice). The next step is then to determine a corresponding controller $K(p)$ and a transition delay $\Delta(p)$ to all $p \in \mathcal{P}^G$. Observe that the construction of $K(p),\ \Delta(p),\ p \in \mathcal{P}^G$ can only be computationally feasible if $\mathcal{P}^G$ is finite and a priori known ($G(k) = G(w_2)(k)$ is known for all $w_2 \in \mathcal{W}_e$ and for all $k \in \mathbb{N}$), since we would otherwise have to compute $K(p),\ \Delta(p)$ for all $p \in \mathcal{P} \supset \mathcal{P}^G$ (which may be a continuum). If $\mathcal{P}^G$ is indeed known, small and finite, then off-line computation of $K(p),\ \Delta(p)$ for all $p \in \mathcal{P}^G$ appears feasible and one could store this information in memory to be employed by the algorithm when in on-line operation. However, $\mathcal{P}^G$ is usually unknown if $G(k)$ is a function of observed signals, e.g. if $G(k)$ describes an advanced refinement scheme (as discussed in Chapter 4, Section 3) and off-line computation of K and Δ is not feasible.

To overcome this problem we now consider the on-line computation of the controller as well as the transition delay. It is important to note that only one controller is active at a time, hence only a single controller and corresponding delay needs to be calculated every time the algorithm performs a switch. This implies that calculating the controller and delay on-line reduces the (possibly infinitely large) computational complexity of determining K and Δ off-line to a single computational operation every time a switch occurs. We can therefore trade off memory size and computational off-line resource versus computational on-line resource, or even implement hybrid schemes.

In principle K and Δ can be any operator satisfying Assumptions 4.1 (note that later in Chapter 5, inequality (5.16) will also constrain Δ), hence even the construction of each $K(p),\ \Delta(p),\ p \in \mathcal{P}^G$ by hand is possible. However, manual construction will not be feasible in many situations, i.e. if $\mathcal{P}^G$ is large or unknown. We will therefore assume that K and Δ are determined by some automated procedure. Automated design procedures for K and Δ can for example be implemented by using (the code from) suitable MATLAB toolboxes with the purpose to automatically construct a stabilising H_∞, LQG, PID controller, or some iterative method to determine Δ that satisfies inequality (5.16) in Chapter 5 (given some $l : \mathcal{P} \to \mathbb{R}^+$). The challenge for the designer then reduces to the problem of setting suitable parameters for the automation.

With Δ and K, determined on- or off-line, we now proceed as follows.

1. Construct the estimators:
   Depending on the signal space we can, for example, choose from Kalman filter estimators or finite horizon estimators. Other estimation algorithms are allowed as long as they satisfy Assumptions 3.4. Note that the Kalman filter implementation as well as the finite horizon implementation of the estimator is recursive. Therefore, for these estimator constructions, only the scalar value of the residual between recursive steps has to be stored for each plant $p \in G$. In the special case of the Kalman filter, additionally the filter state $\hat{x}$ as well as Σ have to be stored for all $p \in G$. The computational costs of evaluating the residual $r_p[k],\ k \in \mathbb{N}$ to plants $p \in G(k)$ depend on the order of P_p and the particular estimation algorithm, but are invariant to $k \in \mathbb{N}$.

   If $\mathcal{P}^G$ is known, a fixed bank of disturbance estimators may be set up off-line for disturbance estimation. However, for a general dynamic G, $\mathcal{P}^G$ is unknown and 'new' disturbance estimators have to be introduced on-line. This is conducted in the following way: if a new plant is introduced to G on-line at time $k \in \mathbb{N}$, the corresponding estimator is iterated forward from zero to time $k \in \mathbb{N}$. For a recursive estimator implementation the computational costs to introduce a new estimator on-line at time $k \in \mathbb{N}$ is therefore k times the costs of a one step iteration. It is therefore more expensive to introduce estimators later.

   The computation of the free switching signal $q_f(k) = \mathrm{argmin}_{p \in G(k)} r_p[k]$ from the residuals $r_p[k]$ is then a simple comparison of $n = |G(k)|$ scalars.

2. Implement the delay:
   Assume the algorithm switched from one controller to another at time $a \in \mathbb{N}$, i.e. $q(a-1) \neq q(a)$. Now compute the delay $\Delta(q(a))$ and store $\Delta(q(a))$ as well as $q(a)$ in memory to be evaluated by some delay routine. Since these operations only apply to a single plant $q(a)$, the computational cost is invariant to the size of $G(k),\ k \in \mathbb{N}$.

3. Compute the control signal:
   As before, assume the algorithm switched from one controller to another at time $a \in \mathbb{N}$. Now compute $K(q(a))$ and compute the control signal $u_2(i)$ via $u_2(i) = C_{K(q(a))} y_2(i),\ a \leq i < b$ where b is the next switching time and the controller is initialised to zero at time a. For example, if the controller is given by $C_c,\ c \in \mathcal{P}_{LTI}$ from equations (4.2)–(4.6) we would let $x_c(a) = 0$ and $w_2^c(i) = w_2(i),\ a \leq i < b$. Since these operations only apply to a single plant $q(a)$, the computational cost is invariant to the size of $G(k),\ k \in \mathbb{N}$.

We now come to the central chapter of this thesis, establishing bounds on the closed loop gain for the given EMMSAC algorithm.

Chapter 5

Stability and gain bound analysis of the nominal closed loop system

In this chapter we will establish $l_r,\ 1 \leq r \leq \infty$ norm bounds on the observation signal $w_2 \in \mathcal{W}_e$ in terms of the external disturbance signal $w_0 \in \mathcal{W}$. A particular feature of the bounds is that they depend on the size and geometry of a 'cover' of the candidate plant set, rather than the plant set itself. This characteristic allows the refinement scaling of plant model sets as a successively increasing fidelity sampling of e.g. a continuum of plants. The main result of this chapter establishes exactly this viewpoint. The following chapter then fully interprets this result and derives many consequences of the gain bound given here, including clear approaches to design.

On the first pass of reading this document the reader is advised to read Section 1 and the statement of the results in Section 5, omitting the detailed construction of the bounds in Section 2, 3, 4 and to follow the argument of Chapter 6 to the end of Section 2 first. There we will fill the objects G, H, U, ν with meaning.

Before we come to our first intermediate result, establishing gain bounds for atomic closed loop systems, we introduce some necessary notation.

1 Preliminaries

1.1 Uncertainty sets and covers

Let

$$U : \mathcal{W}_e \to map(\mathbb{N}, \mathcal{P}^*)$$

be a monotonic plant-generating operator where

$$\mathcal{P}^U := \bigcup_{w_2 \in \mathcal{W}_e} \bigcup_{k \in \mathbb{N}} U(w_2)(k) \subset \mathcal{P}.$$

U has the role of specifying an uncertainty set we seek to control at a given time $k \in \mathbb{N}$.

For example let a plant P_a be given by

$$P_a : y_1(k+1) = a y_1(k) + u_1, \; a \in [-a_{max}, a_{max}] \tag{5.1}$$

where a is an uncertain parameter. For a given finite a_{max} then take U to be constant:

$$U = U(k) = [-a_{max}, a_{max}], \; \forall k \in \mathbb{N}$$

Hence U specifies the uncertainty set. An implementation of an EMMSAC controller will then be based on a plant model set specified by a constant plant generating operator G, where G is a suitable sampling of U. However, for a constant uncertainty set, we will show in Theorem 6.6 that the closed loop gain scales with the uncertainty a_{max}; that is that the algorithm is conservative. A remedy to this problem is to dynamically expand the uncertainty set $U(k)$ (along with $G(k)$) until some performance requirement is met. For the purpose of our example we could, for example, let $U(k) = [-k, k]$. We will show in Chapter 6 that a performance-orientated expansion of U and G leads to gain function bounds that are invariant to the level of uncertainty in the system, and give algorithms of finite computational complexity.

We now consider sets of plants that are close to each other within each $U(k)$.

Let $\chi : \mathcal{P} \times \mathcal{P} \to \mathbb{R}^+$ be as in Assumption 3.4(5). Let

$$H : \mathcal{W}_e \to map(\mathbb{N}, \mathcal{P}^*) \tag{5.2}$$

be a plant-generating operator where

$$\mathcal{P}^H := \bigcup_{w_2 \in \mathcal{W}_e} \bigcup_{k \in \mathbb{N}} H(w_2)(k) \subset \mathcal{P}.$$

Let $\nu : \mathcal{W}_e \to map(\mathbb{N}, map(\mathcal{P}, \mathbb{R}^+))$ be given. As in Chapter 4 we write $U(k), H(k), \nu(k)$ for $U(w_2)(k), H(w_2)(k), \nu(w_2)(k)$.

Now define the ball

$$B_\chi(p, \nu(k)(p)) := \{p\} \cup \{p_1 \in \mathcal{P} \mid \chi(p, p_1) < \nu(k)(p)\} \cap U(k), \; p \in \mathcal{P}, \; k \in \mathbb{N} \tag{5.3}$$

to be the set of plants that reside within a neighbourhood of radius $\nu(k)(p)$, as measured by χ, around $p \in H(k)$ in $U(k)$. For an appropriate choice of H and ν, the union of the corresponding neighbourhoods in U then leads to a cover for U:

Definition 5.1. *(H, ν) is said to be a monotonic cover for a plant-generating operator U if $\forall k \in \mathbb{N}$, $w_2 \in \mathcal{W}_e$:*

1. *H and ν define a cover for U:*

$$U(k) \subset R(k) := \cup_{p \in H(k)} B_\chi(p, \nu(k)(p))), \ \forall k \in \mathbb{N}, \ w_2 \in \mathcal{W}_e.$$

2. *The cover is monotonic:*

$$R(k) \subset R(k+1), \ \forall k \in \mathbb{N}, \ \forall w_2 \in \mathcal{W}_2.$$

(H, ν) is said to be a finite cover if $H(k)$ is a finite set for all $k \in \mathbb{N}$, $w_2 \in \mathcal{W}_e$.

We will establish sufficient conditions for the existence of a finite cover (H, ν) for U in Chapter 6.

Returning to the example in equation (5.1), we can construct a monotonic cover in the following way: Assume $a_{max} = 100$. Let

$$H(k) = \{p_1, p_2, p_3, p_4, p_5\} = \{-100, -50, 0, 50, 100\}, \ \forall k \in \mathbb{N}$$

and $\nu(k) = 2 \max_{1 \le i \le 4} \chi(p_i, p_{i+1}), \ \forall k \in \mathbb{N}$. Then $U \subset R = \cup_{p \in H} B_\chi(p, \nu(p))), \ w_2 \in \mathcal{W}_e$ where the cover is monotonic (since it is constant).

The introduction of (H, ν) is the device by which we are able to express gain bounds which scale in terms of the number of elements of $|H(k)|$ rather than the absolute size of the set $|G(k)|$. This will lead to a notion of 'complexity' of a plant model uncertainty set in the next chapter.

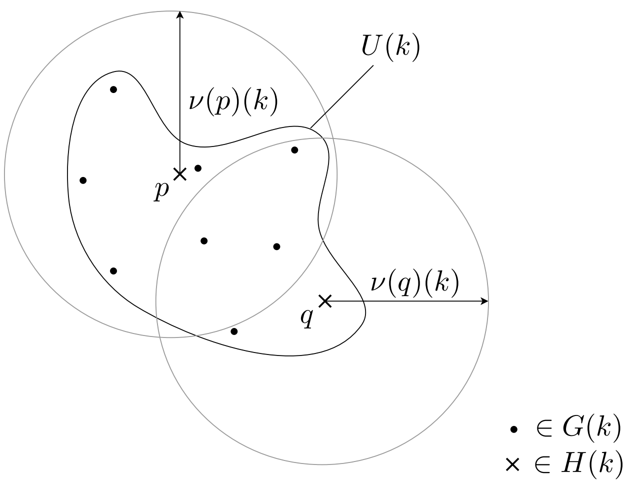

FIGURE 5.1: Uncertainty set $U(k)$, cover $(H(k), \nu(k))$ and sampling $G(k)$

Consider the example in Figure 5.1. We will think about the objects U, H, ν, G in the following way:

- U is the uncertainty, or more precisely U specifies the uncertainty set,
- The cover (H, ν) for U is the device by which we assess the 'complexity' of the uncertainty U,
- G is an appropriate sampling of the uncertainty set U.

1.2 Switching times

Let $q_f \in map(\mathbb{N}, \mathcal{P})$, and let $q = Dq_f$ (equations (4.29)–(4.33)) denote the switching signal. Let

$$L_k := \{l_0 = 0, l_1, l_2 \cdots\} = \{l \in \mathbb{N} \mid q(l-1) \neq q(l),\ 0 \leq l \leq k\} \tag{5.4}$$

be an ordered set, i.e. if $l_i, l_j \in L_k,\ i \leq j$ then $l_i \leq l_j$, interpreted as the set of physical switching times up to time $k \in \mathbb{N}$. These are the times where the algorithm switches from one controller to another.

To every pair of consecutive physical switching times l_i, l_{i+1} define the set of virtual switching times $V(l_i, l_{i+1})$ by

$$V(l_i, l_{i+1}) := \left\{ a \in \mathbb{N} \;\middle|\; \begin{array}{c} \exists b \in \mathbb{N} \text{ s.t. } a = l_i + b\Delta(q(l_i)), \\ l_i < a \leq l_{i+1} - \Delta(q(l_i)) \end{array} \right\}. \tag{5.5}$$

The idea of a virtual switch arises from the fact that if the algorithm switches to a controller $C_{K(q(l_i))}$ and remains switched to that controller for a long interval of time we can interpret this as a series of consecutive switches to the same controller. However note that a virtual switch differs from a physical switch in that the atomic controller is not intentionally initialised to 0 at the virtual switching time. Note that the interval $[l_i, l_{i+1}]$ might in some cases not be of sufficient length to accommodate a virtual switch at all. In that case $V(l_i, l_{i+1})$ is an empty set. Also note that virtual switching times are defined purely for analytical purposes and do not affect the actual switching algorithm whatsoever.

Now define the ordered set of all switching times, physical and virtual,

$$Q_k = \{k_0 = 0, k_1,\ k_2,\ \cdots\},\ 0 \leq k_i \leq k_{i+1} \leq k \tag{5.6}$$

by

$$Q_k := L_k \cup \bigcup_{i \geq 0} \{V(l_i, l_{i+1}) \mid l_i, l_{i+1} \in L_k\}, \tag{5.7}$$

where we treat for the remaining document virtual and physical switches alike. Let

$$Q_k(p) := \{i \in Q_k \mid q(i) = p\} \subset Q_k,\ p \in \mathcal{P}$$

be the switching times where the algorithm switches to a plant p.

Let $p \in H(k)$ and let

$$Q_k(p, \nu(k)(p)) := \cup_{x \in B_\chi(p,\nu(k)(p))}\{Q_k(x)\} \tag{5.8}$$

be the set of all switching times corresponding to the plants in the neighbourhood $B_\chi(p, \nu(k)(p))$ around a plant $p \in H(k)$.

For $p \in H(k)$, let

$$F_k(p, \nu(k)(p)) := \begin{cases} \{\max(Q_k(p, \nu(k)(p)))\} & \text{if} \quad \max(Q_k(p, \nu(k)(p))) \neq \emptyset \\ \emptyset & \text{otherwise} \end{cases} \tag{5.9}$$

be the switching time where the algorithm switches to a plant within the neighbourhood $B_\chi(p, \nu(k)(p))$ for the last time in the interval $[0, k]$. Note that $F_k(p, \nu(k)(p))$ is always defined since $\max Q_k(p, \nu(k)(p)) \leq k$.

Let

$$F_k := \cup_{p \in H(k)} F_k(p, \nu(k)(p)) \tag{5.10}$$

and note that:

$$F_k(p, \nu(k)(p)) \subset F_k \subset Q_k.$$

Let

$$O_k(p, \nu(k)(p)) := \begin{cases} Q(p, \nu(k)(p)) \setminus F_k(p, \nu(k)(p)) & \text{if} \quad Q_k(p, \nu(k)(p)) \neq \emptyset \\ \emptyset & \text{otherwise} \end{cases} \tag{5.11}$$

be the set of all 'ongoing' switching times corresponding to the plants in the neighbourhood $B_\chi(p, \nu(k)(p))$ around the plant p, i.e. the algorithm will switch back to a plant within $B_\chi(p, \nu(k)(p))$ at a later time in the interval $[0, k]$. We let:

$$O_k := \cup_{p \in H(k)} O_k(p, \nu(k)(p)) \tag{5.12}$$

and note that:

$$O_k(p, \nu(k)(p)) \subset O_k \subset Q_k.$$

For example, assume that there are only four plants and they are positioned as in Figure 5.2. Let q be such that the set of switching times Q_k and the set of switching times corresponding to each plant $Q_k(p_i)$, $1 \leq i \leq 4$ are as in the Table 5.1. Note that if a plant lies in more then one neighbourhood, it is counted multiple times. Hence its corresponding final switching time for one neighbourhood may be in the set of ongoing switching times for another neighbourhood, i.e.

$$F_k(p_i, v(k)(p_i)) \in O_k(p_j, v(k)(p_j)),\ i \neq j.$$

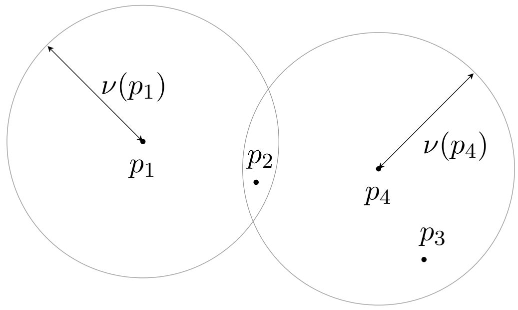

FIGURE 5.2: Neighbourhoods $B_\chi(p_1, \nu(p_1)),\ B_\chi(p_4, \nu(p_4))$

Q_k	1	5	11	22	44	101	202	212	333
$Q_k(p_1)$	1		11		44		202		
$Q_k(p_2)$		5		22					
$Q_k(p_3)$						101		212	
$Q_k(p_4)$									333
$Q_k(p_1, \nu(p_1))$	1	5	11	22	44		202		
$Q_k(p_4, \nu(p_4))$		5		22		101		212	333
$F_k(p_1, \nu(p_1))$							202		
$F_k(p_4, \nu(p_4))$									333
$O_k(p_1, \nu(p_1))$	1	5	11	22	44				
$O_k(p_4, \nu(p_4))$		5		22		101		212	

TABLE 5.1: A switched system with corresponding switching times

Note that this example is only to demonstrate the relationship between these sets of switching times. It is important to note that they will not appear in any gain bound and we do not impose any knowledge on them. In particular we do not impose knowledge about when the algorithm will switch, how often it will switch or to which controller. This is very important since allowing arbitrary disturbances of a certain size acting on the system it is impossible to predict every possible resulting switching sequence. The only time-structure property of these sets that will be used in the argument is that there exists an upper and lower bound on the pause between two switches (see Lemma 5.4). This is a result of the switching delay, the definition of the virtual switching times and the fact that the sets are ordered.

Impose the following constraint on the set $\mathcal{P}^U$:

Assumption 5.2. *Let $\mathcal{P}^U \subset \mathcal{P}$ have the property:*

$$\sigma := \max_{p_1, p_2 \in \mathcal{P}^U} \max\{\sigma(p_1), \sigma(K(p_2))\} < \infty.$$

where we recall that $\sigma(p)$ *and* $\sigma(c)$ *are defined by equation* (4.8) *and* (4.7)*.*

Therefore σ represents the total number of time steps the controller and plant signals need to be observed to uniquely define the initial condition of any closed loop system $[P_{p_1}, C_{K(p_2)}]$, $p_1, p_2 \in \mathcal{P}^U$. This implies finite dimensionality of the plants and controllers.

Furthermore assume the following:

Assumption 5.3. *We assume the delay transition function* $\Delta : \mathcal{P} \to \mathbb{N}$ *satisfies:*

$$\Delta(p) > \sigma, \ \forall p \in \mathcal{P}^U.$$

For all switching times $k_i \in Q_k$ define the intervals

$$A_i := [k_i - \sigma, k_i) \tag{5.13}$$

$$B_i := [k_i, k_{i+1} - \sigma). \tag{5.14}$$

Note that by Lemma 5.4:

$$k_{i+1} - k_i \geq \Delta(q(k_i)) > \sigma$$

hence $k_{i+1} - \sigma > k_i$ and A_i, B_i are defined and form a disjoint cover of $\mathbb{N}$.

Upper and lower bounds on the switching times are now given as follows:

Lemma 5.4. *Suppose* $\Delta : \mathcal{P} \to \mathbb{N}$ *is a given delay transition function and suppose the delay operator* D *is given by equation* (4.30)*–*(4.33)*. Let* $k \in \mathbb{N}$ *and let* $q_f \in map(\mathbb{N}, \mathcal{P}^U)$*. Let* $q = Dq_f$*. Suppose* $k_i \leq k_{i+1}$ *are consecutive switching times,* $k_i, k_{i+1} \in Q_k$ *where* Q_k *is defined by equations* (5.4)*–*(5.7)*. Let* $p = q(k_i)$*. Then:*

$$\Delta(p) \leq k_{i+1} - k_i < 2\Delta(p). \tag{5.15}$$

Proof By the definition of the switching delay in equation (4.32) it follows that $\Delta(p) \leq k_{i+1} - k_i$. If k_{i+1} is a virtual switching time, then $k_{i+1} - k_i = \Delta(q(k_i))$ by equation (5.5), and if k_{i+1} is a physical switching time, then

$$k_i := l_i + b\Delta(q(l_i)) \leq k_{i+1} - \Delta(q(l_i)) < l_i + (b+1)\Delta(q(l_i)) = k_i + \Delta(q(l_i)),$$

hence $k_{i+1} - k_i < 2\Delta(q(k_i))$ and equation (5.15) follows. □

We are now in the position to begin with the construction of the gain bound. The proof is organised into sections as follows.

Section 2. Gain bounds for atomic closed loop systems:

In this section we are concerned with a) the atomic closed loop $[P_{p_*}, C_{K(p)}]$, that is the closed loop system containing the true plant P_{p_*}, $p_* \in \mathcal{P}$ and a controller based on any plant P_p, $p \in \mathcal{P}$, and b) the atomic closed loop $[P_p, C_{K(p)}]$, $p \in \mathcal{P}$, that is the closed loop system containing a matching plant-controller pair. In both cases we will establish a bound on (portions of) the observation w_2 in terms of the disturbance estimate $d_p[k]$. The bounds from a) will be used in the cont ex of final switching times, i.e. $k_i \in F_k$ where the bounds from b) will be used in the context of non-final switching times, i.e. $k_i \in O_k$.

Section 3. Bounds on disturbance estimates:
Since the established bounds on the observation in Section 2 are given in terms of disturbance estimates $d_p[k]$, and the overall goal is to construct a bound on the observation w_2 in terms of the true disturbance w_0, we now establish bounds on the disturbance estimates $d_p[k]$ in terms of the true disturbance w_0.

Section 4. Gain bounds for non-final switching intervals:
In this section we will utilise the results from Section 3. to show that (a series of) disturbance estimates $d_p[k_i]$ corresponding to intervals associated with non-final switching times, i.e. $k_i \in O_k$, can be bounded efficiently in terms of the true disturbance w_0. This leads to a bound on w_2 in terms of w_0 for sequences of intervals associated with non-final switching times.

Section 5. Main result:
Finally, all gain results are collated to the main result that establishes a bound on w_2 in terms of w_0 for both, intervals associated with final and non-final switching times — hence over the whole time axis.

2 Gain bounds for atomic closed loop systems

The first result establishes bounds on the gain from the disturbance signals w_0 to the internal signals w_2 for the atomic closed loop interconnection between the true plant and the controller switched into closed loop at time k_i, i.e. $[P_{p_*}, C_{K(q(k_i))}]$ as depicted in Figure 5.3, on the various intervals of type A_i, B_i, $k_i \in Q_k$.

The two cases $w_2^c|_{A_i} = 0$ and $w_2^c|_{A_i} = w_2|_{A_i}$ correspond to the case whereby the controller is initialised to zero at time k_i i.e. $k_i \in L_k$ (a physical switch) or the case where the controller is not intentionally initialised to zero at time k_i i.e $k_i \in Q_k \setminus L_k$ (a virtual switch). To improve readability we repeat all relevant equations in Table 5.2.

$$\begin{aligned} y_1 &= Pu_1 && (2.7)\\ u_0 &= u_1 + u_2 && (2.8)\\ y_0 &= y_1 + y_2 && (2.9)\\ u_2^c &= C_c y_2^c && (4.1) \end{aligned}$$

$$D : map(\mathbb{N}, \mathcal{P}) \to map(\mathbb{N}, \mathcal{P}) \tag{4.30}$$

$$[k \mapsto q_f(k)] \mapsto [k \mapsto q(k)] \tag{4.31}$$

$$q(k) := \begin{cases} q_f(k) & \text{if} \quad k - k_s(k) \geq \Delta(q(k_s(k))) \\ q(k_s(k)) & \text{else} \end{cases} \tag{4.32}$$

$$k_s(k) := \max\{i \in \mathbb{N} \mid 0 \leq i \leq k,\ q(i) \neq q(i-1)\} \tag{4.33}$$

$$L_k := \{l_0 = 0, l_1, l_2 \cdots\} = \{l \in \mathbb{N} \mid q(l-1) \neq q(l),\ 0 \leq l \leq k\} \tag{5.4}$$

$$V(l_i, l_{i+1}) := \left\{ a \in \mathbb{N} \;\middle|\; \begin{array}{c} \exists b \in \mathbb{N} \text{ s.t. } a = l_i + b\Delta(q(l_i)), \\ l_i < a \leq l_{i+1} - \Delta(q(l_i)) \end{array} \right\} \tag{5.5}$$

$$Q_k = \{k_0 = 0, k_1,\ k_2,\ \cdots\},\ 0 \leq k_i \leq k_{i+1} \leq k \tag{5.6}$$

$$Q_k := L_k \cup \bigcup_{i \geq 0} \{V(l_i, l_{i+1}) \mid l_i, l_{i+1} \in L_k\} \tag{5.7}$$

$$A_i := [k_i - \sigma, k_i) \tag{5.13}$$

$$B_i := [k_i, k_{i+1} - \sigma) \tag{5.14}$$

Table 5.2: Details for Proposition 5.5

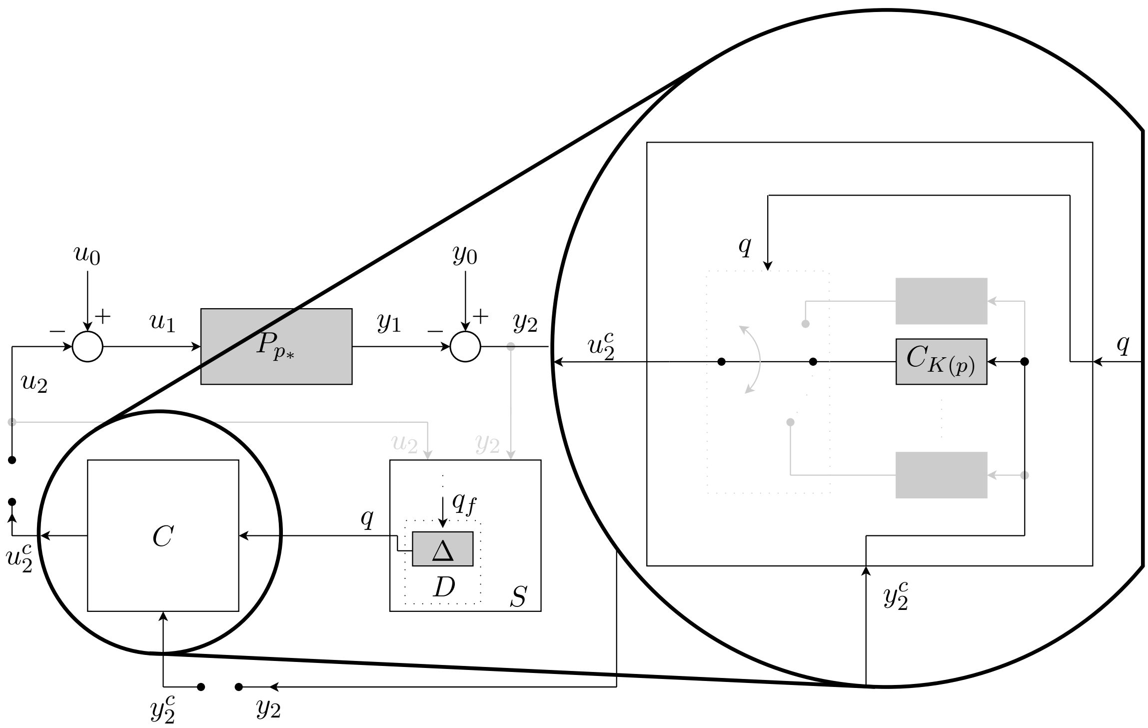

FIGURE 5.3: Closed loop $[P_{p_*}, C_{K(q(k_i))}]$ with magnified switching controller C

Proposition 5.5. *Let* $1 \leq r \leq \infty$. *Suppose* $\mathcal{P}^U \subset \mathcal{P}$ *satisfies Assumption 5.2. Let* $p_* \in \mathcal{P}^U$ *and* $P = P_{p_*}$. *Let* $K : \mathcal{P} \to \mathcal{C}$ *be a given control design satisfying Assumption 4.1(1). Suppose* Δ *is a given delay transition function satisfying Assumption 5.3 and suppose the delay operator* D *is given by equation* (4.30)–(4.33).

Let $k \in \mathbb{N}$ *and let* $q_f \in map(\mathbb{N}, \mathcal{P}^U)$. *Let* $q = Dq_f$. *Suppose* $k_i \leq k_{i+1}$ *are consecutive switching times,* $k_i, k_{i+1} \in Q_k$ *where* Q_k *is defined by equations* (5.4)–(5.7) *and let the intervals* A_i, A_{i+1}, B_i *be given by equations* (5.13),(5.14). *Suppose* $(w_0, w_1, w_2) \in \mathcal{W} \times \mathcal{W}_e \times \mathcal{W}_e$, $w_2^c \in \mathcal{W}_e$ *satisfy equations* (2.7)–(2.9),(4.1) *on the interval* $A_i \cup B_i \cup A_{i+1}$, *where* $p = q(k_i)$, $c = K(p)$ *and either*

$$w_2^c|_{A_i} = 0, \; w_2^c|_{B_i \cup A_{i+1}} = w_2|_{B_i \cup A_{i+1}}$$

or

$$w_2^c|_{A_i \cup B_i \cup A_{i+1}} = w_2|_{A_i \cup B_i \cup A_{i+1}}.$$

Then, in both cases,

$$\|\mathscr{T}_{k_{i+1}-1} w_2\| \quad \leq \quad \gamma_1(p)\|\mathscr{T}_{k_i-1} w_2\| + \gamma_2(p)\|w_0\|$$

where with α, β *from Assumption 4.1 define*

$$\begin{aligned} \gamma_1(p) &= 1 + \sup_{\Delta(p) \leq x \leq 2\Delta(p)} \alpha(p_*, K(p), 0, x) \\ \gamma_2(p) &= \sup_{\Delta(p) \leq x \leq 2\Delta(p)} \beta(p_*, K(p), 0, x). \end{aligned}$$

Proof Let

$$\begin{aligned} I_1 &= A_i = [k_i - \sigma, k_i) \\ I_2 &= \emptyset \\ I_3 &= B_i \cup A_{i+1} = [k_i, k_{i+1}) \end{aligned}$$

Since

$$|I_1| = |A_i| = \sigma \geq \max\{\sigma(p_*), \sigma(K(q(k_i)))\}$$

by Assumption 4.1(1) we have for the closed loop $[P_{p_*}, C_{K(q(k_i))}]$ that

$$\begin{aligned} \|\mathscr{T}_{k_{i+1}-1} w_2\| &\leq \|\mathscr{T}_{k_i-1} w_2\| + \|w_2|_{I_3}\| \\ &\leq \|\mathscr{T}_{k_i-1} w_2\| + \alpha(p_*, K(q(k_i)), 0, |I_3|)\|w_2|_{I_1}\| \\ &\quad +\beta(p_*, K(q(k_i)), 0, |I_3|)\|w_0|_{I_1 \cup I_2 \cup I_3}\| \\ &\leq \|\mathscr{T}_{k_i-1} w_2\| + \alpha(p_*, K(q(k_i)), 0, |I_3|)\|\mathscr{T}_{k_i-1} w_2\| \\ &\quad +\beta(p_*, K(q(k_i)), 0, |I_3|)\|w_0\| \\ &\leq (1 + \alpha(p_*, K(q(k_i)), 0, |I_3|))\|\mathscr{T}_{k_i-1} w_2\| + \beta(p_*, K(q(k_i)), 0, |I_3|)\|w_0\| \end{aligned}$$

By Lemma 5.4 we now have

$$\Delta(p) \leq |I_3| = k_{i+1} - k_i \leq 2\Delta(p).$$

and arrive at

$$\begin{aligned} \|\mathscr{T}_{k_{i+1}-1} w_2\| &\leq (1 + \alpha(p_*, K(p), 0, |I_3|))\|\mathscr{T}_{k_i-1} w_2\| + \beta(p_*, K(p), 0, |I_3|)\|w_0\| \\ &\leq \gamma_1(p)\|\mathscr{T}_{k_i-1} w_2\| + \gamma_2(p)\|w_0\| \end{aligned}$$

as required. □

Before we discuss the next gain bound we give an elementary bound.

For $x, y, c \in \mathbb{R}$ define

$$\lfloor c \rfloor := \max\{n \in \mathbb{Z} \mid n \leq c\} \text{ and } \binom{x}{y} := \frac{x!}{y!(x-y)!}.$$

Lemma 5.6. *Let $1 \leq \xi < \infty$. Let $a, b \geq 0$. Then*

$$(a + b)^\xi \leq J(\xi)(a^\xi + b^\xi).$$

where

$$J(\xi) = \xi \binom{\xi}{\lfloor \xi/2 \rfloor}.$$

Proof Let $v, w \in \mathbb{N}$. Observe that since

$$\begin{aligned} a^v b^w &\leq a^{v+w} \text{ if } b \leq a \\ a^v b^w &\leq b^{v+w} \text{ if } a \leq b \end{aligned}$$

it follows that

$$a^v b^w \leq \max\{a^{v+w}, b^{v+w}\} \leq a^{v+w} + b^{v+w}.$$

We then have

$$\begin{aligned} (a+b)^\xi &= \sum_{i=0}^{\xi} \binom{\xi}{i} a^{\xi-i} b^i \leq \binom{\xi}{\lfloor \xi/2 \rfloor} \sum_{i=0}^{\xi} a^{\xi-i} b^i \\ &= \binom{\xi}{\lfloor \xi/2 \rfloor} \left[(a^\xi + b^\xi) + \sum_{i=1}^{\xi-1} a^{\xi-i} b^i \right] \\ &\leq \binom{\xi}{\lfloor \xi/2 \rfloor} \left[(a^\xi + b^\xi) + \sum_{i=1}^{\xi-1} (a^\xi + b^\xi) \right] \\ &\leq \xi \binom{\xi}{\lfloor \xi/2 \rfloor} (a^\xi + b^\xi) \end{aligned}$$

as required. □

Note that when applying this lemma later, $\xi \in \mathbb{N}$ will be chosen to be

$$\xi = \begin{cases} r & \text{for} \quad 1 \leq r < \infty \\ 1 & \text{for} \quad r = \infty \end{cases}$$

where r determines the space l_r in which the analysis is being conducted.

Up to this point, the transition delay function $\Delta : \mathcal{P} \to \mathbb{N}$ and the controller design procedure $K : \mathcal{P} \to \mathcal{C}$ have not been connected in any way. We will now do so with the help of the so-called attenuation function $l : \mathcal{P} \to [0, 1)$: Let Δ, K, l satisfy:

$$\begin{aligned} J(r)\alpha^r(p, K(p), \Delta(p) - \sigma, \sigma) &\leq l(p) < 1, \ \forall p \in \mathcal{P}^U \text{ if } 1 \leq r < \infty \\ \alpha(p, K(p), \Delta(p) - \sigma, \sigma) &\leq l(p) < 1, \ \forall p \in \mathcal{P}^U \text{ if } r = \infty \end{aligned} \tag{5.16}$$

where Δ satisfies Assumption 5.3 ($\Delta(p) > \sigma, \ p \in \mathcal{P}^U$) and α, β are defined in Assumptions 4.1.

The purpose of the attenuation function l is to define an upper bound on the signal attenuation that is achieved by the atomic closed loop interconnection between the plant P_p and the corresponding controller $C_{K(p)}$ over some interval of length $\Delta(p) - \sigma$. In practise, one would choose a stabilising design procedure K, an attenuation function l, a norm l_r and then compute for all $p \in \mathcal{P}^U$ a corresponding $\Delta(p)$ such that inequality

(5.16) holds, hence note that there always exists such a Δ. Inequality (5.16) therefore establishes a relationship between delay and attenuation. This freedom in choosing Δ can now be utilised in many ways. For example, it can be utilised to decouple control sampling rate and switching rate, e.g. by choosing Δ large we would maintain a high update rate for controller sampling and updating the disturbance estimates, however have a low switching rate between controllers.

The next result establishes bounds on the gain from the disturbance signals w_0^p to the internal signals w_2 for the atomic closed loop $[P_p, C_{K(p)}]$, $p = q(k_i)$ on the various intervals of type A_i, B_i, $k_i \in Q_k$. That is the closed loop loop interconnection between: the controller the algorithm switches to at time k_i, and its corresponding plant — as depicted in Figure 5.4. To improve readability we repeat all relevant equations in Table 5.3.

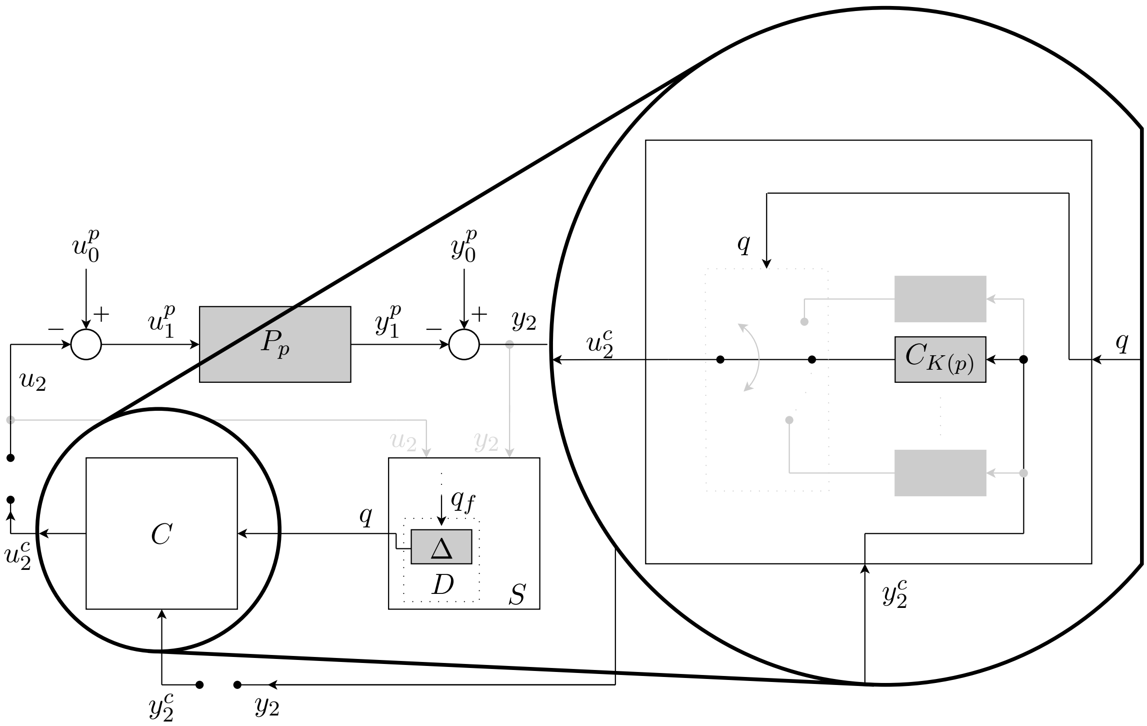

FIGURE 5.4: Closed loop $[P_p, C_{K(p)}]$ with magnified switching controller C

$$y_1^p = P_p u_1^p \tag{3.8}$$
$$u_0^p = u_1^p + u_2 \tag{3.9}$$
$$y_0^p = y_1^p + y_2 \tag{3.10}$$
$$u_2^c = C_c y_2^c \tag{4.1}$$

$$D : map(\mathbb{N}, \mathcal{P}) \rightarrow map(\mathbb{N}, \mathcal{P}) \tag{4.30}$$

$$[k \mapsto q_f(k)] \mapsto [k \mapsto q(k)] \tag{4.31}$$

$$q(k) := \begin{cases} q_f(k) & \text{if} \quad k - k_s(k) \geq \Delta(q(k_s(k))) \\ q(k_s(k)) & \text{else} \end{cases} \tag{4.32}$$

$$k_s(k) := \max\{i \in \mathbb{N} \mid 0 \leq i \leq k,\ q(i) \neq q(i-1)\} \tag{4.33}$$

$$L_k := \{l_0 = 0, l_1, l_2 \cdots\} = \{l \in \mathbb{N} \mid q(l-1) \neq q(l),\ 0 \leq l \leq k\} \tag{5.4}$$

$$V(l_i, l_{i+1}) := \left\{ a \in \mathbb{N} \;\middle|\; \begin{array}{c} \exists b \in \mathbb{N} \text{ s.t. } a = l_i + b\Delta(q(l_i)), \\ l_i < a \leq l_{i+1} - \Delta(q(l_i)) \end{array} \right\} \tag{5.5}$$

$$Q_k = \{k_0 = 0, k_1,\ k_2,\ \cdots\},\ 0 \leq k_i \leq k_{i+1} \leq k \tag{5.6}$$

$$Q_k := L_k \cup \bigcup_{i \geq 0} \{V(l_i, l_{i+1}) \mid l_i, l_{i+1} \in L_k\} \tag{5.7}$$

$$A_i := [k_i - \sigma, k_i) \tag{5.13}$$
$$B_i := [k_i, k_{i+1} - \sigma) \tag{5.14}$$

$$\begin{array}{rcl} J(r)\alpha^r(p, K(p), \Delta(p) - \sigma, \sigma) & \leq & l(p) < 1,\ \forall p \in \mathcal{P}^U \text{ if } 1 \leq r < \infty \\ \alpha(p, K(p), \Delta(p) - \sigma, \sigma) & \leq & l(p) < 1,\ \forall p \in \mathcal{P}^U \text{ if } r = \infty \end{array} \tag{5.16}$$

Table 5.3: Details for Proposition 5.7

Proposition 5.7. *Let $1 \leq r \leq \infty$. Suppose $p \in \mathcal{Q} \subset \mathcal{P}^U \subset \mathcal{P}$, $c = K(p)$ and $\mathcal{P}^U$ satisfies Assumption 5.2. Let $K : \mathcal{P} \to \mathcal{C}$ be a given control design satisfying Assumption 4.1(1),(2). Suppose Δ is a given delay transition function satisfying Assumption 5.3 and suppose the delay operator D is given by equation (4.30)–(4.33). Let $l : \mathcal{P} \to [0, 1)$ be a given attenuation function and suppose that K, Δ, l satisfy inequality (5.16). Let $k \in \mathbb{N}$ and let $q_f \in map(\mathbb{N}, \mathcal{P}^U)$. Let $q = Dq_f$ and suppose $q(k_i) = p$. Suppose $k_i \leq k_{i+1}$ are consecutive switching times, $k_i, k_{i+1} \in Q_k$ where Q_k is defined by equations (5.4)–(5.7). Let the intervals A_i, A_{i+1}, B_i be given by equations (5.13),(5.14). Suppose $(w_0^p, w_1^p, w_2) \in \mathcal{W} \times \mathcal{W}_e \times \mathcal{W}_e$, $w_2^c \in \mathcal{W}_e$ satisfy equations (3.8)–(3.10),(4.1) on the interval $A_i \cup B_i \cup A_{i+1}$ and either*

$$w_2^c|_{A_i} = 0, \;\; w_2^c|_{B_i \cup A_{i+1}} = w_2|_{B_i \cup A_{i+1}} \tag{5.17}$$

or

$$w_2^c|_{A_i \cup B_i \cup A_{i+1}} = w_2|_{A_i \cup B_i \cup A_{i+1}}. \tag{5.18}$$

Then, in both cases, for $1 \leq r < \infty$:

$$\begin{aligned} \|w_2|_{A_{i+1}}\|_r^r &\leq \alpha_{OP}(\mathcal{Q})\|w_2|_{A_i}\|_r^r + \beta_{OP}(\mathcal{Q})\|w_0^{q(k_i)}|_{A_i \cup B_i \cup A_{i+1}}\|_r^r \\ \|w_2|_{B_i}\|_r^r &\leq \alpha_{OS}(\mathcal{Q})\|w_2|_{A_i}\|_r^r + \beta_{OS}(\mathcal{Q})\|w_0^{q(k_i)}|_{A_i \cup B_i}\|_r^r \end{aligned}$$

and for $r = \infty$:

$$\begin{aligned} \|w_2|_{A_{i+1}}\|_\infty &\leq \alpha_{OP}(\mathcal{Q})\|w_2|_{A_i}\|_\infty + \beta_{OP}(\mathcal{Q})\|w_0^{q(k_i)}|_{A_i \cup B_i \cup A_{i+1}}\|_\infty \\ \|w_2|_{B_i}\|_\infty &\leq \alpha_{OS}(\mathcal{Q})\|w_2|_{A_i}\|_\infty + \beta_{OS}(\mathcal{Q})\|w_0^{q(k_i)}|_{A_i \cup B_i}\|_\infty \end{aligned}$$

where for $J(\xi)$ from Lemma (5.6) and α, β from Assumption 4.1 define

$$\begin{aligned} \xi &= \begin{cases} r & \text{for} \quad 1 \leq r < \infty \\ 1 & \text{for} \qquad r = \infty \end{cases} \\ \alpha_{OP}(\mathcal{Q}) &= \max_{p_1 \in \mathcal{Q}} l(p_1) \\ \beta_{OP}(\mathcal{Q}) &= J(\xi) \sup_{p_1 \in \mathcal{Q}} \sup_{\Delta(p_1) \leq x \leq 2\Delta(p_1)} \beta^\xi(p_1, K(p_1), x - \sigma, \sigma) \\ \alpha_{OS}(\mathcal{Q}) &= J(\xi) \sup_{p_1 \in \mathcal{Q}} \sup_{\Delta(p_1) \leq x \leq 2\Delta(p_1)} \alpha^\xi(p_1, K(p_1), 0, x - \sigma) \\ \beta_{OS}(\mathcal{Q}) &= J(\xi) \sup_{p_1 \in \mathcal{Q}} \sup_{\Delta(p_1) \leq x \leq 2\Delta(p_1)} \beta^\xi(p_1, K(p_1), 0, x - \sigma). \end{aligned}$$

Proof By Lemma 5.4, inequality (5.15) we have

$$\Delta(p) \leq |B_i \cup A_{i+1}| = |B_i| + \sigma = k_{i+1} - k_i \leq 2\Delta(p). \tag{5.19}$$

Let

$$\begin{aligned} I_1 &= A_i = [k_i - \sigma, k_i) \\ I_2 &= B_i = [k_i, k_{i+1} - \sigma) \\ I_3 &= A_{i+1} = [k_{i+1} - \sigma, k_{i+1}). \end{aligned}$$

By Assumption 5.2,

$$|I_1| = |A_i| = \sigma \geq \max\{\sigma(p), \sigma(K(p))\},$$

it follows from Assumption 4.1(1) inequality (4.10) that:

$$\begin{aligned} \|w_2|_{A_{i+1}}\|_r^\xi &\leq \big(\alpha(p, K(p), |B_i|, |A_{i+1}|)\|w_2|_{A_i}\|_r \\ &\quad +\beta(p, K(p), |B_i|, |A_{i+1}|)\|w_0^p|_{A_i \cup B_i \cup A_{i+1}}\|_r\big)^\xi \\ &\leq \big(\alpha(p, K(p), \Delta(p) - \sigma, \sigma)\|w_2|_{A_i}\|_r \\ &\quad +\beta(p, K(p), |B_i|, \sigma)\|w_0^p|_{A_i \cup B_i \cup A_{i+1}}\|_r\big)^\xi \end{aligned}$$

where the second inequality follows from the fact that α is monotonically decreasing in the third parameter (Assumption 4.1(2)) and $|B_i| \geq \Delta(p) - \sigma$ (equation (5.19)).

Since K, Δ, l satisfy inequality (5.16) for $1 \leq r \leq \infty$ we have that:

$$J(\xi)\alpha^\xi(p, K(p), \Delta(p) - \sigma, \sigma) \leq l(p) < 1, \ \forall p \in \mathcal{P}^U.$$

Hence by Lemma 5.6 and equation (5.19) we obtain

$$\begin{aligned} \|w_2|_{A_{i+1}}\|_r^\xi &\leq J(\xi)\alpha^\xi(p, K(p), \Delta(p) - \sigma, \sigma)\|w_2|_{A_i}\|_r^\xi \\ &\quad +J(\xi)\beta^\xi(p, K(p), |B_i|, \sigma)\|w_0^p|_{A_i \cup B_i \cup A_{i+1}}\|_r^\xi \\ &\leq l(p)\|w_2|_{A_i}\|_r^\xi + \max_{\Delta(p) \leq x \leq 2\Delta(p)} J(\xi)\beta^\xi(p, K(p), x - \sigma, \sigma)\|w_0^p|_{A_i \cup B_i \cup A_{i+1}}\|_r^\xi \end{aligned}$$

and hence

$$\|w_2|_{A_{i+1}}\|_r^\xi \leq \alpha_{OP}(\mathcal{Q})\|w_2|_{A_i}\|_r^\xi + \beta_{OP}(\mathcal{Q})\|w_0^p|_{A_i \cup B_i \cup A_{i+1}}\|_r^\xi.$$

Now let

$$\begin{aligned} I_1 &= A_i = [k_i - \sigma, k_i) \\ I_2 &= \emptyset \\ I_3 &= B_i = [k_i, k_{i+1} - \sigma). \end{aligned}$$

By Assumption 5.2,

$$|I_1| \geq \sigma \geq \max\{\sigma(p_1), \sigma(K(p_2))\}$$

and it follows from Assumption 4.1(1) (inequality (4.10)), Lemma 5.6 and equation (5.19) that:

$$\begin{aligned}
\|w_2|_{B_i}\|_r^\xi &\leq \big(\alpha(p,K(p),0,|B_i|)\|w_2|_{A_i}\|_r + \beta(p,K(p),0,|B_i|)\|w_0^p|_{A_i\cup B_i}\|_r\big)^\xi \\
&\leq J(\xi)\alpha^\xi(p,K(p),0,|B_i|)\|w_2|_{A_i}\|_r^\xi + J(\xi)\beta^\xi(p,K(p),0,|B_i|)\|w_0^p|_{A_i\cup B_i}\|_r^\xi \\
&\leq \max_{\Delta(p)\leq x\leq 2\Delta(p)} J(\xi)\alpha^\xi(p,K(p),0,x-\sigma)\|w_2|_{A_i}\|_r^\xi \\
&\quad + \max_{\Delta(p)\leq x\leq 2\Delta(p)} J(\xi)\beta^\xi(p,K(p),0,x-\sigma)\|w_0^p|_{A_i\cup B_i}\|_r^\xi \\
&\leq \alpha_{OS}(\mathcal{Q})\|w_2|_{A_i}\|_r^\xi + \beta_{OS}(\mathcal{Q})\|w_0^p|_{A_i\cup B_i}\|_r^\xi
\end{aligned}$$

as required. □

3 Bounds on disturbance estimates

The next proposition follows directly from Assumption (3.4)(5) and gives a bound on a series of disturbance estimates. The idea is to cover a set of plants by a union of sub-covers B_j and then use Assumption (3.4)(5) to bound disturbance estimates corresponding to all plants within a sub-cover B_j by a disturbance estimate of a single plant z_j in B_j. This technique opens up the possibility to use infinitely many plant models since the bound will only depend on the cover and not the (number of) plants covered by it. To improve readability we repeat all relevant equations in Table 5.4.

$E : \mathcal{W}_e \to map(\mathbb{N}, map(\mathcal{P}, map(\mathbb{N},\mathbb{R}^h)))$	(3.14)
$w_2 \mapsto \big[k \mapsto (p \mapsto d_p[k])\big]$	(3.15)

Table 5.4: Details for Proposition 5.8

Proposition 5.8. *Let* $1 \leq r \leq \infty$. *Suppose* $\mathcal{P}^U \subset \mathcal{P}$ *and let* $a, \lambda, m, n \in \mathbb{N},\ m \leq n$. *Suppose* E *is given by equations* (3.14),(3.15) *and satisfies Assumption 3.4(5). Let* $\bar{k}_i, \tilde{k}_i \in \mathbb{N},\ 0 \leq \bar{k}_i \leq \lambda,\ \tilde{k}_i < \tilde{k}_{i+1},\ B_i \subset \mathcal{P}^U,\ m \leq i \leq n$. *Let* $k \geq 1, w_2 \in \mathcal{W}_e$ *and* $d_p[k] = E(w_2)(k)(p),\ p \in \mathcal{P}$. *Suppose* $p_i, z_i \in \mathcal{P}$ *satisfy* $p_i, z_i \in B_i$. *Then:*

$$\|\Phi_{\bar{k}_m} d_{p_m}[\tilde{k}_m], \Phi_{\bar{k}_{m+1}} d_{p_{m+1}}[\tilde{k}_{m+1}], \cdots, \Phi_{\bar{k}_n} d_{p_n}[\tilde{k}_n]\| \leq \|\Phi_{\bar{k}_m} d_{z_m}[\tilde{k}_m], \Phi_{\bar{k}_{m+1}} d_{z_{m+1}}[\tilde{k}_{m+1}], \cdots, \Phi_{\bar{k}_n} d_{z_n}[\tilde{k}_n]\| + \chi\|c\|\|\mathscr{T}_{\tilde{k}_n} w_2\|$$

where

$$\chi = \max_{m\leq i\leq n} \sup_{p,q\in B_i} \chi(p,q) \tag{5.20}$$

and $\chi(\cdot,\cdot), \Phi_j, c$ *are defined as in Assumption 3.4(5).*

Proof Observe that

$$\|a+b, c+d\| \leq \|a, c\| + \|b, d\|, \ \ a, b, c, d \in \mathbb{R}$$

and

$$\|a, b\|_r = \|\|a\|_r, \|b\|_r\|_r, \ \ a, b, \in \mathcal{S}, \ \ 1 \leq r \leq \infty.$$

By Assumption (3.4)(5) we then have

$$\begin{aligned}
&\|\Phi_{\bar{k}_m} d_{p_m}[\tilde{k}_m], \Phi_{\bar{k}_{m+1}} d_{p_{m+1}}[\tilde{k}_{m+1}], \cdots, \Phi_{\bar{k}_n} d_{p_n}[\tilde{k}_n]\| \\
&\qquad \leq \left\| \begin{array}{c} \|\Phi_{\bar{k}_m} d_{z_m}[\tilde{k}_m]\| + \chi(p_m, z_m)\|\Upsilon_{\tilde{k}_m} w_2\|, \\ \|\Phi_{\bar{k}_{m+1}} d_{z_{m+1}}[\tilde{k}_{m+1}]\| + \chi(p_{m+1}, z_{m+1})\|\Upsilon_{\tilde{k}_{m+1}} w_2\|, \\ \vdots, \\ \|\Phi_{\bar{k}_n} d_{z_n}[\tilde{k}_n]\| + \chi(p_n, z_n)\|\Upsilon_{\tilde{k}_n} w_2\| \end{array} \right\| \\
&\qquad \leq \|\Phi_{\bar{k}_m} d_{z_m}[\tilde{k}_m], \Phi_{\bar{k}_{m+1}} d_{z_{m+1}}[\tilde{k}_{m+1}], \cdots, \Phi_{\bar{k}_n} d_{z_n}[\tilde{k}_n]\| \\
&\qquad\qquad + \chi\|\Upsilon_{\tilde{k}_m} w_2, \Upsilon_{\tilde{k}_{m+1}} w_2, \cdots, \Upsilon_{\tilde{k}_n} w_2\|.
\end{aligned}$$

For $1 \leq r < \infty$, and since $\tilde{k}_i < \tilde{k}_{i+1}$ we can write

$$\begin{aligned}
\|\Upsilon_{\tilde{k}_m} w_2, \Upsilon_{\tilde{k}_{m+1}} w_2, \cdots, \Upsilon_{\tilde{k}_n} w_2\|_r &= \left(\sum_{i=m}^{n} \|\Upsilon_{\tilde{k}_i} w_2\|_r^r \right)^{1/r} \\
&\leq \left(\sum_{k=0}^{\tilde{k}_n} \|\Upsilon_k w_2\|_r^r \right)^{1/r} \\
&= \left(\sum_{k=0}^{\tilde{k}_n} \sum_{j=0}^{k} |(\Upsilon_k w_2)(j)|^r \right)^{1/r} \\
&= \left(\sum_{k=0}^{\tilde{k}_n} \sum_{j=0}^{k} |c(k-j) w_2(j)|^r \right)^{1/r} \\
&= \left(\sum_{k=0}^{\tilde{k}_n} |w_2(k)|^r \sum_{j=0}^{\tilde{k}_n - k} |c(j)|^r \right)^{1/r} \\
&\leq \left(\sum_{k=0}^{\tilde{k}_n} |w_2(k)|^r \|c\|_r^r \right)^{1/r} \\
&\leq \|c\|_r \|\mathscr{T}_{\tilde{k}_n} w_2\|_r
\end{aligned}$$

and for $r = \infty$

$$\begin{aligned} \|\Upsilon_{\tilde{k}_m} w_2, \Upsilon_{\tilde{k}_{m+1}} w_2, \cdots, \Upsilon_{\tilde{k}_n} w_2\|_\infty &\leq \max_{0\leq k\leq \tilde{k}_n} \{\|\Upsilon_k w_2\|_\infty\} \\ &= \max_{0\leq k\leq \tilde{k}_n} \max_{0\leq j\leq k} |c(k-j)w_2(j)| \\ &= \max_{0\leq k\leq \tilde{k}_n} |w_2(k)| \max_{0\leq j\leq \tilde{k}_n} |c(j)| \\ &\leq \|c\|_\infty \|\mathscr{T}_{\tilde{k}_n} w_2\|_\infty. \end{aligned}$$

Hence for $1 \leq r \leq \infty$ we have:

$$\begin{aligned} &\|\Phi_{\bar{k}_m} d_{p_m}[\tilde{k}_m], \Phi_{\bar{k}_{m+1}} d_{p_{m+1}}[\tilde{k}_{m+1}], \cdots, \Phi_{\bar{k}_n} d_{\tilde{k}_n}[\tilde{k}_n]\| \leq \\ &\qquad \|\Phi_{\bar{k}_m} d_{z_m}[\tilde{k}_m], \Phi_{\bar{k}_{m+1}} d_{z_{m+1}}[\tilde{k}_{m+1}], \cdots, \Phi_{\bar{k}_n} d_{z_n}[\tilde{k}_n]\| + \chi\|c\|\|\mathscr{T}_{\tilde{k}_n} w_2\| \end{aligned}$$

as required. □

The next proposition shows that the disturbance estimates corresponding to a plant $z \in \mathcal{P}$ at time $x \in \mathbb{N}$ bounds a series of disturbance estimates for the same plant up to time x. To improve readability we repeat all relevant equations in Table 5.5.

$$E : \mathcal{W}_e \to map(\mathbb{N}, map(\mathcal{P}, map(\mathbb{N}, \mathbb{R}^h))) \tag{3.14}$$

$$w_2 \mapsto \big[k \mapsto (p \mapsto d_p[k])\big] \tag{3.15}$$

Table 5.5: Details for Proposition 5.9

Proposition 5.9. *Suppose E is given by equations (3.14),(3.15) and satisfies Assumptions 3.4(3)–(4) for $\lambda \in \mathbb{N}$. Let $x \in \mathbb{N}$. Let $q : \mathbb{N} \to \mathcal{P}$ be a switching signal and let $z = q(x)$. Suppose $\bar{a}_j, \tilde{a}_j \in \mathbb{N}$, $1 \leq j \leq i$, $i \in \mathbb{N}$, satisfy:*

$$\tilde{a}_{j-2} < \tilde{a}_j - \bar{a}_j \tag{5.21}$$

$$\tilde{a}_i \leq x. \tag{5.22}$$

Suppose $w_2 \in \mathcal{W}_e$. Let $d_z[k] = E(w_2)(k)(z)$. Then:

$$\|\Phi_{\bar{a}_0} d_z[\tilde{a}_0], \Phi_{\bar{a}_1} d_z[\tilde{a}_1], \cdots, \Phi_{\bar{a}_i} d_z[\tilde{a}_i]\| \leq \|(1,1)\|\|d_z[x]\|$$

where Φ_j is defined by Assumption 3.4(3).

Proof We first claim that for $1 \leq j \leq i$:

$$\big\|\Phi_{\bar{a}_0} d_z[\tilde{a}_0], \Phi_{\bar{a}_1} d_z[\tilde{a}_1], \cdots, \Phi_{\bar{a}_j} d_z[\tilde{a}_j]\big\| \leq \|d_z[\tilde{a}_{j-1}], d_z[\tilde{a}_j]\|. \tag{5.23}$$

Observe that

$$\|x, y\|_r = \big\|\|x\|_r, \|y\|_r\big\|_r, \; x, y \in \mathcal{S}, \; 1 \leq r \leq \infty. \tag{5.24}$$

This proof is by induction. Let $i = j = 1$. For the ease of notation let

$$\mathscr{R}_\sigma d_z[k] = \mathscr{R}_{\sigma,k} d_z[k].$$

Since

$$\|\mathscr{R}_{\bar{a}_l} d_z[\tilde{a}_l]\| \leq \|d_z[\tilde{a}_l]\|,\ 0 \leq l \leq i, \tag{5.25}$$

we have :

$$\begin{aligned}\|\Phi_{\bar{a}_0} d_z[\tilde{a}_0], \Phi_{\bar{a}_1} d_z[\tilde{a}_1]\| &\overset{\text{Ass. (3.4)(3),(5.24)}}{\leq} \|\|\mathscr{R}_{\bar{a}_0} d_z[\tilde{a}_0]\|, \|\mathscr{R}_{\bar{a}_1} d_z[\tilde{a}_1]\|\| \\ &\overset{(5.25),(5.24)}{\leq} \|d_z[\tilde{a}_0], d_z[\tilde{a}_1]\|.\end{aligned}$$

Therefore the base step is shown.

For the inductive step, assume equation (5.23) holds for $2 \leq j \leq i-1$. Then

$$\begin{aligned}\|\Phi_{\bar{a}_0} d_z[\tilde{a}_0], \Phi_{\bar{a}_1} d_z[\tilde{a}_1], \cdots, \Phi_{\bar{a}_i} d_z[\tilde{a}_i]\| &\overset{(5.23),(5.24)}{\leq} \left\| \begin{matrix} \|d_z[\tilde{a}_{i-3}]\|, \|d_z[\tilde{a}_{i-2}]\|, \\ \|\Phi_{\bar{a}_{i-1}} d_z[\tilde{a}_{i-1}]\|, \|\Phi_{\bar{a}_i} d_z[\tilde{a}_i]\| \end{matrix} \right\| \\ &\overset{\text{Ass. (3.4)(3)}}{\leq} \left\| \begin{matrix} \|d_z[\tilde{a}_{i-3}]\|, \|d_z[\tilde{a}_{i-2}]\|, \\ \|\mathscr{R}_{\bar{a}_{i-1}} d_z[\tilde{a}_{i-1}]\|, \|\mathscr{R}_{\bar{a}_i} d_z[\tilde{a}_i]\| \end{matrix} \right\| \\ &\leq \left\| \begin{matrix} \|d_z[\tilde{a}_{i-3}]\|, \|\mathscr{R}_{\bar{a}_{i-1}} d_z[\tilde{a}_{i-1}]\|, \\ \|d_z[\tilde{a}_{i-2}]\|, \|\mathscr{R}_{\bar{a}_i} d_z[\tilde{a}_i]\| \end{matrix} \right\| \\ &\overset{\text{Ass. (3.4)(4)}}{\leq} \left\| \begin{matrix} \|\mathscr{T}_{\tilde{a}_{i-3}} d_z[\tilde{a}_{i-1}]\|, \|\mathscr{R}_{\bar{a}_{i-1}} d_z[\tilde{a}_{i-1}]\|, \\ \|\mathscr{T}_{\tilde{a}_{i-2}} d_z[\tilde{a}_i]\|, \|\mathscr{R}_{\bar{a}_i} d_z[\tilde{a}_i]\| \end{matrix} \right\| \\ &\overset{(5.21),(5.24)}{\leq} \|d_z[\tilde{a}_{i-1}], d_z[\tilde{a}_i]\|.\end{aligned}$$

This completes the inductive step and establishes the claimed inequality (5.23).

We now bound disturbance estimates $d_z[\tilde{a}_i]$ by $d_z[x]$. We will exploit the fact that $q(x) = z$ where $x \geq \tilde{a}_i$ (inequality (5.22)).

We then have with Assumption 3.4(4) that:

$$\begin{aligned}\|d_z[\tilde{a}_i]\| &\overset{z=q(x)}{=} \|d_{q(x)}[\tilde{a}_i]\| \qquad (5.26) \\ &\overset{\text{Ass.(3.4)(4),(5.22)}}{\leq} \|\mathscr{T}_{\tilde{a}_i} d_{q(x)}[x]\| \\ &\overset{\|\cdot\|}{\leq} \|d_z[x]\| \qquad (5.27)\end{aligned}$$

Hence by inequality (5.23) and inequality (5.27) we have:

$$\begin{aligned}\|\Phi_{\bar{a}_0} d_z[\tilde{a}_0], \Phi_{\bar{a}_1} d_z[\tilde{a}_1], \cdots, \Phi_{\bar{a}_i} d_z[\tilde{a}_i]\| &\overset{(5.23)}{\leq} \|d_z[\tilde{a}_{i-1}], d_z[\tilde{a}_i]\| \\ &\overset{(3.4)(4)}{\leq} \|\mathscr{T}_{\tilde{a}_{i-1}} d_z[\tilde{a}_i], d_z[\tilde{a}_i]\| \\ &\overset{\|\cdot\|}{\leq} \|(1,1)\|\|d_z[x]\|.\end{aligned}$$

as required. □

The next key proposition is short and shows that if the algorithm switches at time x to a plant z that the disturbance estimate at this time x, given by $d_z[x] = E(w_2)(z)(x)$, can be bounded by the real disturbance w_0 — as indicated by the gray squares in Figure 5.5. To improve readability we repeat all relevant equations in Table 5.6.

$$y_1 = Pu_1 \tag{2.7}$$
$$u_0 = u_1 + u_2 \tag{2.8}$$
$$y_0 = y_1 + y_2 \tag{2.9}$$
$$E : \mathcal{W}_e \to map(\mathbb{N}, map(\mathcal{P}, map(\mathbb{N}, \mathbb{R}^h))) \tag{3.14}$$
$$w_2 \mapsto \big[k \mapsto (p \mapsto d_p[k])\big] \tag{3.15}$$
$$N : map(\mathbb{N}, map(\mathcal{P}, map(\mathbb{N}, \mathbb{R}^h))) \to map(\mathbb{N}, map(\mathcal{P}, \mathbb{R}^+)) \tag{3.16}$$
$$\big[k \mapsto (p \mapsto d_p[k])\big] \mapsto \big[k \mapsto (p \mapsto \|d_p[k]\| = r_p[k])\big] \tag{3.17}$$
$$G : \mathcal{W}_e \to map(\mathbb{N}, \mathcal{P}^*) \tag{4.25}$$
$$M : (map(\mathbb{N}, map(\mathcal{P}, \mathbb{R}^+)), map(\mathbb{N}, \mathcal{P}^*)) \to map(\mathbb{N}, \mathcal{P}^*) \tag{4.26}$$
$$\big[k \mapsto (p \mapsto r_p[k]), k \mapsto G(k)\big] \mapsto \big[k \mapsto q_f(k)\big] \tag{4.27}$$
$$q_f(k) := \operatorname*{argmin}_{p \in G(k)} r_p[k], \ \forall k \in \mathbb{N} \tag{4.28}$$
$$D : map(\mathbb{N}, \mathcal{P}) \to map(\mathbb{N}, \mathcal{P}) \tag{4.30}$$
$$[k \mapsto q_f(k)] \mapsto [k \mapsto q(k)] \tag{4.31}$$
$$q(k) := \begin{cases} q_f(k) & \text{if} \quad k - k_s(k) \geq \Delta(q(k_s(k))) \\ q(k_s(k)) & \text{else} \end{cases} \tag{4.32}$$
$$k_s(k) := \max\{i \in \mathbb{N} \mid 0 \leq i \leq k, \ q(i) \neq q(i-1)\} \tag{4.33}$$
$$L_k := \{l_0 = 0, l_1, l_2 \cdots\} = \{l \in \mathbb{N} \mid q(l-1) \neq q(l), \ 0 \leq l \leq k\} \tag{5.4}$$
$$V(l_i, l_{i+1}) := \left\{ a \in \mathbb{N} \ \middle| \ \begin{array}{c} \exists b \in \mathbb{N} \text{ s.t. } a = l_i + b\Delta(q(l_i)), \\ l_i < a \leq l_{i+1} - \Delta(q(l_i)) \end{array} \right\} \tag{5.5}$$
$$Q_k = \{k_0 = 0, k_1, \ k_2, \ \cdots\}, \ 0 \leq k_i \leq k_{i+1} \leq k \tag{5.6}$$
$$Q_k := L_k \cup \bigcup_{i \geq 0} \{V(l_i, l_{i+1}) \mid l_i, l_{i+1} \in L_k\} \tag{5.7}$$

TABLE 5.6: Details for Proposition 5.10

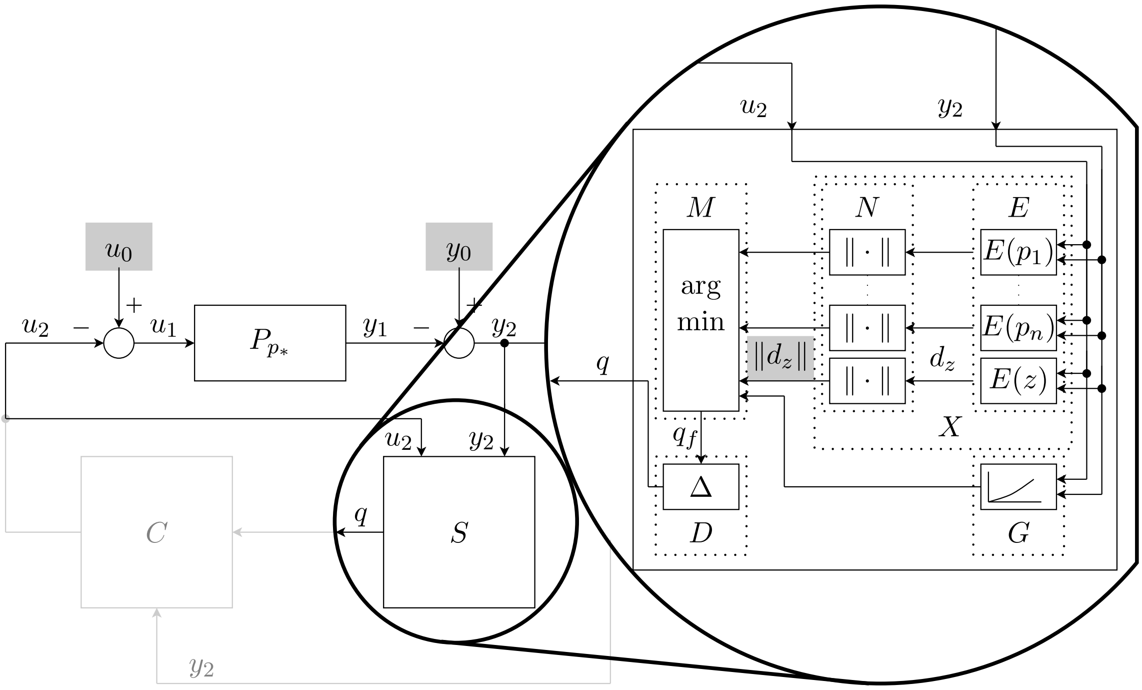

FIGURE 5.5: Bounding $d_z[x|$ in terms of $w_0 = (u_0, y_0)^\top$ for $z = q(x) = DM(X,G)(x)$

Proposition 5.10. *Let* $1 \leq r \leq \infty$*. Suppose* $\mathcal{P}^U \subset \mathcal{P}$ *and let* $p_* \in \mathcal{P}^U$*. Suppose* Δ *is a given delay transition function and suppose the delay operator* D *is given by equations* (4.30)–(4.33)*. Suppose* G *is a plant-generating operator. Suppose* E *satisfies Assumptions 3.4(1)–(4) for some* $\lambda \in \mathbb{R}$ *and the switching operator* $S = DM(NE, G)$ *is given by equations* (3.14)–(3.17),(4.26)–(4.28),(4.30)–(4.33)*. Let* $k \in \mathbb{N}$*. Suppose* $(w_0, w_1, w_2) \in \mathcal{W} \times \mathcal{W}_e \times \mathcal{W}_e$ *satisfy equations* (2.7)–(2.9) *for* $P = P_{p_*}$*. Let* $x \in Q_k,\ z = q(x) = S(w_2)(x)$ *where* Q_k *is defined by equations* (5.4)–(5.7) *and suppose* $p_* \in G(x)$*. Then:*

$$\|E(w_2)(z)(x)\| = \|d_z[x]\| \leq \mu\|w_0\|.$$

Proof By the definition of the switching algorithm $q_f(t) = M(NE, G)(t),\ t \in \mathbb{N}$, will always point to the plant which corresponding disturbance estimates are minimal. Since $p_* \in G(x)$ and by the definition of M we have

$$\|d_{q_f(x)}[x]\| = \inf_{p \in G(x)} \|d_p[x]\| \leq \|d_{p_*}[x]\|.$$

Since by the definition of D, $q_f(x) = q(x)$ it follows that

$$\|d_{q(x)}[x]\| = \|d_{q_f(x))}[x]\| \leq \|d_{p_*}[x]\| \tag{5.28}$$

hence

$$\begin{aligned}
\|d_z[x]\| &= \|d_{q(x)}[x]\| \\
&\overset{(5.28)}{\leq} \|d_{p_*}[x]\| \\
&\overset{\text{Ass.}(3.4)(2)}{\leq} \mu\|\mathscr{T}_x w_0\| \\
&\overset{\|\cdot\|}{\leq} \mu\|w_0\|
\end{aligned}$$

as required. □

4 Gain bounds for non-final switching intervals

Before we commence with establishing gain bounds, we give a intermediate result that is self-contained and purely combinatorial.

Proposition 5.11. *Let $1 \leq r \leq \infty$ and*

$$\xi = \begin{cases} r & \text{for} \quad 1 \leq r < \infty \\ 1 & \text{for} \quad r = \infty \end{cases} .$$

Let $z, f, \beta, \epsilon : \mathbb{N} \to \mathbb{R}^+$ and $a, b, d, e \in \mathbb{R}^+$, $a < 1$. Let $m, n \in \mathbb{N}$ and suppose for all $m \leq i \leq n$:

$$z_{i+1}^{\xi} \leq a z_i^{\xi} + d\beta_i^{\xi} \tag{5.29}$$

$$f_i^{\xi} \leq b z_i^{\xi} + e\epsilon_i^{\xi}. \tag{5.30}$$

Then:

$$\left\| z|_{[m+1,n+1]}, f|_{[m,n]} \right\| \leq \tilde{\gamma}_3(\mathcal{G})|z_m| + \tilde{\gamma}_4(\mathcal{G})\|\beta|_{[m,n]}\| + \tilde{\gamma}_5(\mathcal{G})\|\epsilon|_{[m,n]}\|$$

where

$$\begin{aligned}
\mathcal{G} &= (a, b, d, e) \\
\tilde{\gamma}_3(\mathcal{G}) &= \begin{cases} (1 + b^{1/r})\left(\frac{a}{1-a}\right)^{1/r} + b^{1/r} & \text{for} \quad 1 \leq r < \infty \\ \max\{1, b\}a + b & \text{for} \quad r = \infty \end{cases} \\
\tilde{\gamma}_4(\mathcal{G}) &= \begin{cases} (1 + b^{1/r})\left(\frac{d}{1-a}\right)^{1/r} & \text{for} \quad 1 \leq r < \infty \\ \max\{1, b\}\frac{d}{1-a} & \text{for} \quad r = \infty \end{cases} \\
\tilde{\gamma}_5(\mathcal{G}) &= \begin{cases} e^{1/r} & \text{for} \quad 1 \leq r < \infty \\ e & \text{for} \quad r = \infty \end{cases}
\end{aligned}$$

Proof Let $1 \leq \xi = r < \infty$. Define for $i \in \mathbb{N}$

$$\hat{z}_i := z_i^r,\ \hat{f}_i := f_i^r,\ \hat{\beta}_i := \beta_i^r,\ \hat{\epsilon}_i := \epsilon_i^r.$$

By equation (5.29) we have

$$\begin{aligned}
\hat{z}_{m+1} &\leq a\hat{z}_m + d\hat{\beta}_m \\
\hat{z}_{m+2} &\leq a^2\hat{z}_m + d\left(a\hat{\beta}_m + \hat{\beta}_{m+1}\right) \\
\hat{z}_{m+3} &\leq a^3\hat{z}_m + d\left(a^2\hat{\beta}_m + a\hat{\beta}_{m+1} + \hat{\beta}_{m+2}\right) \\
\vdots\ &\ \vdots\ \ \vdots \\
\hat{z}_{n+1} &\leq a^{n-m+1}\hat{z}_m + d\left(\hat{\beta}_m a^{n-m} + \hat{\beta}_{m+1}a^{n-m-1} + \cdots + \hat{\beta}_{n-1}a + \hat{\beta}_n\right).
\end{aligned}$$

Summing vertically gives us

$$\begin{aligned}
\sum_{i=m+1}^{n+1} \hat{z}_i &\leq \hat{z}_m \sum_{i=1}^{n-m+1} a^i + d\left(\hat{\beta}_m \sum_{i=0}^{n-m} a^i + \hat{\beta}_{m+1} \sum_{i=0}^{n-m-1} a^i + \cdots + \hat{\beta}_{n-1}\sum_{i=0}^{1} a^i + \hat{\beta}_n\right) \\
&\leq \hat{z}_m \sum_{i=1}^{n-m+1} a^i + d\sum_{j=m}^{n} \hat{\beta}_j \sum_{i=0}^{n-j} a^i
\end{aligned}$$

Since $a < 1$ we have that for any $j > 0$

$$\sum_{i=0}^{j} a^i \leq \frac{1}{1-a} \text{ and } \sum_{i=1}^{j} a^i \leq \frac{a}{1-a},$$

hence

$$\sum_{i=m+1}^{n+1} \hat{z}_i \leq \frac{1}{1-a}\left(a\hat{z}_m + d\sum_{i=m}^{n} \hat{\beta}_i\right)$$

and therefore

$$\begin{aligned}
\|z|_{[m+1,n+1]}\|_r &= \left(\sum_{i=m+1}^{n+1} \hat{z}_i\right)^{1/r} \\
&\leq \left(\frac{1}{1-a}\right)^{1/r}\left(a\hat{z}_m + d\sum_{i=m}^{n} \hat{\beta}_i\right)^{1/r} \\
&= \left(\frac{1}{1-a}\right)^{1/r}\left(az_m^r + d\sum_{i=m}^{n} \beta_i^r\right)^{1/r} \\
&\leq \left(\frac{1}{1-a}\right)^{1/r}\left(a^{1/r}|z_m| + d^{1/r}\left\|\beta|_{[m,n]}\right\|_r\right). \qquad (5.31)
\end{aligned}$$

By inequality (5.30) we have

$$\begin{aligned}
\left\|f|_{[m,n]}\right\|_r &\leq \left(b\sum_{i=m}^{n} z_i^r + e\sum_{i=m}^{n}\epsilon_i^r\right)^{1/r} \\
&\leq b^{1/r}\left\|z|_{[m,n]}\right\|_r + e^{1/r}\left\|\epsilon|_{[m,n]}\right\|_r. \qquad (5.32)
\end{aligned}$$

By inequalities (5.32) and equations (5.31) we arrive at

$$\begin{aligned}
\left\|z|_{[m+1,n+1]}, f|_{[m,n]}\right\|_r &\leq \left\|z|_{[m+1,n+1]}\right\|_r + \left\|f|_{[m,n]}\right\|_r \\
&\leq \left\|z|_{[m+1,n+1]}\right\|_r + b^{1/r}\left\|z|_{[m,n]}\right\|_r + e^{1/r}\left\|\epsilon|_{[m,n]}\right\|_r \\
&\leq \left(1+b^{1/r}\right)\left\|z|_{[m+1,n+1]}\right\|_r + b^{1/r}|z_m| + e^{1/r}\left\|\epsilon|_{[m,n]}\right\|_r \\
&\leq \left(1+b^{1/r}\right)\left(\frac{1}{1-a}\right)^{1/r}\left(a^{1/r}|z_m| + d^{1/r}\left\|\beta|_{[m,n]}\right\|_r\right) \\
&\quad + b^{1/r}|z_m| + e^{1/r}\left\|\epsilon|_{[m,n]}\right\|_r \\
&\leq \left((1+b^{1/r})\left(\frac{1}{1-a}\right)^{1/r} a^{1/r} + b^{1/r}\right)|z_m| \\
&\quad + (1+b^{1/r})\left(\frac{1}{1-a}\right)^{1/r} d^{1/r}\|\beta|_{[m,n]}\| + e^{1/r}\left\|\epsilon|_{[m,n]}\right\|_r \\
&\leq \tilde{\gamma}_3(\mathcal{G})|z_m| + \tilde{\gamma}_4(\mathcal{G})\|\beta|_{[m,n]}\|_r + \tilde{\gamma}_5(\mathcal{G})\|\epsilon|_{[m,n]}\|_r
\end{aligned}$$

as required.

Let $r=\infty$, so $\xi = 1$. By equation (5.29) we have

$$\begin{aligned}
z_{m+1} &\leq a z_m + d\beta_m \\
z_{m+2} &\leq a^2 z_m + d\left(a\beta_m + \beta_{m+1}\right) \\
z_{m+3} &\leq a^3 z_m + d\left(a^2\beta_m + a\beta_{m+1} + \beta_{m+2}\right) \\
\vdots &\quad \vdots \quad \vdots \\
z_{n+1} &\leq a^{n-m+1} z_m + d\left(\beta_m a^{n-m} + \beta_{m+1}a^{n-m-1} + \cdots + \beta_{n-1}a + \beta_n\right).
\end{aligned}$$

Taking norms leads to

$$\begin{aligned}
\left\|z|_{[m+1,n+1]}\right\|_\infty &= \max_{m+1\leq j\leq n+1}|z_j| \\
&\leq a|z_m| + d\sum_{i=0}^{n-m} a^i\|\beta|_{[m,n]}\|_\infty \\
&\leq a|z_m| + \frac{d}{1-a}\|\beta|_{[m,n]}\|_\infty.
\end{aligned}$$

Furthermore by equation (5.30) we have

$$\begin{aligned}\left\|f|_{[m,n]}\right\|_\infty &= \max_{m\le j\le n} |f_j| \\ &\le b\left\|z|_{[m,n]}\right\|_\infty + e\left\|\epsilon|_{[m,n]}\right\|_\infty .\end{aligned}$$

Substitutions lead to

$$\begin{aligned}\left\|z|_{[m+1,n+1]}, f|_{[m,n]}\right\|_\infty &\le \max\{\|z|_{[m+1,n+1]}\|_\infty, b\|z_{[m,n]}\|_\infty + e\|\epsilon_{[m,n]}\|_\infty\} \\ &\le \max\{\|z|_{[m+1,n+1]}\|_\infty, b\|z_{[m+1,n+1]}\|_\infty + b|z_m| + e\|\epsilon_{[m,n]}\|_\infty\} \\ &\le \max\{1,b\}\|z|_{[m+1,n+1]}\|_\infty + b|z_m| + e\|\epsilon_{[m,n]}\|_\infty \\ &\le \max\{1,b\}\big(a|z_m| + \frac{d}{1-a}\|\beta|_{[m,n]}\|_\infty\big) + b|z_m| + e\|\epsilon_{[m,n]}\|_\infty \\ &\le \big(\max\{1,b\}a + b\big)|z_m| + \max\{1,b\}\frac{d}{1-a}\|\beta|_{[m,n]}\|_\infty + e\|\epsilon_{[m,n]}\|_\infty \\ &\le \tilde\gamma_3(\mathcal{G})|z_m| + \tilde\gamma_4(\mathcal{G})\|\beta|_{[m,n]}\|_\infty + \tilde\gamma_5(\mathcal{G})\|\epsilon_{[m,n]}\|_\infty\end{aligned}$$

as required. □

In Proposition 5.7 we established a gain relationship between w_2 and disturbance signals w_0^p which are consistent with $p \in \mathcal{P}$ and $w_2 \in \mathcal{W}_e$ over some finite interval. Since it is the overall goal to establish a bound on the gain from the real world disturbances w_0 to the internal signals w_2 we need to bound the consistent disturbance signals w_0^p by the real world disturbances w_0.

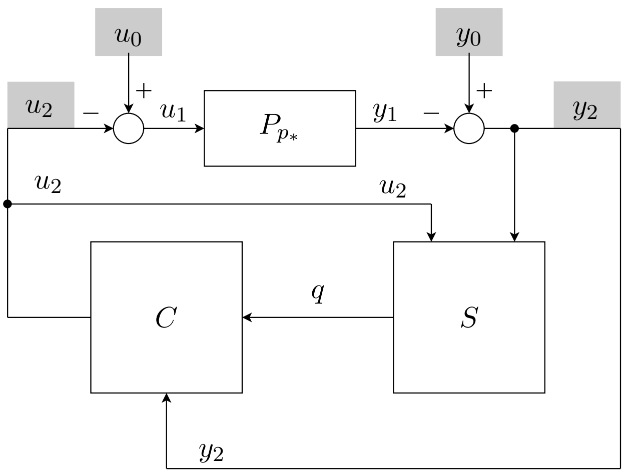

FIGURE 5.6: Bounding intervals of $w_2 = (u_2, y_2)^\top$, corresponding to ongoing switching times, in terms of $w_0 = (u_0, y_0)^\top$

We do this by considering intervals $[k_m, k_n],\ m,n \in \mathbb{N},\ m \le n,\ k_m, k_n \in Q_k$ where all intermediate switching times are ongoing, i.e. $k_i \in O_k,\ m \le i \le n$ and then use the fact that after a series of ongoing switches there must follow a final switch hence Proposition 5.10 is applicable. The next result establishes bounds on intervals of w_2 in terms of w_0 — as indicated by the gray squares in Figure 5.6. Before we give the statement we make the following definition. To improve readability we repeat all relevant equations in Table 5.7 and Table 5.8.

$$X : \mathcal{W}_e \to map(\mathbb{N}, map(\mathcal{P}, \mathbb{R}^+)) : w_2 \mapsto [k \to (p \mapsto r_p[k])] \tag{3.7}$$

$$G : \mathcal{W}_e \to map(\mathbb{N}, \mathcal{P}^*) \tag{4.25}$$

$$M : (map(\mathbb{N}, map(\mathcal{P}, \mathbb{R}^+)), map(\mathbb{N}, \mathcal{P}^*)) \to map(\mathbb{N}, \mathcal{P}^*) \tag{4.26}$$

$$\big[k \mapsto (p \mapsto r_p[k]), k \mapsto G(k)\big] \mapsto \big[k \mapsto q_f(k)\big] \tag{4.27}$$

$$q_f(k) := \operatorname*{argmin}_{p \in G(k)} r_p[k], \;\; \forall k \in \mathbb{N} \tag{4.28}$$

$$D : map(\mathbb{N}, \mathcal{P}) \to map(\mathbb{N}, \mathcal{P}) \tag{4.30}$$

$$[k \mapsto q_f(k)] \mapsto [k \mapsto q(k)] \tag{4.31}$$

$$q(k) := \begin{cases} q_f(k) & \text{if} \quad k - k_s(k) \geq \Delta(q(k_s(k))) \\ q(k_s(k)) & \text{else} \end{cases} \tag{4.32}$$

$$k_s(k) := \max\{i \in \mathbb{N} \mid 0 \leq i \leq k, \;\; q(i) \neq q(i-1)\} \tag{4.33}$$

$$C : \mathcal{Y}_e \to \mathcal{U}_e : y_2 \mapsto u_2 \tag{4.34}$$

$$u_2(k) = C_{K(q(k))}(y_2 - \mathscr{T}_{k_s(k)-1} y_2)(k) \tag{4.35}$$

$$\begin{array}{rcl} J(r)\alpha^r(p, K(p), \Delta(p) - \sigma, \sigma) & \leq & l(p) < 1, \;\; \forall p \in \mathcal{P}^U \text{ if } 1 \leq r < \infty \\ \alpha(p, K(p), \Delta(p) - \sigma, \sigma) & \leq & l(p) < 1, \;\; \forall p \in \mathcal{P}^U \text{ if } r = \infty \end{array} \tag{5.16}$$

TABLE 5.7: Details for the definition of standard EMMSAC in Definition 5.12

Definition 5.12. *An EMMSAC algorithm is said to be standard if it satisfies:*

- $K : \mathcal{P} \to \mathcal{C}$ *is a given control design satisfying Assumption 4.1(1),(2)*
- $\Delta : \mathcal{P} \to \mathbb{N}$ *is a delay transition function satisfying Assumption* (5.3) *and the delay operator D is given by equations* (4.30)–(4.33)
- K, Δ *and a given attenuation function $l : \mathcal{P} \to [0, 1)$ satisfy inequality* (5.16)
- E *satisfies Assumptions 3.4(1)–(5) where*

$$\lambda = \max_{p \in \mathcal{P}^U} (2\Delta(p) + \sigma) \tag{5.33}$$

- *The switching operator $S = DM(X, G)$ is given by equations* (3.7),(4.26)–(4.28) *and* (4.30)–(4.33)
- *The switching controller C is defined by equations* (4.34),(4.35).

$$
\begin{aligned}
y_1 &= Pu_1 && (2.7)\\
u_0 &= u_1 + u_2 && (2.8)\\
y_0 &= y_1 + y_2 && (2.9)\\
u_2 &= Cy_2 && (2.10)
\end{aligned}
$$

$$L_k := \{l_0 = 0, l_1, l_2 \cdots\} = \{l \in \mathbb{N} \mid q(l-1) \neq q(l),\ 0 \leq l \leq k\} \tag{5.4}$$

$$V(l_i, l_{i+1}) := \left\{ a \in \mathbb{N} \;\middle|\; \begin{array}{c} \exists b \in \mathbb{N} \text{ s.t. } a = l_i + b\Delta(q(l_i)), \\ l_i < a \leq l_{i+1} - \Delta(q(l_i)) \end{array} \right\} \tag{5.5}$$

$$Q_k = \{k_0 = 0, k_1,\ k_2,\ \cdots\},\ 0 \leq k_i \leq k_{i+1} \leq k \tag{5.6}$$

$$Q_k := L_k \cup \bigcup_{i \geq 0} \{V(l_i, l_{i+1}) \mid l_i, l_{i+1} \in L_k\} \tag{5.7}$$

TABLE 5.8: Details for Proposition 5.13

Proposition 5.13. *Let $1 \leq r \leq \infty$. Suppose $p_* \in \mathcal{P}^U \subset \mathcal{P}$ where $\mathcal{P}^U$ satisfies Assumption 5.2. Let $P = P_{p_*}$. Let U be a monotonic plant generating operator and suppose (H, ν) defines a monotonic cover for U. Suppose the EMMSAC algorithm is standard. Let $k \in \mathbb{N}$. Let G be a plant generating operator that satisfies $G(j) \subset U(j),\ j \leq k$. Suppose $(w_0, w_1, w_2) \in \mathcal{W} \times \mathcal{W}_e \times \mathcal{W}_e$ satisfy the closed loop $[P, C]$ equations (2.7)–(2.10) over the interval $[0, k)$. Let $k_i,\ i \in \mathbb{N}$ be defined by equations (5.4)–(5.7) and suppose $k_{n+1} \leq k$. Let $m, n \in \mathbb{N}$, suppose $F_k \cap [k_m - \sigma, k_{n+1}] = \emptyset$. If $p_* \in G(j),\ j \geq k_m$,*

$$\|c\|\Big(\gamma_4(U(j)) + \gamma_5(U(j))\Big)\chi_\nu(H(j), \nu(j)) < 1,\ \forall j \leq k \tag{5.34}$$

and $\alpha_{OP}(U(k)) < 1$ then

$$\|\mathscr{T}_{k_{n+1}-1} w_2\| \;\leq\; \gamma_6(U(k), H(k), \nu(k))\|\mathscr{T}_{k_m - 1} w_2\| + \gamma_7(U(k), H(k), \nu(k))\|w_0\|$$

where for $\mathcal{Q}_1 \subset \mathcal{P}^U,\ \mathcal{Q}_2 \subset \mathcal{P}^H,\ \epsilon : \mathcal{P} \to \mathbb{R}^+$:

$$
\begin{aligned}
\chi_\nu(\mathcal{Q}_2, \epsilon) &= 2 \sup_{p \in \mathcal{Q}_2} \epsilon(p) && (5.35)\\
\gamma_3(\mathcal{Q}_1) &= \begin{cases} (1 + \alpha_{OS}^{1/r}(\mathcal{Q}_1))\left(\frac{\alpha_{OP}(\mathcal{Q}_1)}{1-\alpha_{OP}(\mathcal{Q}_1)}\right)^{1/r} + \alpha_{OS}^{1/r}(\mathcal{Q}_1) & \text{if } \ 1 \leq r < \infty \\ \max\{1, \alpha_{OS}(\mathcal{Q}_1))\}\alpha_{OP}(\mathcal{Q}_1) + \alpha_{OS}(\mathcal{Q}_1) & \text{if } \quad r = \infty \end{cases}\\
\gamma_4(\mathcal{Q}_1) &= \begin{cases} (1 + \alpha_{OS}^{1/r}(\mathcal{Q}_1))\left(\frac{\beta_{OP}(\mathcal{Q}_1)}{1-\alpha_{OP}(\mathcal{Q}_1)}\right)^{1/r} & \text{if } \ 1 \leq r < \infty \\ \max\{1, \alpha_{OS}(\mathcal{Q}_1)\})\frac{\beta_{OP}(\mathcal{Q}_1)}{1-\alpha_{OP}(\mathcal{Q}_1)} & \text{if } \quad r = \infty \end{cases}\\
\gamma_5(\mathcal{Q}_1) &= \begin{cases} \beta_{OS}^{1/r}(\mathcal{Q}_1) & \text{if } \ 1 \leq r < \infty \\ \beta_{OS}(\mathcal{Q}_1) & \text{if } \quad r = \infty \end{cases}\\
\gamma_6(\mathcal{Q}_1, \mathcal{Q}_2, \epsilon) &= \frac{1 + \gamma_3(\mathcal{Q}_1)}{1 - \|c\|\big(\gamma_4(\mathcal{Q}_1) + \gamma_5(\mathcal{Q}_1)\big)\chi_\nu(\mathcal{Q}_2, \epsilon)}\\
\gamma_7(\mathcal{Q}_1, \mathcal{Q}_2, \epsilon) &= \frac{2^{1/r}\mu|\mathcal{Q}_2|^{1/r}\big(\gamma_4(\mathcal{Q}_1) + \gamma_5(\mathcal{Q}_1)\big)}{1 - \|c\|\big(\gamma_4(\mathcal{Q}_1) + \gamma_5(\mathcal{Q}_1)\big)\chi_\nu(\mathcal{Q}_2, \epsilon)}
\end{aligned}
$$

and $\alpha_{OP}, \alpha_{OS}, \beta_{OP}, \beta_{OS}$ *are from Proposition 5.7 and* c *is as in Assumption 3.4(5).*

Proof Let $1 \leq r \leq \infty$ and

$$\xi = \begin{cases} r & \text{if} \quad 1 \leq r < \infty \\ 1 & \text{if} \quad r = \infty \end{cases} .$$

Let $k \in \mathbb{N}$. Let $(w_0, w_1, w_2) \in \mathcal{W} \times \mathcal{W}_e \times \mathcal{W}_e$ denote the solution to the closed loop equations (2.7)–(2.10) with $P = P_{p*}$ and C as in equations (4.34),(4.35). Let the intervals

$$A_i = [k_i - \sigma, k_i), \;\; B_i = [k_i, k_{i+1} - \sigma)$$

be defined by equations (5.13),(5.14). In particular $(w_0, w_1, w_2) \in \mathcal{W} \times \mathcal{W}_e \times \mathcal{W}_e$ satisfy equations (2.7)–(2.10) on the intervals $A_i \cup B_i \cup A_{i+1}$ where

$$A_i \cup B_i \cup A_{i+1} \subseteq [k_m - \sigma, k_{n+1}) \subseteq [0, k)$$

for $m \leq i \leq n$.

For $k_i \in Q_k$, define $\bar{k}_i, \; \tilde{k}_i$ as follows. Let

$$\begin{aligned} \bar{k}_i &= k_{i+1} - k_i + \sigma - 1 \\ \tilde{k}_i &= k_{i+1} - 1 \end{aligned}$$

and note that $A_i \cup B_i \cup A_{i+1} = [\tilde{k}_i - \bar{k}_i, \tilde{k}_i]$.

We now intend to apply Proposition 5.7. For that purpose we first observe the following facts.

By Lemma 5.4, Assumption 5.3 and equation (5.33) we have

$$0 \leq \bar{k}_i = k_{i+1} - k_i + \sigma - 1 \leq 2\Delta(q(k_i)) + \sigma \leq \lambda \tag{5.36}$$

Let $p = q(k_i)$. Define

$$w_0^p(k) \;\; = \;\; \begin{cases} \Phi_{\bar{k}_i} d_p[\tilde{k}_i](k) & \text{if} \qquad k \in A_i \cup B_i \cup A_{i+1} \\ 0 & \text{otherwise} \end{cases} .$$

By Assumption 3.4(3) we know that:

$$\Phi_{\bar{k}_i} d_p[\tilde{k}_i] \in \mathcal{N}_p^{[\tilde{k}_i - \bar{k}_i, \tilde{k}_i]}(w_2).$$

For every $k_i \in Q_k$ let $w_2^c \in \mathcal{W}_e$ satisfy

$$w_2^c(k) = \begin{cases} w_2(k) & \text{if} \quad k \in B_i \cup A_{i+1} \quad \text{and} \quad k_i \in L_k \\ w_2(k) & \text{if} \quad k \in A_i \cup B_i \cup A_{i+1} \quad \text{and} \quad k_i \in Q_k \setminus L_k \\ 0 & \text{otherwise} \end{cases}.$$

Note that w_2, w_2^c satisfy equations (5.17),(5.18) of Proposition 5.7.

There exists a $w_1^p \in \mathcal{W}_e$ such that

$$(w_0^p, w_1^p, w_2) \in \mathcal{W}_\times \mathcal{W}_e \times \mathcal{W}_e$$

satisfies equations (2.7)–(2.10) for $P = P_p$ and C as in equation (4.35) on the intervals

$$A_i \cup B_i \cup A_{i+1} = [k_i - \sigma, k_{i+1}) = [\tilde{k}_i - \bar{k}_i, \tilde{k}_i].$$

To see this observe that w_2 is generated by the special structure of C, i.e. the controller C_c at time k_i is initialised to zero if $k_i \in L_k$ and inherits an initial value at time k_i determined from $w_2|_{A_i}$ if $k_i \in Q_k \setminus L_k$.

Define

$$\begin{aligned} a &= \alpha_{OP}(U(k)) < 1 \\ b &= \alpha_{OS}(U(k)) \\ d &= \beta_{OP}(U(k)) \\ e &= \beta_{OS}(U(k)) \\ z_i &= \|w_2|_{A_i}\|_r \geq \|w_2^c|_{A_i}\|_r \\ f_i &= \|w_2|_{B_i}\|_r = \|w_2^c|_{B_i}\|_r \\ \beta_i &= \|w_0^{q(k_i)}|_{A_i \cup B_i \cup A_{i+1}}\|_r = \|\Phi_{\bar{k}_i} d_{q(k_i)}[\tilde{k}_i]\|_r \\ \epsilon_i &= \|w_0^{q(k_i)}|_{A_i \cup B_i}\|_r \leq \beta_i = \|\Phi_{\bar{k}_i} d_{q(k_i)}[\tilde{k}_i]\|_r \end{aligned}$$

where we note that since U is monotonic, hence $G(k_i) \subset U(k_i) \subset U(k)$, it follows for all $k_i \in Q_k$ that:

$$\begin{aligned} \alpha_{OP}(G(k_i)) &\leq \alpha_{OP}(U(k_i)) \leq \alpha_{OP}(U(k)) \\ \alpha_{OS}(G(k_i)) &\leq \alpha_{OS}(U(k_i)) \leq \alpha_{OS}(U(k)) \\ \beta_{OP}(G(k_i)) &\leq \beta_{OP}(U(k_i)) \leq \beta_{OP}(U(k)) \\ \beta_{OS}(G(k_i)) &\leq \beta_{OS}(U(k_i)) \leq \beta_{OS}(U(k)). \end{aligned}$$

Since $\|w_2|_{B_i}\|_r = \|w_2^c|_{B_i}\|_r$ it follows from Proposition 5.7 that:

$$\begin{aligned} z_{i+1}^{\xi} &\leq az_i^{\xi} + d\beta_i^{\xi} \\ f_i^{\xi} &\leq bz_i^{\xi} + e\epsilon_i^{\xi}. \end{aligned}$$

Since $\epsilon_i \leq \beta_i$ it follows that $\|\epsilon|_{[m,n]}\| \leq \|\beta|_{[m,n]}\|$ and by Proposition 5.11 we then have for $1 \leq r \leq \infty$ that:

$$\begin{aligned} \|w_2|_{[k_m,k_{n+1})}\| &= \Big\| \|w_2|_{A_{m+1}}\|, \|w_2|_{A_{m+2}}\|, \cdots, \|w_2|_{A_{n+1}}\|, \\ &\quad \|w_2|_{B_m}\|, \|w_2|_{B_{m+1}}\|, \cdots, \|w_2|_{B_n}\| \Big\| \qquad (5.37) \\ &= \Big\| z|_{[m+1,n+1]}, f|_{[m,n]} \Big\| \\ &\leq \gamma_3(U(k))|z_m| + \gamma_4(U(k))\|\beta|_{[m,n]}\| + \gamma_5(U(k))\|\epsilon|_{[m,n]}\| \\ &\leq \gamma_3(U(k))|z_m| + \big(\gamma_4(U(k)) + \gamma_5(U(k))\big)\|\beta|_{[m,n]}\|. \qquad (5.38) \end{aligned}$$

It remains to show that $\big\|\beta|_{[m,n]}\big\|$, $|z_m|$ are bounded by $\|w_0\|$ and $\|w_2\|$.

Recall that

$$\mathscr{R}_i d_p[j] := \mathscr{R}_{i,j} d_p[j], \ i \leq j, \ p \in \mathcal{P},$$

also recall that

$$\|x, y\|_r = \big\| \|x\|_r, \|y\|_r \big\|_r, \ x, y \in \mathcal{S}, \ 1 \leq r \leq \infty. \qquad (5.39)$$

By Assumption 3.4(3) we have

$$\|w_0^{q(k_i)}|_{A_i \cup B_i \cup A_{i+1}}\| = \|\Phi_{\bar{k}_i} d_{q(k_i)}[\tilde{k}_i]\| \leq \|\mathscr{R}_{\bar{k}_i} d_{q(k_i)}[\tilde{k}_i]\|.$$

Let $m \leq i \leq n$. For $k_i \in Q_k$ let $p_i \in H(k)$ such that:

$$q(k_i) \in B_\chi(p_i, \nu(k)(p_i))$$

where the existence of such a $p_i \in H(k)$ is guaranteed since

$$q(k_j) \in G(k_j) \subset U(k_j) \subset R(k_j) \subset R(k),$$

where $R(j) = \cup_{p \in H(j)} B_\chi(p, \nu(j)(p))$.

Let $z_i = q(F_k(p_i))$ and let $B_i = B_\chi(p_i, \nu(k)(p_i))$ hence $q(k_i), z_i \in B_i$. Observe that $\tilde{k}_i = k_{i+1} - 1 \leq \tilde{k}_{i+1} = k_{i+2} - 1$ and that $0 \leq \bar{k}_i \leq \lambda$ (equation (5.36)).

We are now in the position to apply Proposition 5.8. With equation (5.39) we obtain:

$$\begin{aligned} \|\beta|_{[m.n]}\| &= \| \|\Phi_{\bar{k}_m} d_{q(k_m)}[\tilde{k}_m]\|, \|\Phi_{\bar{k}_{m+1}} d_{q(k_{m+1})}[\tilde{k}_{m+1}]\|, \cdots, \|\Phi_{\bar{k}_n} d_{q(k_n)}[\tilde{k}_n]\| \| \\ &= \|\Phi_{\bar{k}_m} d_{q(k_m)}[\tilde{k}_m], \Phi_{\bar{k}_{m+1}} d_{q(k_{m+1})}[\tilde{k}_{m+1}], \cdots, \Phi_{\bar{k}_n} d_{q(k_n)}[\tilde{k}_n]\| \\ &\leq \|\Phi_{\bar{k}_m} d_{z_m}[\tilde{k}_m], \Phi_{\bar{k}_{m+1}} d_{z_{m+1}}[\tilde{k}_{m+1}], \cdots, \Phi_{\bar{k}_n} d_{z_n}[\tilde{k}_n]\| \\ &\quad + \chi_\nu(H(k), \nu(k))\|c\|\|\mathscr{T}_{\tilde{k}_n} w_2\| \qquad (5.40) \end{aligned}$$

where by equations (5.20) and equation (5.35)

$$\chi = \max_{m\leq i\leq n} \sup_{p,q\in B_i} \chi(p,q) \leq 2 \sup_{p\in H(k)} \nu(k)(p) = \chi_\nu(H(k),\nu(k)).$$

We now bound the term $\|\Phi_{\bar{k}_m} d_{z_m}[\tilde{k}_m],...\|$ in terms of $\|w_0\|$ using Proposition 5.9.

Let $p \in H(k)$. Let

$$\{a_0, a_1, \cdots, a_{i-1}, a_i\} = [k_m, k_n] \cap Q_k(p, \nu(k)(p))$$

be the ordered set ($a_j \leq a_{j+1},\ 0 \leq j \leq i-1$) of switching times corresponding to the plants within the set $B_\chi(p, \nu(k)(p))$ over the interval $[k_m, k_n]$. For $j \in \mathbb{N}$, let:

$$\begin{aligned} \bar{a}_j &= a_{j+1} - a_j + \sigma - 1 \\ \tilde{a}_j &= a_{j+1} - 1. \end{aligned}$$

First observe that

$$\tilde{a}_{j-2} < \tilde{a}_j - \bar{a}_j.$$

Since $F_k \cap [k_m - \sigma, k_{n+1}] = \emptyset$, it follows that $a_j \in O_k(p, \nu(k)(p)),\ 0 \leq j \leq i$, hence the switching sequence $q(a_j)$ will switch back to a plant within the neighbourhood $B_\chi(p, \nu(k)(p)$ for one final time in $[0, k]$, after $\tilde{a}_i$, i.e. there exists a time $x \in Q_k$ such that:

$$k \geq x = F_k(p, \nu(k)(p)) \geq \tilde{a}_i = a_{i+1} - 1 \geq k_n$$

and $z = q(x) \in B_\chi(p, \nu(k)(p))$.

Define

$$\psi(p) = \Phi_{\bar{a}_0} d_z[\tilde{a}_0], \Phi_{\bar{a}_1} d_z[\tilde{a}_1], \cdots, \Phi_{\bar{a}_i} d_z[\tilde{a}_i].$$

Since $x \in Q_k$ and since $p_* \in G(x)$ we have by Proposition 5.10 that

$$d_z[x] \leq \mu\|w_0\|.$$

Hence by Proposition 5.9:

$$\|\psi(p)\| = \|\Phi_{\bar{a}_0} d_z[\tilde{a}_0], \Phi_{\bar{a}_1} d_z[\tilde{a}_1], \cdots, \Phi_{\bar{a}_i} d_z[\tilde{a}_i]\| \leq \|(1,1)\|\|d_z[x]\| \leq \|(1,1)\|\mu\|w_0\|.$$

Let $\{p_1, p_2, \cdots, p_a\} = H(k),\ a = |H(k)|$. Since

$$[k_m, k_n] \cap O_k = \cup_{p\in H(k)}\{[k_m, k_n] \cap O_k(p, \nu(k)(p))\}$$

it follows that

$$\begin{aligned}
\|\Phi_{\bar{k}_m} d_{z_m}[\tilde{k}_m], \Phi_{\bar{k}_{m+1}} d_{z_{m+1}}[\tilde{k}_{m+1}], \cdots, \Phi_{\bar{k}_n} d_{z_n}[\tilde{k}_n]\| &\leq \|\psi(p_1), \psi(p_2), \cdots, \psi(p_a)\| \\
&= \|\|\psi(p_1)\|, \|\psi(p_2)\|, \cdots, \|\psi(p_a)\|\| \\
&\leq \|\underbrace{1, 1, \cdots, 1}_{2|H(k)|}\|\mu\|w_0\| \\
&= 2^{1/r}|H(k)|^{1/r}\mu\|w_0\|
\end{aligned}$$

hence by inequality (5.40) and equation (5.39):

$$\|\beta|_{[m,n]}\| \leq 2^{1/r}|H(k)|^{1/r}\mu\|w_0\| + \chi_\nu(H(k), \nu(k))\|c\|\|\mathscr{T}_{k_{n+1}-1}w_2\|. \tag{5.41}$$

By inequality (5.41) and inequality (5.38) and since

$$|z_m| = \|w_2|_{A_m}\| \leq \|\mathscr{T}_{k_m-1}w_2\|$$

we have

$$\begin{aligned}
\|\mathscr{T}_{k_{n+1}-1}w_2\| &\leq \|\mathscr{T}_{k_m-1}w_2\| + \|w_2|_{[k_m,k_{n+1})}\| \\
&\leq \|\mathscr{T}_{k_m-1}w_2\| + \gamma_3(U(k))|z_m| + \big(\gamma_4(U(k)) + \gamma_5(U(k))\big)\|\beta|_{[m,n]}\| \\
&\leq (1 + \gamma_3(U(k)))\|\mathscr{T}_{k_m-1}w_2\| + \big(\gamma_4(U(k)) + \gamma_5(U(k))\big) \\
&\quad \cdot\Big(2^{1/r}|H(k)|^{1/r}\mu\|w_0\| + \chi_\nu(H(k), \nu(k))\|c\|\|\mathscr{T}_{k_{n+1}-1}w_2\|\Big)
\end{aligned}$$

Since inequality (5.34) holds, we can now rearrange to obtain:

$$\begin{aligned}
\|\mathscr{T}_{k_{n+1}-1}w_2\| &\leq \frac{1}{1 - \|c\|\big(\gamma_4(U(k)) + \gamma_5(U(k))\big)\chi_\nu(H(k), \nu(k))} \\
&\quad \cdot\Big((1 + \gamma_3(U(k)))\|\mathscr{T}_{k_m-1}w_2\| \\
&\qquad + 2^{1/r}\big(\gamma_4(U(k)) + \gamma_5(U(k))\big)|H(k)|^{1/r}\mu\|w_0\|\Big) \\
&\leq \gamma_6(U(k), H(k), \nu(k))\|\mathscr{T}_{k_m-1}w_2\| + \gamma_7(U(k), H(k), \nu(k))\|w_0\|
\end{aligned}$$

as required. □

In Chapter 6 we will establish sufficient conditions that inequality (5.34) can be satisfied by a finite cover (H, ν) for U.

5 Main result

Define the two time intervals $[0, k_*)$ and $[k_*, \infty]$ where the inclusion time $k_* \in \mathbb{N} \cup \infty$ is the time at which the parameter p_*, corresponding to the unknown true plant $P = P_{p_*}$, belongs to the set of available parameters for the first time (see equation (5.42)). Note that in the classical setup (e.g. see French and Trenn (2005); Fisher-Jeffes (2003); Hespanha et al. (2003); Morse (1996, 1997)) we have $p_* \in G(k) = G$, $\forall k \in \mathbb{N}$ so $k_* = 0$.

In Proposition 5.13 we have established gain bounds for sequences of intervals (ongoing intervals) relating to ongoing switches, i.e. $m, n \in \mathbb{N},\ k_i \in O_k,\ 0 \leq m \leq i \leq n$ and in Proposition 5.5 we have established gain bounds which can be applied to intervals (final intervals) relating to final switches, i.e. $k_i \in F_k$. Now observe the following: to every $p \in H(k)$, provided that $Q_k(p, \nu(k)(p))$ is not empty, there exists a plant z in the neighbourhood $B_\chi(p, \nu(k)(p))$, such that the algorithm switches to that plant for the last time on the interval $[0, k]$, i.e. $z = q(F_k(p)),\ z \in B_\chi(p, \nu(k)(p))$. This implies that none, one, or a sequences of ongoing intervals is always followed by a final interval. This progression may repeat itself a maximum of $|H(k)|$ times since there can be only a maximum of $|F_k| = |H(k)|$ final switches.

These facts will be used in the following main result establishing gain bounds on w_2 in terms of w_0 for dynamic and static EMMSAC — as indicated by the gray squares in Figure 5.7. To improve readability we repeat all relevant equations in Table 5.9.

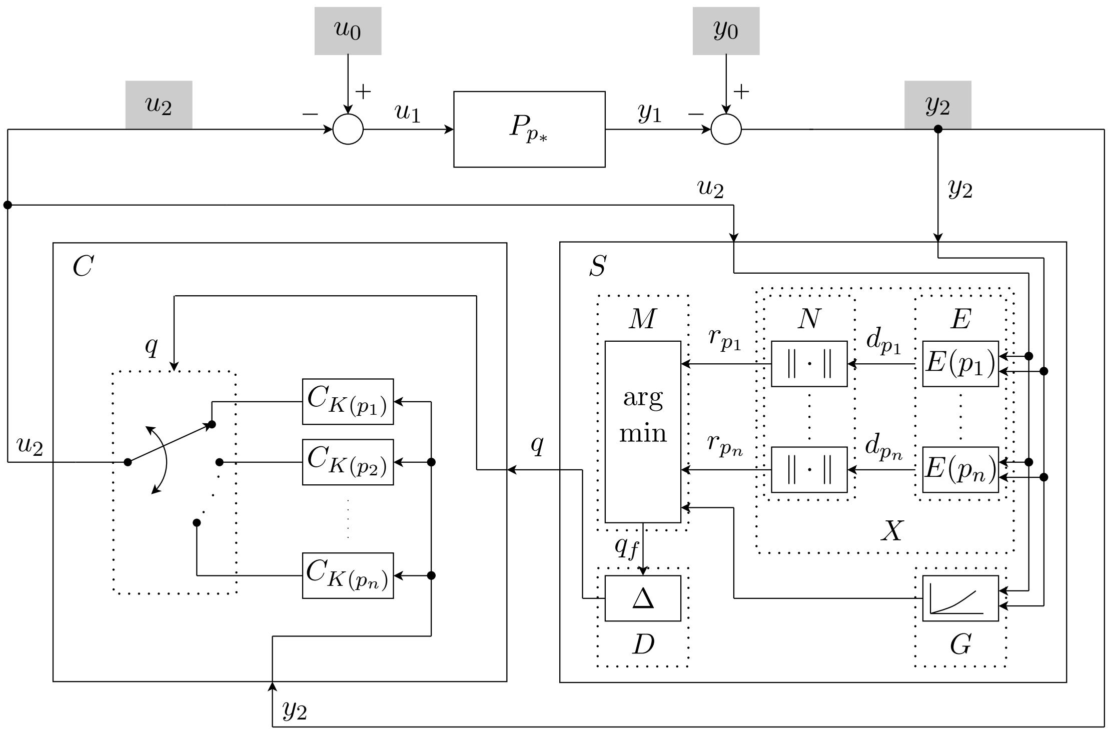

FIGURE 5.7: Bounding $w_2 = (u_2, y_2)^\top$ in terms of $w_0 = (u_0, y_0)^\top$

$$
\begin{aligned}
y_1 &= Pu_1 && (2.7)\\
u_0 &= u_1 + u_2 && (2.8)\\
y_0 &= y_1 + y_2 && (2.9)\\
u_2 &= Cy_2 && (2.10)
\end{aligned}
$$

TABLE 5.9: Details for Theorem 5.14

Theorem 5.14. *Let $1 \le r \le \infty$. Suppose $p_* \in \mathcal{P}^U \subset \mathcal{P}$ where $\mathcal{P}^U$ satisfies Assumption 5.2. Let $P = P_{p_*}$. Let U be a monotonic plant generating operator and suppose (H, ν) defines a monotonic cover for U. Suppose the EMMSAC algorithm is standard. Let $k \in \mathbb{N}$. Let G be a plant generating operator that satisfies $G(j) \subset U(j),\ j \le k$. Suppose $(w_0, w_1, w_2) \in \mathcal{W} \times \mathcal{W}_e \times \mathcal{W}_e$ satisfy the closed loop equations (2.7)–(2.10). Let*

$$
k_* = \begin{cases} \min\{i \in Q_\infty \mid p_* \in G(i)\} & \text{if } \exists i \text{ s.t. } p_* \in G(j),\ \forall j \ge i \\ \infty & \text{if not} \end{cases} \tag{5.42}
$$

and suppose $k_* < \infty$. If

$$
\|c\| \Big(\gamma_4(U(j)) + \gamma_5(U(j))\Big) \chi_\nu(H(j), \nu(j)) < 1,\ \forall j \le k \tag{5.43}
$$

and $\alpha_{OP}(U(k)) < 1$ then:

$$
\|\mathscr{T}_k w_2\| \;\le\; \beta(U(k), H(k), \nu(k), p_*)\|\mathscr{T}_{k_*-1} w_2\| + \hat{\gamma}(U(k), H(k), \nu(k), p_*)\|w_0\|
$$

where for $\mathcal{Q}_1 \subset \mathcal{P}^U,\ \mathcal{Q}_2 \subset \mathcal{P}^H,\ \epsilon : \mathcal{P} \to \mathbb{R}^+$:

$$
\begin{aligned}
\alpha_{OP}(\mathcal{Q}) &= \sup_{p_1 \in \mathcal{Q}} l(p_1)\\
\beta_{OP}(\mathcal{Q}) &= J(\xi) \sup_{p_1 \in \mathcal{Q}} \sup_{\Delta(p_1) \le x \le 2\Delta(p_1)} \beta^\xi(p_1, K(p_1), x - \sigma, \sigma)\\
\alpha_{OS}(\mathcal{Q}) &= J(\xi) \sup_{p_1 \in \mathcal{Q}} \sup_{\Delta(p_1) \le x \le 2\Delta(p_1)} \alpha^\xi(p_1, K(p_1), 0, x - \sigma)\\
\beta_{OS}(\mathcal{Q}) &= J(\xi) \sup_{p_1 \in \mathcal{Q}} \sup_{\Delta(p_1) \le x \le 2\Delta(p_1)} \beta^\xi(p_1, K(p_1), 0, x - \sigma)\\
\gamma_1(p) &= 1 + \sup_{\Delta(p) \le x \le 2\Delta(p)} \alpha(p_*, K(p), 0, x)\\
\gamma_2(p) &= \sup_{\Delta(p) \le x \le 2\Delta(p)} \beta(p_*, K(p), 0, x)\\
\gamma_3(\mathcal{Q}_1) &= \begin{cases} (1 + \alpha_{OS}^{1/r}(\mathcal{Q}_1)) \left(\frac{\alpha_{OP}(\mathcal{Q}_1)}{1 - \alpha_{OP}(\mathcal{Q}_1)}\right)^{1/r} + \alpha_{OS}^{1/r}(\mathcal{Q}_1) & \text{if } \ 1 \le r < \infty \\ \max\{1, \alpha_{OS}(\mathcal{Q}_1)\} \alpha_{OP}(\mathcal{Q}_1) + \alpha_{OS}(\mathcal{Q}_1) & \text{if } \quad r = \infty \end{cases}\\
\gamma_4(\mathcal{Q}_1) &= \begin{cases} (1 + \alpha_{OS}^{1/r}(\mathcal{Q}_1)) \left(\frac{\beta_{OP}(\mathcal{Q}_1)}{1 - \alpha_{OP}(\mathcal{Q}_1)}\right)^{1/r} & \text{if } \ 1 \le r < \infty \\ \max\{1, \alpha_{OS}(\mathcal{Q}_1)\} \frac{\beta_{OP}(\mathcal{Q}_1)}{1 - \alpha_{OP}(\mathcal{Q}_1)} & \text{if } \quad r = \infty \end{cases}
\end{aligned}
$$

$$
\begin{aligned}
\gamma_5(\mathcal{Q}_1) &= \begin{cases} \beta_{OS}^{1/r}(\mathcal{Q}_1) & \text{if} \quad 1 \le r < \infty \\ \beta_{OS}(\mathcal{Q}_1) & \text{if} \quad r = \infty \end{cases} \\
\chi_\nu(\mathcal{Q}_2, \epsilon) &= 2 \sup_{p \in \mathcal{Q}} \epsilon(p) \\
\gamma_6(\mathcal{Q}_1, \mathcal{Q}_2, \epsilon) &= \frac{1 + \gamma_3(\mathcal{Q}_1)}{1 - \|c\| \big(\gamma_4(\mathcal{Q}_1) + \gamma_5(\mathcal{Q}_1)\big) \chi_\nu(\mathcal{Q}_2, \epsilon)} \\
\gamma_7(\mathcal{Q}_1, \mathcal{Q}_2, \epsilon) &= \frac{2^{1/r} \mu |\mathcal{Q}_2|^{1/r} \big(\gamma_4(\mathcal{Q}_1) + \gamma_5(\mathcal{Q}_1)\big)}{1 - \|c\| \big(\gamma_4(\mathcal{Q}_1) + \gamma_5(\mathcal{Q}_1)\big) \chi_\nu(\mathcal{Q}_2, \epsilon)} \\
\beta(\mathcal{Q}_1, \mathcal{Q}_2, \epsilon) &= \gamma_6^{|\mathcal{Q}_2|}(\mathcal{Q}_1, \mathcal{Q}_2, \epsilon) \prod_{p \in \mathcal{Q}_2} \gamma_1(p) \\
\hat{\gamma}(\mathcal{Q}_1, \mathcal{Q}_2, \epsilon, p_*) &= \gamma_6^{|\mathcal{Q}_2|}(\mathcal{Q}_1, \mathcal{Q}_2, \epsilon) \prod_{p \in \mathcal{Q}_2} \gamma_1(p) \left(|\mathcal{Q}_2| \gamma_7(\mathcal{Q}_1, \mathcal{Q}_2, \epsilon, p_*) + \sum_{p \in \mathcal{Q}_2} \gamma_2(p) \right)
\end{aligned}
$$

where c is as in Assumption 3.4(5) and $J(\xi)$ is from Lemma (5.6).

Proof Let $1 \le r \le \infty$. Suppose $0 \le k \le k_* - 1$. Observe that since the gain

$$a = \gamma_3(U(j)) \ge 0, \ j \le k$$

and

$$0 \le b = \|c\| \big(\gamma_4(U(j)) + \gamma_5(U(j))\big) \chi_\nu(H(j), \nu(j)) < 1, \ j \le k$$

by assumption, it follows that

$$\gamma_6(U(j), H(j), \nu(j)) = \frac{1 + a}{1 - b} \ge 1, \ j \le k.$$

Also observe that since

$$\alpha(p_*, K(p), 0, x) \ge 0, \ p \in \mathcal{P}^U, \ x \in \mathbb{N}$$

it follows that

$$\gamma_1(p) = 1 + \sup_{\Delta(p) \le x \le 2\Delta(p)} \alpha(p_*, K(p), 0, x) \ge 1, \ p \in \mathcal{P}^U$$

therefore

$$\beta(U(j), H(j), \nu(j), p_*) = \gamma_6^{|H(k)|}(U(j), H(j), \nu(j), p_*) \prod_{p \in H(j)} \gamma_1(p) \ge 1, \ j \le k$$

and we have

$$\|\mathscr{T}_k w_2\| \ \le \ \|\mathscr{T}_{k_*-1} w_2\| \le \beta(U(j), H(j), \nu(j), p_*) \|\mathscr{T}_{k_*-1} w_2\| + \gamma(U(j), H(j), \nu(j)) \|w_0\|$$

as required.

Now suppose $k \geq k_*$. Let

$$\{k_{f_0} = k_*, k_{f_1}, \cdots, k_{f_m}\} = \cup_{p \in H(k)} \{\max(O_k(p))\} \cup \{k_*\} \cup F_k$$

be an ordered set of switching times, i.e. $k_{f_i} \leq k_{f_{i+1}},\ 0 \leq i < m$.

Observe that the algorithm might not switch to some neighbourhood $B_\chi(p, \nu(k)(p)),\ p \in H(k)$ at all, i.e. there might exist a $p \in H(k)$ such that $F_k(p) = O_k(p) = \emptyset$, and indeed $O_k(p_i) \cap F_k(p_j)$ may not be empty for all $i, j \leq k$ however

$$m = |F_k| + |\cup_{p \in H(k)} \{\max(O_k(p))\}| \leq 2|H(k)|.$$

Let

$$\begin{aligned} a_{f_i} &= \begin{cases} \gamma_6(U(k), H(k), \nu(k)) & \text{if} \quad k_{f_i} \in O_k \\ \gamma_1(q(k_{f_i})) & \text{if} \quad k_{f_i} \in F_k \end{cases} \\ b_{f_i} &= \begin{cases} \gamma_7(U(k), H(k), \nu(k)) & \text{if} \quad k_{f_i} \in O_k \\ \gamma_2(q(k_{f_i})) & \text{if} \quad k_{f_i} \in F_k \end{cases} \end{aligned}$$

where $a_{f_i} \geq 0$ since $\gamma_1, \gamma_6 \geq 1$, as previously. Now define

$$k_{f_{m+1}} = \min\{a > k_{f_m} \mid a \in Q_a\}$$

and observe that $k_{f_m} \leq k < k_{f_{m+1}}$ where $k_{f_i} \in Q_k \subset Q_{k_{f_{m+1}}},\ 0 \leq i \leq m$ and $k_{f_m}, k_{f_{m+1}} \in Q_{k_{f_{m+1}}}$. We then have with

$$z_{f_i} \quad = \quad \|\mathscr{T}_{k_{f_i}-1} w_2\|_r,\ 0 \leq i \leq m+1.$$

and by Propositions 5.5, 5.13 that:

$$z_{f_{i+1}} \quad \leq \quad a_{f_i} z_{f_i} + b_{f_i} \|w_0\|,\ 0 \leq i \leq m.$$

Since $k_{f_0} = k_*$ there follows that $z_{f_0} = \|\mathscr{T}_{k_*-1} w_2\|$ hence we obtain

$$\begin{aligned}
\|\mathscr{T}_k w_2\| &\leq z_{f_{m+1}} \\
&\leq \prod_{i=0}^{m} a_{f_i} z_{f_0} + \left(\prod_{i=1}^{m} a_{f_i} b_{f_0} + \prod_{i=2}^{m} a_{f_i} b_{f_1} + \cdots + \prod_{i=m}^{m} a_{f_i} b_{f_{m-1}} + b_{f_m} \right) \|w_0\| \\
&\leq \prod_{i=0}^{m} a_{f_i} \left(z_{f_0} + \sum_{i=0}^{m} b_{f_i} \|w_0\| \right) \\
&\leq \gamma_6^{|H(k)|}(U(k), H(k), \nu(k)) \prod_{p \in H(k)} \gamma_1(p) \\
&\quad \cdot \left(z_{f_0} + \Big(|H(k)| \gamma_7(U(k), H(k), \nu(k)) + \sum_{p \in H(k)} \gamma_2(p) \Big) \|w_0\| \right) \\
&\leq \beta(U(k), H(k), \nu(k), p_*) \|\mathscr{T}_{k_*-1} w_2\| + \hat{\gamma}(U(k), H(k), \nu(k), p_*) \|w_0\|
\end{aligned}$$

as required. □

We will establish sufficient conditions for the existence of a finite cover (H, ν) satisfying inequality (5.43) in Chapter 6.

Theorem 5.14 establishes gain bounds for the case where $p_* \in G(j),\ j \geq k_*$. In the case where U describes a finite constant set the theorem is directly applicable taking $G = U$, e.g. for the case of an integrator with an unknown sign, i.e. $\mathcal{P}^G = \mathcal{P}^U = \{(1, -1, 1), (1, +1, 1)\}$, a robustness guarantee can be given via Theorem 2.12.

In the case where U describes a continuum, one can still take $G = U$, however the controller may not have a finite dimensional realisation as a continuum of estimators is involved. However, in the next chapter, we will establish results where G represents a finite sampling of the uncertainty set U, and gives rise to a feasible controller for implementation. With appropriate constructions and under mild conditions, it will be shown that such controllers robustly stabilise all plants $p_* \in U$.

Finally, we claim that if the plant model set contains a plant model of the form

$$P_p : x_p(k+1) = A_p x_p(k) + B_p u_1^p(k),\ y_1^p(k) = C_p x(k) + D_p u_1^p(k),\ k \in \mathbb{N} \tag{5.44}$$

where $A_p = -2, B_p = 1, C_p = -\frac{2}{p+1}, D_p = 1,\ p > 0$, the given bounds have the property that they scale unboundedly for $p \to \infty$. This is due to a loss of observability in P_p for large p, i.e. $C_p \to 0$ as $p \to \infty$. Equivalently, the corresponding transfer function

$$P_p : \frac{y_1^p}{u_1^p} = \frac{z + 2\frac{p}{p+1}}{z + 2},$$

tends towards a pole/zero cancellation for increasingly large p.

To show that this loss of observability causes the gain bound to behave badly, we will first show that α in Assumptions 4.1 scales unboundedly for $p \to \infty$. Recall from Assumptions 4.1 that $l_2, l_3, l_4 \in \mathbb{N},\ l_2 \leq l_3 < l_4,\ I_3 = [l_3, l_4)$. It follows from the output equation in (5.44) that

$$x_p(l_2) = C_p^{-1}\big(y_1^p(l_2) - D_p u_1^p(l_2)\big),\ l_2 \in \mathbb{N}.$$

Since $C_p^{-1} \to \infty$ as $p \to \infty$ it follows that for any non-zero w_1^p, $\|x_p(l_2)\| \to \infty$ as $p \to \infty$. Given l_2, l_3, l_4 we have that $\|w_2|_{I_3}\|$ (in equation (4.10)) must have the property that $\|w_2|_{I_3}\| \to \infty$ as $p \to \infty$ since the closed loop signal $w_2|_{I_3}$ is a function of the previous state $x_p(l_2)$. Hence there cannot exist an $M < \infty$ such that

$$\alpha(p, K(p), a, x) < M,\ \forall a, x \text{ as } p \to \infty.$$

To see that this is reflected by the α in Lemma 4.3, note that for P_p as in equation 5.44 the observability matrix O_p in the proof of Lemma 4.3 is given by $O_p = C_p$. Hence $O_p^+ = C_p^{-1} = -\frac{p+1}{2}$ and $Y_p \to \infty$ as $p \to \infty$ and also $\alpha(p, K(p), a, x) \to \infty,\ \forall a, x$ as $p \to \infty$.

Now observe that in order to satisfy the attenuation ineqality 5.16 for increasingly large values of α we have to choose increasingly large delays $\Delta(p)$, i.e. $\Delta(p) \to \infty$ as $p \to \infty$. Since $\alpha(p, K(p), \cdot, a),\ a \in \mathbb{N}$ is an increasing function in a, this implies that if $p \in \mathcal{Q}$ we have for

$$\alpha_{OS}(\mathcal{Q}) = J(\xi) \sup_{p_1 \in \mathcal{Q}} \sup_{\Delta(p_1) \leq x \leq 2\Delta(p_1)} \alpha^{\xi}(p_1, K(p_1), 0, x - \sigma)$$

from Theorem 5.14 that

$$\alpha_{OS}(\mathcal{Q}) \to \infty \text{ as } p \to \infty$$

and hence γ_3, γ_6 as well as $\hat{\gamma}$ and β grow unboundedly as $p \to \infty$. This establishes the claim.

Chapter 6

Design

The outcome of any design process in multiple model switched control must include a (possibly time-varying) plant model set that allows the algorithm to achieve some performance objective. Hence a designer is necessarily confronted with the following design questions:

1. How many plant models are needed?
2. How should they be (geometrically) distributed over the uncertainty set?
3. How can a conservative design be avoided?

At this point we emphasise that even though this thesis has been presented in a different order, the driving questions that lead to the analysis as it is, i.e. the introduction of the plant-generating operator G and the cover (H, ν) for the uncertainty U, have been precisely the ones asked above.

To find answers to these questions is considered to be one of the key outstanding issues in the field of multiple model control. As mentioned in the introduction, the first two questions are for example addressed in Fekri et al. (2006), where the authors ask: "How to divide the initial parameter uncertainty set into N smaller subsets, how large should N be, etc." and then provide an explicit, however sub-optimal, design procedure to find a constant plant model set, based on the atomic closed-loop performance of matching plant and controller pairs. Anderson et al. (2000) make first steps towards a principled construction of a constant plant model set, whereby they construct a cover for the uncertainty set from local robust stability margins of atomic plant and controller pairs. Furthermore in Anderson (2005) similar questions to Fekri et al. (2006) are asked: "How many plants (models) should be chosen, how does one choose a representative set of plants (plant model set), etc.".

The third question for a non-conservative design of the plant model set, however, has not been addressed previously for MMSAC and is considered to be a key contribution of this thesis. For example, Morse (2004) assumes the uncertainty set to be compact and known. This implies conservativeness since the achieved bounds scale with the size of the uncertainty set. However, no discussion of this issue has yet been conducted in the literature although we note that one of the key reason why adaptive control algorithms are employed at all is their potential for non-conservativeness.

This chapter will give explicit, performance-orientated approaches to answering all of these questions for EMMSAC. We start by attaching some further meaning to the objects H, ν and U.

1 Uncertainty, information and complexity

The purpose of the plant-generating operator U is to specify the uncertainty of the true plant $P = P_{p_*}$ in terms of a plant model set. Let U be a constant. If we have complete information about p_* we would let $U = \{p_*\} \subset \mathcal{P}$ hence U is a single plant. Usually we are uncertain about p_*, however we may have enough information to confine p_* to a region in $\mathcal{P}$, i.e. $p_* \in U \subset \mathcal{P}$. For example for $P_a = \frac{1}{s+a}$, $a \in [-a_{max}, a_{max}]$ let $U(k) = [-a_{max}, a_{max}]$, $\forall k \in \mathbb{N}$.

There is also the possibility that there is no information about the 'size' of the uncertainty available, but only on its structure. For example if $P_a = \frac{1}{s+a}$, $a \in (0, \infty)$. This scenario motivates a dynamic U, as discussed later in this chapter. However for now we will confine ourselves to the case where U is constant.

We now employ a suitable measure to quantify the amount of information that is represented by an uncertainty set specified by U. We will denote this quantity the "metric entropy" or the complexity of U. A higher complexity implies less prior information. This concept of interlinking information with complexity is due to Zames (1998), where it is utilised to seek to define the term 'adaptive' in a control context.

For our purpose this connection is important since the complexity of U, as measured by H and ν, determines the gain bound $\hat{\gamma}$ from Theorem 5.14. The purpose of this chapter is to address design, e.g. how to choose a suitable sampling G of U for the actual implementation in order to ensure robust stabilisation of all $p_* \in U$. The resulting gain bound for the implemented algorithm will then depend on the complexity of U rather than the absolute number of the allowable plant models in G. For example in the case of static EMMSAC, we will give conditions for G which guarantee robust stabilisation of all $p_* \in U$, together with a complexity-dependent gain bound.

1.1 Complexity and metric entropy

The following definitions of complexity are interpretations via metric entropy, which is the minimum number of elements that are required to approximate any given subset in a metric space, given an error bound ϵ.

We define the metric entropy

$$\mathcal{C}_E : (\mathcal{P}^*, \mathbb{R}^+) \to \mathbb{N}$$

by

$$\mathcal{C}_E(A, \epsilon) = \{n \in \mathbb{N} \mid n = \min(|h|) \ s.t. \ A \subset \cup_{p \in h} B_{\mathcal{X}}(p, \epsilon), \ h \in \mathcal{P}^*\}$$

where we note that the size of the neighbourhoods $\epsilon \in \mathbb{R}^+$ is required to equal for all $p \in h$.

The Kolmogorov (1956) 'n-width' is the inverse concept, hence returning the size ϵ of the neighbourhoods in terms of the number of neighbourhoods n. Define the n-width

$$\mathcal{C}_N : (\mathcal{P}^*, \mathbb{N}) \to \mathbb{R}^+$$

by

$$\mathcal{C}_N(A, n) = \{a \in \mathbb{R} \mid a = \inf(\epsilon) \ s.t. \ A \subset \cup_{p \in h} B_{\mathcal{X}}(p, \epsilon), \ h \in \mathcal{P}^*, \ |h| = n\} .$$

Note that for a given constant, compact plant-generating operator U the choice of ϵ in $\mathcal{C}_E(U, \epsilon)$ or n in $\mathcal{C}_N(U, n)$ defines a cover (h, ϵ) for U and we could therefore rewrite the gain bound in Theorem 5.14 in terms of the complexity $\mathcal{C}_E$ or $\mathcal{C}_N$ of U by letting $(U, H, \nu) = (U, h, \epsilon)$. Such covers will satisfy inequality (5.43) for sufficiently large choices of n in $\mathcal{C}_E(U, n)$ and sufficiently small choices of ϵ in $\mathcal{C}_N(U, \epsilon)$. We can therefore relate the complexity of the uncertainty set U to performance. However at this stage we have no handle on how to choose n or ϵ. A further minimisation of the gain-bound with respect to either n or ϵ could then be performed. Additionally, these classic definitions are limited since they provide a cover with neighbourhoods of a common size ϵ where a cover in terms of (H, ν) is more flexible and allows for neighbourhoods of different sizes for each $p \in H$.

We now combine the idea of measuring complexity in terms of a cover with the constraint imposed by the gain bound $\hat{\gamma}$ in Theorem 5.14.

Let $A \subset \mathcal{P}$ be compact and assume that $k_* = 0$ hence $\beta = 0$. Let $\hat{\gamma}$ be defined as in Theorem 5.14. Now define the smallest achievable gain bound $\hat{\gamma}_{OPT}$ with respect to $\hat{\gamma}$ by

$$\hat{\gamma}_{OPT} : \mathcal{P}^* \to \mathbb{R}^+$$

and

$$\hat{\gamma}_{OPT}(A) = \left\{ \gamma^* \in \mathbb{R} \;\middle|\; \begin{array}{c} \gamma^* = \inf(\hat{\gamma}(A, h, \epsilon, p_*)),\ p_* \in A \\ s.t.\ A \subset \cup_{p \in h} B_\chi(p, \epsilon(p)),\ h \in \mathcal{P}^*,\ \epsilon : \mathcal{P} \to \mathbb{R}^+ \\ \text{and } (U, H, \nu) = (A, h, \epsilon) \text{ satisfy (5.43)} \end{array} \right\}. \quad (6.1)$$

So as before, for a constant, compact plant-generating operator U, $\hat{\gamma}_{OPT}(U)$ defines a (in general non-unique) cover (h, ϵ) for U. However, this cover is minimal in respect to $\hat{\gamma}$. This makes the design problem of constructing a cover explicit. $\hat{\gamma}_{OPT}$ will critically depend on the behaviour of the gain bound $\hat{\gamma}$. For example assume that for some algorithm, $\hat{\gamma}(U, h, \epsilon, p_*)$ does not depend on the number of elements in h, but only on the size of ϵ. This algorithm will for an uncertainty set U with many plants (or a continuum) achieve a lower $\hat{\gamma}_{OPT}$ than an algorithm where $\hat{\gamma}(U, h, \epsilon, p_*)$ scales with $|A|$. For a general, time-varying plant-generating operator U, covers can be constructed by evaluating $\hat{\gamma}_{OPT}(U(k))$ for all $k \in \mathbb{N}$.

Let $A \subset \mathcal{P}$ be a compact plant model set and (h, ϵ) provide a cover for A. The complexity of A, as evaluated by (h, ϵ), is therefore given by $C_C(A) := |h|$. For example if the cover $(h_{OPT}, \epsilon_{OPT})$ is minimising $\hat{\gamma}_{OPT}(A)$ then the complexity of A, as measured by the minimising cover $(h_{OPT}, \epsilon_{OPT})$, is given by $C_C(A) = |h_{OPT}|$. The given gain bounds are therefore implicitly functions of complexity.

In general it is not possible to solve the optimisation problem in equation (6.1) and to determine the cover (h, ϵ) that minimises $\hat{\gamma}$ for $A = U$ explicitly. This, however, is not necessary since we may utilise any suitable (possibly non-minimal) cover for U in practice. With the true gain γ we therefore arrive at $\gamma \leq \hat{\gamma}_{OPT} \leq \hat{\gamma}$.

2 Scaling

The overall objective of this chapter is to indicate performance-orientated design strategies for the plant generating operator G. Since no measure of the actual performance γ is available, we seek to optimise upper bounds on γ, e.g. $\hat{\gamma}_{OPT}$ or $\hat{\gamma}$, with respect to the plant-generating operator G instead. For the sake of argument, this section will introduce two key 'scaling geometries' of G and investigate the behaviour of the bounds in their respect. In particular these scalings describe either a refinement or an expansion in the parameter space of G.

Consider the following example. Let U be a constant plant-generating operator defined by

$$U = [(-l, 1, 1), (+l, 1, 1)] \subset \bar{\mathcal{P}}_{LTI} \subset \mathbb{R}^3,\ l > 0.$$

A water tank could have such an uncertainty set, where U describes the uncertainty of the flow rate in or out of the tank.

A possible sampling of U is then given as follows. Let the parameter bound $l > 0,\ l \in \mathbb{R}$ and the parameter step $m > 0,\ m \in \mathbb{R}$ define the plant model set

$$\mathcal{P}_{l,m} = \left\{ (i,1,1) \in \bar{\mathcal{P}}_{LTI} \subset \mathbb{R}^3 |\ i = \pm am,\ a \in \mathbb{N},\ |i| \leq l \right\}. \tag{6.2}$$

All elements in $\mathcal{P}_{l,m}$ are therefore bounded by l, and m apart where we observe that the constant plant generating operator $G = \mathcal{P}_{l,m}$ is a subset of U for all $m > 0$.

We are now interested in how the algorithm behaves when the number of plant models under consideration is large, e.g. the number of elements of $P_{l,m}$ is large. In particular consider the two cases depicted in Figure 6.1 where 1. l is constant and m is increasingly small, and 2. m is constant and l is increasingly large. In a geometrical sense we will observe in the first case a 'refinement' in parameter space and an 'expansion' in the latter. These geometries are motivated by the following observations:

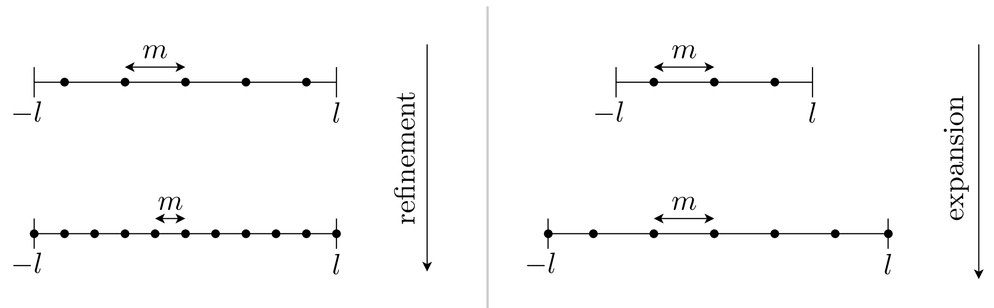

FIGURE 6.1: Increasing the number of elements in $\mathcal{P}_{l,m}$ by scaling

- Assume that computational resource is not an issue and we can implement as many plant models (with corresponding estimators) as we like. We might now be interested in choosing a very fine grid of plant models, corresponding to a small m, in the hope that making plant models available to the algorithm that are very close to the true plant improves performance. This case is concerned with the refinement geometry.
- Assume that the amount of available computational resource is limited. We then might ask the question: What level of refinement $m > 0$ is required to achieve a given performance objective? This case is also concerned with the refinement geometry.
- Assume that for the true plant $P = P_{p*} = P_{(i,1,1)}$ the parameter i is poorly known, i.e. we only know that $|i| < i_{max}$ where i_{max} is large. We are then concerned with providing a stabilising plant model set for all possible values of i. Therefore the plant model has to be expanded for increasingly large values of i_{max}. This case is concerned with the expansion geometry.

We now study the effects of the given scaling scenarios on the algorithm where we utilise different fixed plant model sets $\mathcal{P}_{l,m}$ for the argument. Let $\mathcal{U} = \mathcal{Y} = l_r,\ 1 \leq r \leq \infty$:

- How does the established gain behave with respect to the refinement geometry? E.g. for $G = \mathcal{P}_{l,m}$ with $l \in \mathbb{R}$ and $m \to 0$.
- How does the established gain behave with respect to the expansion geometry? E.g. for $G = \mathcal{P}_{l,m}$ with $m \in \mathbb{N}$ and $l \to \infty$.

The answers to these questions will heavily influence the design of the plant model set.

3 Refinement scaling

Observe that all previously established EMMSAC gain bounds in the literature scale with the number of elements in the plant set (e.g. see French and Trenn (2005)). This is analogous to choosing $p_* \in G = H$ and $\nu = 0$ in Theorem 5.14. Then $\hat{\gamma}$ will also scale exponentially with the number of elements in G. In Hespanha et al. (2001), which is concerned with an observer based multiple model switched adaptive algorithm in the style of Morse, the established bound on the size of the state as well as the robustness margin also scale with the number of elements in the plant model set. The authors then propose a modification to the switching logic to circumvent this analytic issue.

However, is there reason to believe that the actual closed loop gain of multiple model switched adaptive control algorithms is not well behaved in respect to refinement scaling?

Consider this: let $l \in \mathbb{R}$ and consider an arbitrary plant $P_p,\ p \in \mathcal{P}_{l,m}$. If $m > 0$ becomes small, an increasing number of plants will accumulate in neighbourhoods of P_p. However, since all plants in small neighbourhoods of P_p are naturally ‘close’ to P_p, we could attempt to model them as a single plant P_p with a small (time-varying) perturbation. Therefore, if we specify a finite number of neighbourhoods covering the whole of $\mathcal{P}_{l,m}$, any plant in $\mathcal{P}_{l,m}$ can be modeled by perturbations to central cover elements for an arbitrarily small $m > 0$ — as depicted in Figure 6.2.

We have already introduced a suitable device to formally express this intuitive idea for EMMSAC. Observe that $B_{\mathcal{X}}(p, \nu(p))$ specifies a single neighbourhood with radius $\nu(p)$ around the plant $p \in H$. H then specifies the centre of all neighbourhoods that cover U, hence we say that (H, ν) provides a cover for U. Since the gain bound in Theorem 5.14 holds for any $p_* \in G \subset U$, where U can be a continuum, we are potentially allowed to use an arbitrary number of plant models within G. However, observe that the bound of Theorem 5.14 scales with the number of elements in H, where (H, ν) is required to

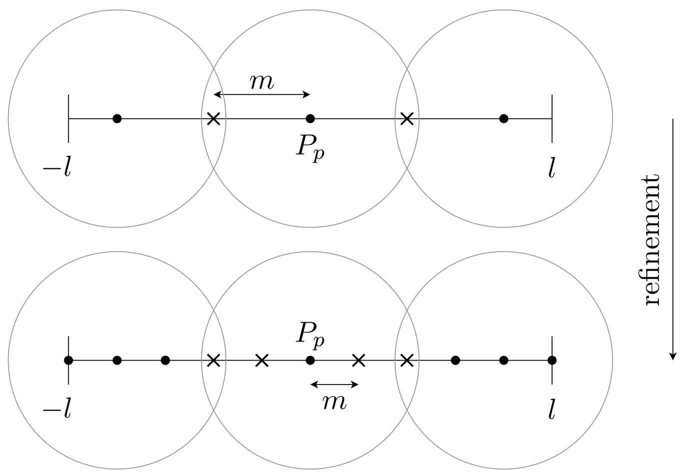

FIGURE 6.2: Covering $P_{l,m}$ by neighbourhoods: The plants labelled $\times$ are modelled as perturbations of the central plant P_p

satisfy inequality (5.43), i.e.

$$2\sup_{p\in H}\nu(p) = \chi_\nu(H,\nu) < \frac{1}{\|c\|\Big(\gamma_4(U)+\gamma_5(U)\Big)}.$$

This inequality implies a constraint on the size of the neighbourhoods $\nu(p),\ p\in H$ (an upper bound), where c depends on the estimator and γ_4,γ_5 depend on the order and the stabilising effect of all atomic closed loops $[P_p,\ C_{K(p)}],\ p\in\cup_{p\in H}B_\chi(p,\nu(p))$, given the controller design procedure K. The fact that the allowable size of the neighbourhoods is a function of the uncertainty set specified by U poses the question if to a compact plant operating operator U, there always exists a finite cover (H,ν), hence a finite $\hat{\gamma}$ in Theorem 5.14.

That this is indeed the case is shown next:

Definition 6.1. *Let $\sigma\in\mathbb{N}$. Let U be a plant-generating operator. Let α,β be defined by Assumptions 4.1 and $\Delta:\mathcal{P}\to\mathbb{N}$, $l:\mathcal{P}\to\mathbb{R}^+$ satisfy inequality* (5.16)*. A control design $K:\mathcal{P}\to\mathcal{C}$ is said to be U regular if for all $\Delta(p)\le x\le 2\Delta(p)$, the functions $l(p)$, $\beta(p_1,K(p),x-\sigma,\sigma)$, $\alpha(p_1,K(p),0,x-\sigma)$, $\beta(p_1,K(p),0,x-\sigma)$, $x\in\mathbb{N}$ are continuous with respect to all $p_1,p\in\mathcal{P}^U$.*

Proposition 6.2. *Let U be a compact plant-generating operator and suppose K is U regular. Suppose $\chi|_{\mathcal{P}^U}$ is continuous. Then there exists a finite cover (H,ν) of U which satisfies inequality* (5.43) *.*

Proof Let $j\in\mathbb{N}$. Since U is compact and K is U regular, the suprema

$$\alpha_{OP}(\mathcal{Q}) = \sup_{p_1\in\mathcal{Q}} l(p_1),\ \ \mathcal{Q}\subset\mathcal{P}^U$$

exists and $\alpha_{OP}(U(j)) < 1$. Also $\alpha_{OS}(U(j)) < \infty$ and $\beta_{OP}(U(j)) < \infty$. Therefore there exist $\epsilon_j > 0$ such that

$$\epsilon_j < \frac{1}{2\|c\|\Big(\gamma_4(U(j)) + \gamma_5(U(j))\Big)}.$$

Recall from Chapter 5, equation 5.3 that

$$B_\chi(p, \epsilon_j) = \{p\} \cup \{p_1 \in \mathcal{P} \mid \chi(p, p_1) < \epsilon_j\} \cap U(j), \ p \in \mathcal{P}.$$

Since $\chi|_{\mathcal{P}^U}$ is continuous, $B_\chi(p, \epsilon_j)$ is open and hence $\{B_\chi(p, \epsilon_j)\}_{p \in U(j)}$ is an open cover of $U(j)$ with respect to the subspace topology of $U(j)$. Since $U(j)$ is compact, there exists a finite set $h_j \subset U(j)$ such that $\{B_\chi(p, \epsilon_j)\}_{p \in h_j}$ covers $U(j)$.

Let $\nu_j(p) = \epsilon_j, \ \forall p \in \mathcal{P}$ hence $(h_j, \nu_j) \in (\mathcal{P}^*, map(\mathcal{P}, \mathbb{R}^+))$ is a finite cover of $U(j)$. Since ν_j is constant it follows that $\epsilon_j = \frac{1}{2}\chi_{\nu_j}(h_j, \nu_j)$. Hence

$$\frac{1}{2}\chi_{\nu_j}(h_j, \nu_j) = \epsilon_j < \frac{1}{2\|c\|\Big(\gamma_4(U(j)) + \gamma_5(U(j))\Big)}.$$

We can therefore construct a monotonic cover (H, ν) by letting $H(k) = \cup_{j \le k} h_j$, $\nu(k)(p) = \min_{j \le k} \epsilon_j, \ \forall p \in \mathcal{P}^H$. It is straightforward to verify that (H, ν) satisfies inequality (5.43) as required. □

We will now show for an example that the existence of a finite cover allows the construction of EMMSAC gain bounds that are invariant to refinement scaling.

3.1 Example

Let the true plant be given by $P_{p_*}, \ p_* = (0, 1, 1)$ where the constant plant-generating operator U specifying the uncertainty set, is given by $U = [(-l, 1, 1), (+l, 1, 1)], \ l > 0$. Apply Proposition 6.2 to give a finite, constant cover (H, ν) for U. Let the constant plant-generating operator G be given by equation (6.2), i.e. $G = \mathcal{P}_{l,m}$, and suppose that $p_* \in G$. Observe that $\mathcal{P}_{l,m}$ describes a sampling of U and therefore $G \subset U$.

Let the plant models $P_p : \mathcal{U}_e \to \mathcal{Y}_e, \ p \in \mathcal{P}_{l,m}$ be given by

$$P_{(a,b,c)} : x_p(k+1) = a x_p(k) + b u_1^p(k), \ y_1^p(k) = c x_p(k), \ x_p(-k) = 0, \ \forall k \in \mathbb{N}. \quad (6.3)$$

Let the controller design procedure K corresponding to the plant P_p, as defined in equation (6.3), be such that $C_c : \mathcal{Y}_2 \to \mathcal{U}_e$ satisfies:

$$C_{K(p)} : u_2(k) = -i y_2(k), \ \forall p = (i, 1, 1). \quad (6.4)$$

Observe that for all $p \in \mathcal{P}_{l,m}$, $[P_p, C_{K(p)}]$ is gain stable. $C_{K(p)}$ is a so-called dead-beat controller since it has the property that if applied to P_p it will bring the plant output y_1^p

to zero after one time step, assuming zero disturbances, i.e. if $q(k) = p$ and $[P_p, C_{K(q(k))}]$ then $y_1^p(k+1) = 0$ assuming $(u_0, y_0)^\top = 0$.

Let the switching controller be defined by

$$C[\mathcal{P}_{l,m}] : \mathcal{Y}_e \to \mathcal{U}_e : y_2 \mapsto u_2 \tag{6.5}$$

with

$$u_2(k) = (C[\mathcal{P}_{l,m}]y_2)(k) := (C_{K(q(k))}y_2)(k),\ k \in \mathbb{N} \tag{6.6}$$

where $q(k) = S(w_2)(k) = DM(X,G)(w_2)(k)$ is given by equations (3.7),(4.25)–(4.33) and $\Delta = 1$. Observe that equations (4.34),(4.35) reduce to equations (6.5),(6.6) for the special case where all plant models and controllers are dead-beat (stabilisable) since we do not have to consider an initialisation at switching times and can simplify.

Now since $p_* \in G$, Theorem 5.14 applies where $k_* = 0$, $\beta = 0$ and $\hat{\gamma} = \hat{\gamma}(U, H, \nu, p_*) < \infty$. Most importantly, since $\hat{\gamma}$ is invariant to G, the bound can be achieved for any refinement level $m > 0$.

This has the following important implication: if we are not limited by implementation considerations we can arbitrarily increase the number of plant models in $\mathcal{P}_{l,m}$, whilst maintaining a common gain bound $\hat{\gamma}$. However note that this does not necessarily mean that the actual closed loop gain $\gamma = \|\Pi_{P_{p*}//C[\mathcal{P}_{l,m}]}\|$ is minimised as $m \to 0$ but only that it does not grow unboundedly in the refinement scaling geometry, i.e. $\gamma \leq \hat{\gamma},\ \forall m > 0$.

4 Sampling of the uncertainty set

Up to this point we assumed that p_* is in the plant model set G or that there exists a time $k_* \in \mathbb{N}$ such that p_* is in $G(k_*)$. For any $p_* \in U \subset \mathcal{P}$ this can only be ensured by the choice $G = U \subset \mathcal{P}$ and hence G may possibly describe a continuum in $\mathcal{P}$. In order for the EMMSAC design to be feasible, we would have to construct estimators that can provide residuals for a continuum of plant models and are bounded in computational complexity. Note that the estimator constructions (estimator A and B) in Chapter 3 are not suitable for a direct implementation of such plant model sets. Hence for the purpose of this thesis we only consider an EMMSAC design to be feasible if G is finite:

Definition 6.3. *An EMMSAC controller is said to be feasible if the underlying plant-generating operator G is finite.*

The construction of estimators that are able to deal with a continuum of plant models goes beyond the scope of this thesis, however note that common plant model sets provide a lot of exploitable structure. For example Morse (1996, 1997, 2004) utilises a state shared observer to allow plant model sets that form a continuum.

In the next section we will show how this implementation issue can be overcome and how a finite plant model set G may be constructed by sampling (possibly continuous) uncertainty sets U. The results will establish conditions under which feasible EMMSAC controllers robustly stabilise any $p_* \in U$, and further that these designs are invariant to refinement scaling.

4.1 Sampling of a constant uncertainty set U

Consider a bound $\bar{\gamma}$ on the closed loop gain that holds for all $p_* \in G$ and has the property that it scales with the number of elements in G. All previously established EMMSAC gain bounds have this property, e.g. see French and Trenn (2005), which is equivalent to taking $U = G = H$, $\nu = 0$ and gives $\bar{\gamma} = \sup_{p \in \mathcal{P}^G} \hat{\gamma}(G, G, 0, p)$ in Theorem 5.14 (on the appropriate class of systems). Notwithstanding their scaling behaviour, such gain bounds lead with Theorem 2.12 to a global robust stability margin $b_{P,C} = \bar{\gamma}^{-1}$. Note that we utilise the bound $\bar{\gamma} \geq \hat{\gamma}$ for design, since $\hat{\gamma}(G, G, 0, p_*)$ depends on the true plant P_{p_*}, $p_* \in G$, which is unknown.

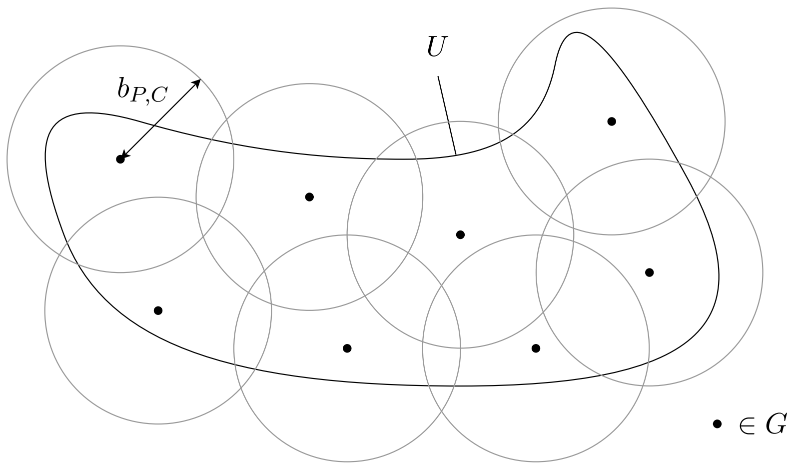

FIGURE 6.3: Covering U by neighbourhoods of size $b_{P,C}$ around $p \in G$

Given some constant, compact plant-generating operator U, we now would like to construct G such that the robustness margins $b_{P,C} = \bar{\gamma}^{-1}$ around each $p \in G$ combine to a cover[1] for U, i.e. $U \subset \cup_{p \in G} \cup_{\delta(p,p_1) < b_{P,C}} \{p_1\}$. If such a finite G can be constructed then a feasible EMMSAC design exists that robustly stabilises all plants in U — as depicted in Figure 6.3.

[1]This cover construction follows essentially the same idea as the local cover construction (H, ν) for U. However, note that ν defines neighbourhoods in the structured uncertainty set U in order to give a notion of complexity of U and to be able to deal with infinitely many plant models, whereas $b_{P,C} = \bar{\gamma}^{-1}$ defines (global) robustness margins (gap balls) in $\mathcal{P}$.

However, since $\bar{\gamma}$ scales with the number of elements in G, and $b_{P,C} = \bar{\gamma}^{-1}$ specifies the size of the robustness margins around each $p \in G$, $b_{P,C}$ may shrink to zero as the number of plant models in G increases and there may not exist a finite cover at all (see Figure 6.4 (A-C)).

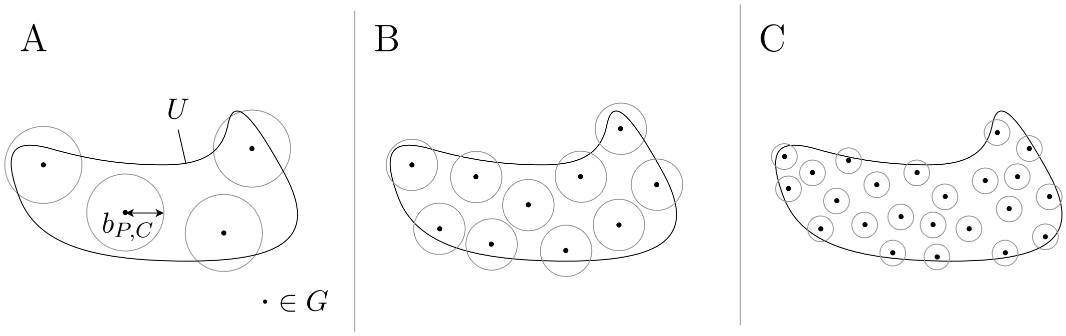

FIGURE 6.4: Attempt to cover U by neighbourhoods $b_{P,C}$, where $b_{P,C}$ scales with $|G|$

Hespanha et al. (2001) essentially face the same problem for their observer-based multiple model switched adaptive algorithm. They note: "[...] if the range of parametric uncertainty is large [...] then the amount of unmodeled dynamics that the switching controller can tolerate becomes small, and might not be sufficient to cover the entire family of admissible process models." In order to rectify this issue, the authors then propose a modification to their switching logic.

To establish robustness margins that are well-behaved in the refinement geometry is therefore not just a theoretical exercise, but it is in fact essential to ensure the existence of a feasible design for the uncertainty set U.

Now assume that (H, ν) provides a suitable constant cover for U which satisfies inequality (5.43). Furthermore assume initially that $p_* \in G \subset U$. We then have by Theorem 5.14 that $k_* = 0$, $\beta = 0$ and $\bar{\gamma} = \sup_{p \in \mathcal{P}^U} \hat{\gamma}(U, H, \nu, p) < \infty$. Since this $\bar{\gamma}$ is well behaved in the refinement geometry and invariant to $G \subset U$, a feasible EMMSAC design exists. This brings us to our next main result:

Theorem 6.4. *Let U be a constant, compact plant-generating operator and suppose $p_* \in \mathcal{P}^U$. Suppose the controller design procedure $K : \mathcal{P} \to \mathcal{C}$ is U-regular. Assume the EMMSAC algorithm is standard where (H, ν) is a constant cover for U which satisfies inequality* (5.43)*. Let $\hat{\gamma}$ be as in Theorem 5.14. Then there exists a constant, finite plant generating operator G satisfying $\mathcal{P}^G \subset \mathcal{P}^U$ and $\bar{\gamma} d < 1$, where*

$$\begin{aligned} \bar{\gamma} &= \sup_{p \in \mathcal{P}^U} \hat{\gamma}(U, H, \nu, p), \\ d &= \sup_{p_2 \in \mathcal{P}^U} \inf_{p_1 \in \mathcal{P}^G} \vec{\delta}(p_1, p_2) \end{aligned}$$

where the standard EMMSAC design based on K and G stabilises all $P = P_{p_*}$ and

$$\|\Pi_{P//C}\| \leq \bar{\gamma}\frac{1+d}{1-\bar{\gamma}d}. \tag{6.7}$$

Proof Since U is compact there exists a constant, finite plant generating operator G such that $\mathcal{P}^G \subset \mathcal{P}^U$ and such that $\bar{\gamma}d < 1$. Let $p_1 \in \mathcal{P}^G$ be such that $\vec{\delta}(p_*, p_1) < d$. Since $\mathcal{P}^G \subset \mathcal{P}^U$ and U, G are constant it follows that $G(j) \subset U(j)$, $\forall j \in \mathbb{N}$, and hence by Theorem 5.14 that

$$\|\Pi_{P_{p_1}//C}\| \leq \hat{\gamma}(U, H, \nu, p_1) \leq \bar{\gamma} < \infty.$$

Since $\vec{\delta}(p_*, p_1) < d < \bar{\gamma}^{-1} = b_{P,C}$ the result follows from Theorem 2.12 as required. □

It is important to note the following facts:

- A refinement of G is always possible since the bound on the closed loop gain in inequality 6.7 holds for any refinement level d such that $d < \bar{\gamma}^{-1}$. Hence we have a positive answer to our first scaling question regarding refinement in a general setting.
- The bound in inequality 6.7 also holds for any plant $p_* \in \hat{U} \supset U$ where

  $$\hat{U} = \cup_{p_1 \in \mathcal{P}^G}\{p \in \mathcal{P} \mid \vec{\delta}(p_1, p) < d\}.$$

  Since $\vec{\delta}$ describes gap-balls in $\mathcal{P}$, this implies that the EMMSAC algorithm robustly stabilises all $p_* \in U$.
- Recall that Theorem 5.14 allowed plant models of the form

  $$P_p : \frac{y_1^p}{u_1^p} = \frac{z + 2\frac{p}{p+1}}{z+2}, \; p > 0,$$

  although we have shown that then the corresponding gain bounds scale unboundedly for increasingly large p (see Page 143). However in Theorem 6.4 such plant models are excluded by the assumption that U is compact, i.e. in the limit, P_p is not observable hence not contained in $U \subset \mathcal{P}$ which, by Assumption 5.2, is a set of observable plants.

Note that Theorem 5.14 only requires that $p_* \in G \subset U$. So for a constant U, G may be time-varying. This leads to the following result:

Theorem 6.5. *Let U be a constant, compact plant-generating operator and suppose $p_* \in \mathcal{P}^U$. Suppose the controller design procedure $K : \mathcal{P} \to \mathcal{C}$ is U-regular. Assume the EMMSAC algorithm is standard where (H, ν) is a constant cover for U which satisfies* *inequality* (5.43)*. Let* $\hat{\gamma}$ *be as in Theorem 5.14. Then there exists a finite plant generating operator* G *satisfying* $\mathcal{P}^G \subset \mathcal{P}^U$ *and* $\bar{\gamma}d < 1$*, where*

$$\begin{aligned} \bar{\gamma} &= \sup_{p \in \mathcal{P}^U} \hat{\gamma}(U, H, \nu, p), \\ d &= \sup_{k \in \mathbb{N}} \sup_{p_2 \in \mathcal{P}^U} \inf_{p_1 \in G(k)} \vec{\delta}(p_1, p_2) \end{aligned}$$

where the standard EMMSAC design based on K *and* G *stabilises all* $P = P_{p_*}$ *and*

$$\|\Pi_{P//C}\| \leq \bar{\gamma}\frac{1+d}{1-\bar{\gamma}d}.$$

Proof The proof is identical to the one of Theorem 6.4, where we construct all $G(k),\ k \in \mathbb{N}$ sufficiently dense that the robustness margins around $p \in G(k)$, given by $b_{P,C} = \bar{\gamma}^{-1}$, cover U. □

This opens the EMMSAC algorithm up to a large class of on-line refinement schemes, so-called dynamic EMMSAC, as discussed in Chapter 4, Section 3 and later in Section 8 of this chapter.

We conclude this section by observing the following facts, which apply to the setting of compact, constant U:

- If there is an infinite amount of computational resource available we may include as many plant models in $G \subset U$ as we like without weakening the gain bounds from Theorem 6.4 and Theorem 6.5. Furthermore for $G = U$ the bounds are minimised, e.g. G being a continuum, and collapse to the one in Theorem 5.14. However, note that this does not imply that the true gain is minimised for $G = U$ but only that it remains bounded.
- If there is sufficient but finite amount of computational resource available, we can always construct a feasible EMMSAC design.
- If the algorithm does not stabilise a plant $p_* \in U$, the only explanation is that the plant model sets $G(k),\ k \in N$ are not dense enough.

These results only hold if a finite cover exists. Sufficient conditions for such covers are given in Proposition 6.2, which includes the requirement that χ is continuous. See the discussion in Chapter 3, Section 3.

5 Expansion scaling and the cause of conservativeness

We now return to our scaling example (equations (6.2)–(6.6)) for a fixed $m > 0$ and some $l > 0$. Observe that all previously established gain bounds, e.g. the one in Theorem

5.14, scale with the size of the candidate plant set $G = \mathcal{P}_{l,m}$ hence with l. The reason for this behaviour may be that either these bounds are unnecessarily weak or that it is in fact the actual closed loop gain that scales with the size of the uncertainty set and therefore its upper bounds. In the following we will prove that it is indeed the true closed loop gain that behaves badly for large l.

Intuitively this can be explained in the following way: Observe that for a parametric uncertainty of 'level' $l \in \mathbb{R}$ the controller set will have to include controllers which are able to deal with a true plant of the worst case parameter value l. This implies, since in our example l represents a bound on the gain of the true plant, we will have to introduce for increasingly large gains l, controllers with increasingly large controller gain to the controller set in order to provide a stabilising controller. If we now manage to confuse the algorithm by a suitable choice of disturbance and convince it to switch the controller with the highest gain into closed loop, we will experience high closed loop gains. We can therefore potentially show that the closed loop gain scales with l — that it is conservative.

We will now show for a simple example that the static EMMSAC algorithm indeed has this undesirable property. Note that although the argument applies to the EMMSAC algorithm, one would expect similar phenomena for other multiple model schemes e.g. designs in the sense of Morse etc.

Theorem 6.6. *Let $m > 0$ and let the parameter set $\mathcal{P}_{l,m}$ be given by equation* (6.2)*. Suppose the EMMSAC algorithm is standard where $\Delta = 1$, $\lambda = 2$ and $G = \mathcal{P}_{l,m}$. Let the atomic plant and controller be defined by equations* (6.3),(6.4)*. Let the switching controller $C[\mathcal{P}_{l,m}]$ be given by equations* (6.5),(6.6)*. Then for $p_* = (1,1,1)$, $P = P_{p_*}$ the closed loop system $[P, C[\mathcal{P}_{l,m}]]$ has the property that there does not exist $M > 0$ such that*

$$\left\|\Pi_{P//C[\mathcal{P}_{l,m}]}\right\| \leq M, \ \forall l \geq 1.$$

Proof Let $m = 1$. Then the set of plants under consideration parametrised by the uncertainty level $l \geq 1$ is given by:

$$\begin{aligned}\mathcal{P}_{l,1} = \mathcal{P}_{l,m} &= \{(-l,1,1),\ldots,(-2,1,1),(-1,1,1),(1,1,1),(-2,1,1),\ldots,(l,1,1)\} \\ &= \{p_l,\ldots,p_4,p_2,p_1 = p_*,p_3,\ldots,p_{l-1}\}\end{aligned}$$

and

$$G(k) = \mathcal{P}_{l,1} = const., \ \forall k \in \mathbb{N}.$$

Observe that with plant and controller being defined by equations (6.3),(6.4) the closed loop $[P_p, C_{K(p)}]$ is gain stable for all $p \in \mathcal{P}_{l,1}$.

The proof is now in two steps. First we will show that we can always make the switching algorithm switch to the controller corresponding to the plant with the largest possible

k	$\begin{pmatrix} u_0 \\ y_0 \end{pmatrix}$	$\begin{pmatrix} u_1 \\ y_1 \end{pmatrix}$	$\begin{pmatrix} u_2 \\ y_2 \end{pmatrix}$
0	$\begin{pmatrix} 0 \\ B \end{pmatrix}$	$\begin{pmatrix} B \\ 0 \end{pmatrix}$	$\begin{pmatrix} -B \\ B \end{pmatrix}$
1	$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} -B \\ B \end{pmatrix}$	$\begin{pmatrix} B \\ -B \end{pmatrix}$
2	$\begin{pmatrix} 0 \\ B - lB \end{pmatrix}$	$\begin{pmatrix} l(B - lB) \\ 0 \end{pmatrix}$	$\begin{pmatrix} l(lB - B) \\ B - lB \end{pmatrix}$
3	$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} \times \\ l(B - lB) \end{pmatrix}$	$\begin{pmatrix} \times \\ \times \end{pmatrix}$

TABLE 6.1: Signals for the true plant $P = P_{p_*}$ up to time $k = 3$

$a \in \mathbb{N},\ (a, 1, 1) \in \mathcal{P}_l$, that is $a = l \in \mathbb{N}$. Second we will show that this condition leads to the unbounded increase in the gain of the closed loop operator as l increases.

Let $p_b = (b, 1, 1), p_l = (l, 1, 1) \in \mathcal{P}_{l,1},\ 1 \leq b < l$. Now consider the closed-loop system $[P, C[\mathcal{P}_{l,1}]]$ and let

$$\begin{pmatrix} u_0 \\ y_0 \end{pmatrix} = \left(\begin{pmatrix} 0 \\ B \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ B - lB \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \dots \right)$$

where $B > 0$.

We now claim that these disturbances make the algorithm switch to the controller C_{p_l} in two time steps, i.e. $q(2) = p_l = q_f(2) = S(w_2)(2)$, and that the signals in Table 6.1 are consistent with

$$\begin{pmatrix} u_1 \\ y_1 \end{pmatrix} = \Pi_{P//C[\mathcal{P}_{l,1}]} \begin{pmatrix} u_0 \\ y_0 \end{pmatrix},\ u_0 = u_1 + u_2,\ y_0 = y_1 + y_2$$

as well as

$$\begin{aligned} P = P_{p_*} : y_1(k+1) &= y_1(k) + u_1(k) \\ P_{p_b} : y_1^b(k+1) &= b y_1^b(k) + u_1^b(k) \\ P_{p_l} : y_1^l(k+1) &= l y_1^l(k) + u_1^l(k) \\ y_1(0) &= y_1^1(0) = y_1^l(0) = 0. \end{aligned}$$

To see this we argue as follows. At time $k = 0$ the disturbance estimates $d_p[0],\ p \in \{p_b, p_l\}$ are forced by $y_0(0) = B$ and zero initial conditions to be

$$d_p[0] = \begin{pmatrix} \times \\ y_0^p(0) \end{pmatrix} = \begin{pmatrix} \times \\ B \end{pmatrix},\ p \in \{p_b, p_l\}.$$

Here and throughout this proof, a vector with an entry marked $\times$ indicates that the entry is irrelevant to the calculation that follows.

Consequently $\|d_p[0]\| = B,\ p \in \{p_b, p_l\}$ and since the switching operator S returns the parametrisation $p_i \in \mathcal{P}_{l,m}$ with the lowest index i if there exist multiple minimal residuals, we have $q(0) = p_* = p_1$.

To order $\mathcal{P}_{l,m}$ in this way is a choice we made earlier, however observe that if for example the order is reversed and the lowest index is assigned such that $p_1 = (l, 1, 1)$ we would have shown the fist step right away. Hence assume the original definition which is most favourable to the algorithm. With $u_2(0) = -y_2(0) = -B$ and $u_0(0) = 0$ we then have $u_1(0) = B$.

At time $k = 1$ we have $y_1(1) = B$ and with $y_0(1) = 0$ there holds $y_2(1) = -B$. The smallest disturbance $d_p[1],\ p \in \{p_b, p_l\}$ consistent with $(\mathscr{T}_0 u_2, \mathscr{T}_1 y_2)$ and P_{p_b}, P_{p_l} can, by the general property $\|d_p[k]\| \leq \|d_p[k+1]\|,\ p \in \mathcal{P},\ k \in \mathbb{N}$ be found to be

$$d_p[1] = \left(\begin{pmatrix} 0 \\ B \end{pmatrix} \begin{pmatrix} \times \\ 0 \end{pmatrix} \right),\ p \in \{p_b, p_l\}.$$

Since $\|d_{p_l}[1]\| = \|d_{p_b}[1]\|$, $q(1) = p_*$ and no switch occurs. Furthermore with $u_2(1) = -y_2(1) = B$ and $u_0(0) = 0$ we have $u_1(1) = -B$.

At time $k = 2$ we have $y_1(2) = 0$ and with $y_0(2) = B - lB$ there holds $y_2(2) = B - lB$. Now, the smallest disturbance estimate for $d_{p_l}[2]$ consistent with $(\mathscr{T}_1 u_2, \mathscr{T}_2 y_2)$ and P_{p_l} is

$$d_{p_l}[2] = \left(\begin{pmatrix} 0 \\ B \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} \begin{pmatrix} \times \\ 0 \end{pmatrix} \right)$$

since similarly minimality is ensured by consistency and $\|d_{p_l}[2]\| = \|d_{p_l}[1]\|$. In fact, the disturbances (u_0, y_0) are not arbitrary but have been chosen so that this holds.
Since $y_0^{p_b}(0) = B$, $\|d_{p_b}[2]\| \geq \|d_{p_l}[2]\|$, however the choice $d_{p_b}[2] = d_{p_l}[2],\ p_b \neq p_l$ is not possible since the trajectories would have the property that

$$\Pi_{C[\mathcal{P}_{l,1}]//P_{p_b}} d_{p_b}[2] = \Pi_{C[\mathcal{P}_{l,1}]//P_{p_b}} d_{p_l}[2] \neq (\mathscr{T}_1 u_2, \mathscr{T}_1 y_2).$$

This can be seen by choosing

$$d_{p_b}[2] = \left(\begin{pmatrix} 0 \\ B \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} \begin{pmatrix} \times \\ y_0^{p_b}(2) \end{pmatrix} \right).$$

In this case we have $y_1^{p_b}(2) = bB - B$. With $y_2^{p_b}(2) = B - lB$ from above we would have to choose

$$y_0^{p_b}(2) = bB - lB \neq 0,\ \forall b \neq l$$

to be consistent with $(\mathscr{T}_1 u_2, \mathscr{T}_2 y_2)$ and P_{p_b}. Therefore we can conclude that

$$\|d_{p_l}[2]\| = B < \|d_{p_b}[2]\|.$$

Consequently we have $q(2) = l$ and obtain $u_2(2) = l(lB - B)$. Furthermore with $u_0(2) = 0$ there follows $u_1(2) = l(B - lB)$.

At time $k = 3$ we are only interested in $y_1(3) = l(B - lB)$, which can be calculated directly. This establishes the first claim.

We now show that this leads to the unbounded increase of the gain of the closed loop operator as l increases. From the definition of $\Pi_{P//C[\mathcal{P}_{l,1}]}$ we have

$$\begin{aligned}\|\Pi_{P//C[\mathcal{P}_{l,1}]}\| &= \sup_{w_0 \in \mathcal{W} \setminus \{0\}} \frac{\|\Pi_{P//C[\mathcal{P}_{l,1}]}(w_0)\|}{\|w_0\|} \\ &\geq \frac{\|w_1\|}{\|w_0\|} \geq \frac{|y_1(3)|}{\|w_0\|} \\ &= \frac{B|(l - l^2)|}{B\|1, 1 - l, 0, 0, \cdots\|}.\end{aligned}$$

Furthermore there exist scalars $L > 1$, $\alpha, \beta > 0$ such that

$$|l - l^2| \geq \alpha l^2, \ \forall l \geq L$$

and

$$\|1, 1 - l, 0, 0, \cdots\|_{l_p} \leq \beta l, \ \forall l \geq L.$$

Therefore with

$$\|\Pi_{P_*//C[\mathcal{P}_{l,1}]}\| \geq \frac{\alpha}{\beta} l, \ \forall l \geq L$$

and the fact that the analysis can be repeated for all $m > 0$, the proof is complete. □

Since the closed loop system is homogeneous, i.e. $\Pi_{P//C}(\alpha w_0) = \alpha \Pi_{P//C}(w_0)$, we have shown that we can make the algorithm switch to an arbitrary controller in the presence of an arbitrarily small disturbance. In order to do so, we exploited the zero initial condition assumption on the system and 'simulated' the output of the plant P_{p_l} by inducing appropriate disturbances. Finally we proved that the algorithm is conservative since the actual closed loop gain has been shown to scale with the uncertainty level $l \geq 1$.

Note that a clever choice of a cover (H, ν) will in this case not provide a finite gain bound since we established a lower bound for the actual closed loop gain. (H, ν) is merely a theoretical device to establish a bound which is invariant to the number of elements in G, e.g. in the refinement scaling scenario.

6 Tackling conservativeness

In this section we will discuss the role of a time-varying plant-generating operator G in dynamic EMMSAC and present a particular construction of G for which we give a gain function bound that is invariant to uncertainty level information of the plant (that is, it is universal). We will achieve this by dynamically expanding the plant model set, which is motivated by the following observation of natural adaptive systems.

When children learn to ride a bike for the first time, they usually approach this rather complex control problem in the following way: They first trial various 'careful' control strategies, i.e. they drive slowly, and fail to control the system since a bike is rather difficult to control at low speeds. Then, they become more and more vigorous until the control strategy is of appropriate aggressiveness (speed) to control the bike with satisfying performance. This strategy is not restricted to riding a bike but proves to be a successful one in 'learning' many physical activities. The approach is known in the literature as the "windsurfer approach" which is due to Lee et al. (1993), where in a different context it is proposed to gradually increase the bandwidth of a controller in order to improve the performance of a closed loop system. Exceptions arise, where we have a priori knowledge on how much vigour is needed, and then approach the problem appropriately from the start.

To replicate this strategy for EMMSAC we have to evaluate the performance of the algorithm given the current plant model set and then, if not sufficient, include more plant models, e.g. resulting in higher gain controllers. There are a number of possibilities for evaluating the current performance of the algorithm. One strong indicator of performance, assuming reasonably small disturbances $\|w_0\|$, is the size of the disturbance estimates. If they are all rather large in size, none of the plant models is very close to the true plant and we can usually expect bad performance. Another more direct and arguably crude performance measure, which we will be using subsequently, is the size of the observation $\|\mathscr{T}_k w_2\|$ at some time $k \in \mathbb{N}$. This choice is based on the observation that if there is no adequate controller in the controller set for the true plant, we expect large closed loop signals and small closed loop signals if controlled sufficiently.

Assume that the uncertainty set, as specified by the plant-generating operator U, is finite. We can therefore let $U = G = H$ and achieve a feasible EMMSAC design. This leads to the following construction of a dynamic EMMSAC algorithm.

Let a plant level set, representing the 'learning level' of the algorithm, be given by

$$\mathcal{P}_i \in \mathcal{P}^*,\ \emptyset \neq \mathcal{P}_j \subset \mathcal{P}_{j+1},\ 1 \leq j < i,\ i \in \mathbb{N} \tag{6.8}$$

where we assume that all $\mathcal{P}_i,\ i \in \mathbb{N}$ are finite and that there exists an index $i \in \mathbb{N}$ such that $p_* \in \mathcal{P}_l,\ \forall l \geq i$.

Let

$$\tilde{\gamma}(\mathcal{Q}) = \max_{p \in \mathcal{Q}} \left(\hat{\gamma}(\mathcal{Q}, \mathcal{Q}, 0, p) + \beta(\mathcal{Q}, \mathcal{Q}, 0, p) \right),\ \mathcal{Q} \subset \mathcal{P}^G$$

where $\hat{\gamma}$ and β are from Theorem 5.14.

Let with $v > 2$ the expansion rule be given by

$$G(k) = \mathcal{P}_{i(k)},\ k \in \mathbb{N} \tag{6.9}$$

where

$$i(k) = \begin{cases} \max\{a \in \mathbb{N} \mid \tilde{\gamma}^v(\mathcal{P}_a) - \tilde{\gamma}^v(\mathcal{P}_1) \leq \|\mathscr{T}_k w_2\|\} & \text{if } 0 \leq k < \infty \\ \infty & \text{if } k = \infty \end{cases}. \tag{6.10}$$

Theorem 5.14 applies with the choice $G(k) = U(k) = H(k),\ \nu = 0$.

This brings us to our next result:

Theorem 6.7. *Let $k \in \mathbb{N}$. Let $\mathcal{P}_i$ be given by equations (6.8) and suppose that there exists $i \in \mathbb{N}$ such that $p_* \in \mathcal{P}_l,\ l \geq i$. Let the expansion rule be given by equation (6.10) which gives the plant-generating operator G via equation (6.9). Suppose the EMMSAC algorithm is standard. Suppose $(w_0, w_1, w_2) \in \mathcal{W} \times \mathcal{W}_e \times \mathcal{W}_e$ satisfy the closed loop equations (2.7)–(2.10). Then for all $w_0 \in \mathcal{W}$:*

$$\|w_2\| \leq \gamma_{mod}(\|w_0\|)$$

where $\gamma_{mod} : \mathbb{R}^+ \to \mathbb{R}^+$ is given by

$$\gamma_{mod}(r) = \beta_1 + \beta_2 r + \beta_3 r^2,$$

with

$$\begin{aligned} \tilde{\gamma}(\mathcal{Q}) &= \max_{p \in \mathcal{Q}} (\hat{\gamma}(\mathcal{Q}, \mathcal{Q}, 0, p) + \beta(\mathcal{Q}, \mathcal{Q}, 0, p)) \\ \beta_1 &= \tilde{\gamma}^{v+2}(\mathcal{P}_N) + \tilde{\gamma}(\mathcal{P}_N)\tilde{\gamma}^v(\mathcal{P}_1) \\ \beta_2 &= 2\tilde{\gamma}^2(\mathcal{P}_N) + \tilde{\gamma}^{1-v}(\mathcal{P}_N)\tilde{\gamma}^v(\mathcal{P}_1) \\ \beta_3 &= \tilde{\gamma}^{2-v}(\mathcal{P}_N) \end{aligned}$$

where β and $\hat{\gamma}$ are from Theorem 5.14 and

$$N := \min\{i \geq 1 \mid p_* \in \mathcal{P}_i\}.$$

Proof Let $w_0 \in \mathcal{W}$ and let k_* be given by equation (5.42). By equation (6.10)

$$\|\mathscr{T}_k w_2\| \leq \tilde{\gamma}^v(\mathcal{P}_{i(k)+1)}) - \tilde{\gamma}^v(\mathcal{P}_1) \leq \tilde{\gamma}^v(\mathcal{P}_{i(k)+1}),\ \forall k \in \mathbb{N}. \tag{6.11}$$

By the fact that

$$i(k_*) \geq N \geq i(k_* - 1) + 1, \tag{6.12}$$

which follows from the definition of k_* and since $\tilde{\gamma}(\mathcal{P}_i)$ is monotonically increasing with i, we can write equation (6.11) with $k = k_* - 1$ as

$$\|\mathscr{T}_{k_*-1}w_2\| \leq \tilde{\gamma}^v(\mathcal{P}_{i(k_*-1)+1}) \leq \tilde{\gamma}^v(\mathcal{P}_N). \tag{6.13}$$

We now have to consider the two possibilities that

1. $k_* = \infty$
2. $k_* < \infty$.

Note that case 1 can occur if no disturbances are acting on the system, i.e. $w_0 = 0$ since it will then rest at the initial condition and $\mathscr{T}_k w_2 = 0$ for all $k \in \mathbb{N}$. We then have by equation (6.10) that no plants can be introduced to G hence there does not exits a $k_* \in \mathbb{N}$ such that $p_* \in G(k_*)$.

We then have that

$$\beta_1 \geq \tilde{\gamma}^{v+2}(\mathcal{P}_N) \geq \tilde{\gamma}^v(\mathcal{P}_N)$$

hence

$$\|w_2\| = \|\mathscr{T}_{k_*-1}w_2\| \leq \tilde{\gamma}^v(\mathcal{P}_N) \leq \beta_1.$$

In case 2 with $k \leq k_* - 1$ it follows similarly to 1. that

$$\|\mathscr{T}_k w_2\| \leq \beta_1.$$

For $k > k_* - 1$ we have by equations (6.10), Theorem 5.14 and inequality (6.13) that

$$\begin{aligned}\tilde{\gamma}^v(\mathcal{P}_{i(k)}) &\leq \|\mathscr{T}_k w_2\| + \tilde{\gamma}^v(\mathcal{P}_1)\\ &\leq \tilde{\gamma}(\mathcal{P}_{i(k)})(\|\mathscr{T}_{k_*-1}w_2\| + \|w_0\|) + \tilde{\gamma}^v(\mathcal{P}_1)\\ &\leq \tilde{\gamma}(\mathcal{P}_{i(k)})(\tilde{\gamma}^v(\mathcal{P}_N) + \|w_0\|) + \tilde{\gamma}^v(\mathcal{P}_1).\end{aligned}$$

Multiplication with $\tilde{\gamma}^{1-v}(\mathcal{P}_{i(k)}) > 0$ yields

$$\tilde{\gamma}(\mathcal{P}_{i(k)}) \leq \tilde{\gamma}^{2-v}(\mathcal{P}_{i(k)})(\|w_0\| + \tilde{\gamma}^v(\mathcal{P}_N)) + \tilde{\gamma}^{1-v}(\mathcal{P}_{i(k)})\tilde{\gamma}^v(\mathcal{P}_1).$$

Furthermore, since $\tilde{\gamma}(\mathcal{P}_i)$ is monotonically increasing with i, we have with equation (6.12) that $\tilde{\gamma}(\mathcal{P}_N) \leq \tilde{\gamma}(\mathcal{P}_{i(k)})$. Hence

$$\tilde{\gamma}^{q-v}(\mathcal{P}_{i(k)}) \leq \tilde{\gamma}^{q-v}(\mathcal{P}_N),\ \forall q < v$$

and we obtain

$$\tilde{\gamma}(\mathcal{P}_{i(k)}) \leq \tilde{\gamma}^2(\mathcal{P}_N) + \tilde{\gamma}^{2-v}(\mathcal{P}_N)\|w_0\| + \tilde{\gamma}^{1-v}(\mathcal{P}_N)\tilde{\gamma}^v(\mathcal{P}_1). \tag{6.14}$$

By Theorem 5.14, inequality (6.14) and inequality (6.13) we now have that:

$$\begin{aligned}
\|\mathscr{T}_k w_2\| &\leq \tilde{\gamma}(\mathcal{P}_{i(k)})(\|\mathscr{T}_{k_*-1} w_2\| + \|w_0\|) \\
&\leq \big(\tilde{\gamma}^2(\mathcal{P}_N) + \tilde{\gamma}^{2-v}(\mathcal{P}_N)\|w_0\| + \tilde{\gamma}^{1-v}(\mathcal{P}_N)\tilde{\gamma}^v(\mathcal{P}_1)\big)(\|\mathscr{T}_{k_*-1} w_2\| + \|w_0\|) \\
&\leq \big(\tilde{\gamma}^2(\mathcal{P}_N) + \tilde{\gamma}^{2-v}(\mathcal{P}_N)\|w_0\| + \tilde{\gamma}^{1-v}(\mathcal{P}_N)\tilde{\gamma}^v(\mathcal{P}_1)\big)(\tilde{\gamma}^v(\mathcal{P}_N) + \|w_0\|) \\
&\leq \tilde{\gamma}^{v+2}(\mathcal{P}_N) + \tilde{\gamma}(\mathcal{P}_N)\tilde{\gamma}^v(\mathcal{P}_1) + (2\tilde{\gamma}^2(\mathcal{P}_N) + \tilde{\gamma}^{1-v}(\mathcal{P}_N)\tilde{\gamma}^v(\mathcal{P}_1))\|w_0\| \\
&\quad + \tilde{\gamma}^{2-v}(\mathcal{P}_N)\|w_0\|^2 \\
&\leq \beta_1 + \beta_2\|w_0\| + \beta_3\|w_0\|^2.
\end{aligned}$$

We observe that the bound is independent of k and therefore

$$\|w_2\| \leq \beta_1 + \beta_2\|w_0\| + \beta_3\|w_0\|^2.$$

as required. □

Now observe that the given dynamic EMMSAC algorithm is universal. This important fact follows directly from Theorem 6.7. The constants $\beta_1, \beta_2, \beta_3$ are invariant to any uncertainty level information and only depend on $\mathcal{P}_i$ and N where N defines the smallest 'learning level' i such that the true plant p_* is included in $\mathcal{P}_N$. Hence we have a positive answer to our second scaling question regarding expansion.

We are now in the position to compare these result for dynamic EMMSAC to the ones obtained in Theorem 5.14 for static EMMSAC, and to the counter example in Theorem 6.6.

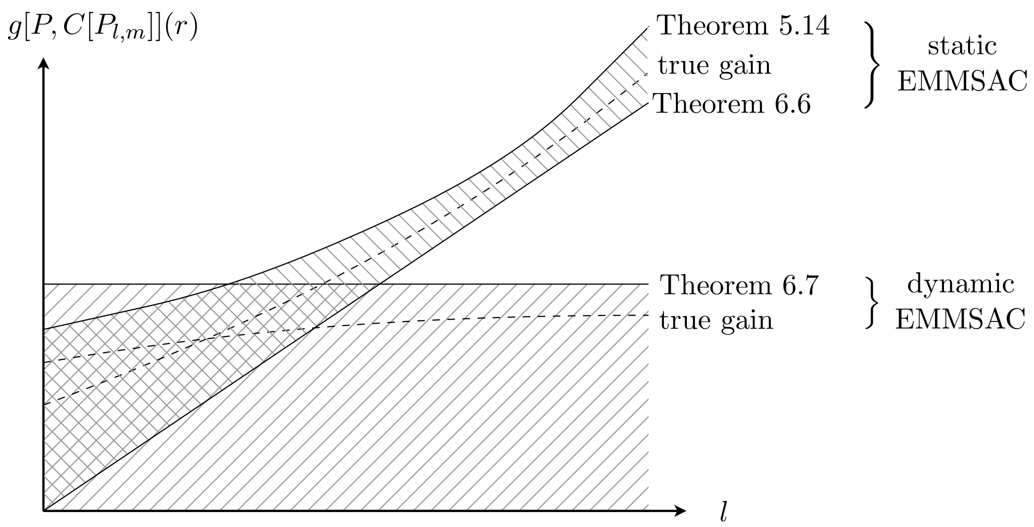

FIGURE 6.5: Gain comparison for EMMSAC under parametric uncertainty of level l

Consider Figure 6.5. In Theorem 6.6 we have discussed how the algorithm behaves in the presence of an increasingly large parametric uncertainty $l \in \mathbb{R}$, represented by the plant model set $G = \mathcal{P}_{l,1}$ with

$$\mathcal{P}_{l,1} = \{(-l,1,1),\dots,(-2,1,1),(-1,1,1),(1,1,1),(-2,1,1),\dots,(l,1,1)\}$$

and we concluded that the actual closed loop gain $\|\Pi_{P_{p*}//C[\mathcal{P}_{l,1}]}\|$ scales at least linearly with the uncertainty level $l \in \mathbb{R}$. This gives a lower bound on the closed loop gain in Figure 6.5 where

$$g[P_{p_*}, C[\mathcal{P}_{l,m}]](r) = \sup_{\|w_0\|<r} \|\Pi_{P_{p*}//C[\mathcal{P}_{l,m}]} w_0\|, \; r \in \mathbb{R}$$

is the worst case gain from the disturbances w_0 to the internal signals w_2 at a disturbance level $r \in \mathbb{R}$, as a function of l. Now observe that an increasingly large l in $G = \mathcal{P}_{l,1}$ corresponds to an increasingly large constant U since $G \subset U$. This however means that the bound $\hat{\gamma}$ in Theorem 5.14 scales with l — as depicted in Figure 6.5.

In contrast we have show in Theorem 6.7 that for a special (dynamic) choice of G we obtain a gain (function) bound which is invariant to l. We therefore conclude that for large parametric uncertainties, dynamic EMMSAC allows for better performance than static EMMSAC.

7 Dynamic versus static EMMSAC

We will now discuss in detail when dynamic EMMSAC promises tighter gain bounds than static EMMSAC and vice versa. First recall that for a constant, compact plant-generating operator U and a corresponding constant cover (H, ν), assuming $p_* \in G \subset U$, there follows $k_* = 0$ hence $\|\mathscr{T}_{k_*-1} w_2\| = 0$. By Theorem 5.14 we then obtain a (linear) gain bound (Figure 6.6 (A)) of the form

$$\|w_2\| \leq \hat{\gamma}(U, H, \nu, p_*)\|w_0\|,$$

where the gain $\hat{\gamma}$ depends on the uncertainty set specified by U and the corresponding cover (H, ν). From Theorem 6.7, we have for a dynamic construction of $U = G = H, \nu = 0$, assuming that there exists a $k_* < \infty$ such that $p_* \in G(k_*)$, a gain function bound of the form

$$\|w_2\| \leq \beta_1 + \beta_2\|w_0\| + \beta_3\|w_0\|^2$$

where $\beta_1, \beta_2, \beta_3$ are constant and depend on $v > 2$, the design of the level set $\mathcal{P}_i$ and the true plant $P = P_{p_*}$ (Figure 6.6 (B)). Since our goal is to optimise the bound on the signal amplification from the disturbances $\|w_0\|$ to the internal signals $\|w_2\|$, we can now intersect these two curves and argue by Figure 6.6 (C) that for disturbances $\|w_0\| < a$, $\|w_0\| > b$ the gain bound obtained for static EMMSAC is tighter than the gain bound for dynamic EMMSAC where for $a < \|w_0\| < b$ the converse relation holds.

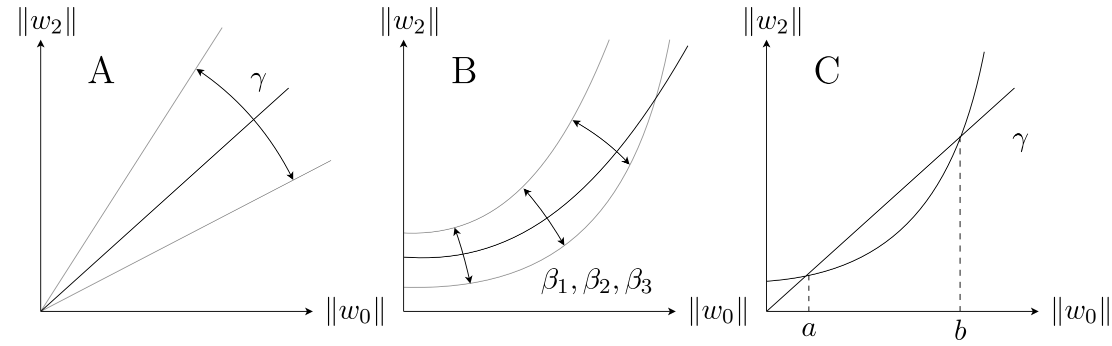

FIGURE 6.6: Gain bound comparison of static and dynamic EMMSAC

Usually there is no exact information on the size of the disturbances available. However in many cases we will have a rough idea how the size of the disturbances is distributed, e.g. small disturbances might be more likely than larger ones, and we can now use this information to trade off the two approaches. Also note that the intersection points a, b depend on $\hat{\gamma}$ and $\beta_1, \beta_2, \beta_3$ where in some scenarios they do not intersect at all, i.e. for $\hat{\gamma} < \beta_2$, and a constant plant set should be preferred over a time-varying one. In all other cases the two curves will intersect for sufficiently large $\|w_0\|$ since the (quadratic) gain function grows faster then the (linear) gain. This implies that for high noise environments, i.e. where large disturbances are very likely, a constant plant model set should be preferred over a time-varying one.

The defining entities of the gain function bound that we are able to influence are therefore:

1. The constant $v > 2$ in

$$\begin{aligned}\beta_1 &= \tilde{\gamma}^{v+2}(\mathcal{P}_N) + \tilde{\gamma}(\mathcal{P}_N)\tilde{\gamma}^{v}(\mathcal{P}_1)\\ \beta_2 &= 2\tilde{\gamma}^{2}(\mathcal{P}_N) + \tilde{\gamma}^{1-v}(\mathcal{P}_N)\tilde{\gamma}^{v}(\mathcal{P}_1)\\ \beta_3 &= \tilde{\gamma}^{2-v}(\mathcal{P}_N).\end{aligned}$$

2. The design of G (and U, H, ν) for dynamic EMMSAC.

For 1. observe that increasingly large v will effectively straighten the curve since β_3 will become increasingly small and the influence of the quadratic term is diminished. However the offset β_1 will increase. Alternatively, small v will lead to small offsets and a faster quadratic growth. The choice of $v > 2$ is therefore dominated by the available information on the size of $\|w_0\|$, i.e. if $\|w_0\|$ is expected to be large it is advantageous to choose v large since then the gain function curve is more linear, which leads to smaller signal amplification. However if $\|w_0\|$ is expected to be small, v should be small since we have to compete with the zero offset of the gain bound for a constant G.

In the next section we give exemplar designs for (time-varying) plant model sets G.

8 Example

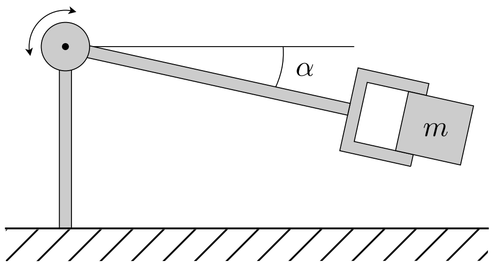

FIGURE 6.7: Robotic arm handling uncertain loads

Consider the example in Figure 6.7. Assume that a robotic arm with 1 degree of freedom is handling items of weight $m \in \{m_1, m_2, m_3\}$, $m_1 \leq m_2 \leq m_3$, where the control objective is to keep the arm perpendicular to the base, i.e. to keep α small. Furthermore assume the robotic arm to be of neglectable mass and that due to the manufacturing process there is a tolerance on the weight of $t\%$. Let the parametrised uncertainty set be given by

$$\bigcup_{0 \leq j \leq t} \bigcup_{m \in \{m_1, m_2, m_3\}} \{(1 \pm j)m\},\ j \in \mathbb{R}$$

where the true, unknown plant is given by $p_* = 1.04m_1$. Assume that the design objective is to stabilise any true plant p_* in the uncertainty set, where the implementation of plant models is computationally expensive and we prefer a small number of them. The control algorithm is reset before every pickup.

8.1 Static EMMSAC

Assume that the tolerance t is known and finite. Therefore let the plant generating operator U be constant and defined by

$$\mathcal{P}^U = \bigcup_{0 \leq j \leq t} \bigcup_{m \in \{m_1, m_2, m_3\}} \{(1 \pm j)m\},\ j \in \mathbb{R}.$$

Also let the cover (H, ν) for U be constant. The following designs of plant model sets are constant therefore G is a constant plant-generating operator.

Consider Figure 6.8 and a constant plant generation operator $\hat{U}$ specifying the uncertainty set $\mathcal{P}^{\hat{U}} = [m_1(1-t), m_3(1+t)] \supset \mathcal{P}^U$. Construct a cover $(\hat{H}, \hat{\nu})$ for $\hat{U}$, satisfying inequality (5.43), and compute $\bar{\gamma} = \sup_{p \in \mathcal{P}^{\hat{U}}} \hat{\gamma}(\hat{U}, \hat{H}, \hat{\nu}, p)$ where $\hat{\gamma}$ is as in Theorem 5.14. Let $G = G_1$ be such that it spans a grid over the uncertainty set $\hat{U}$ and make G dense enough that the global robustness margins $b_{P,C} = \bar{\gamma}^{-1}$ overlap (Figure 6.8 with $G = G_1$).

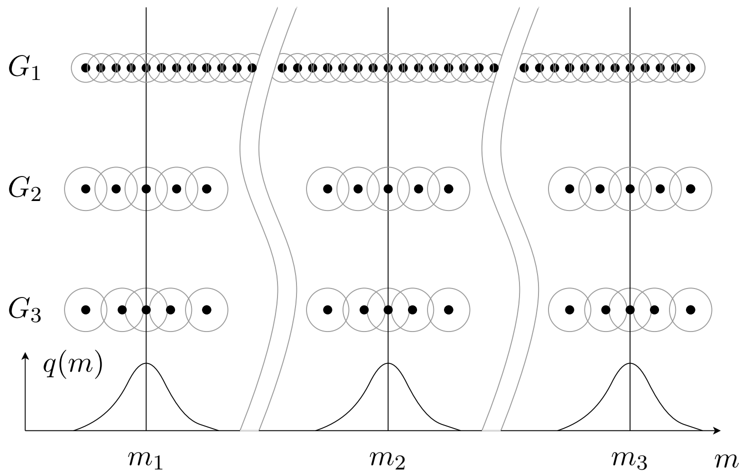

FIGURE 6.8: Sensible choices for G; G_3 in respect to the probability distribution $q(m)$

Observe that although this construction of G provides stability for all $p_* \in U$, it is conservative since by considering an overly large uncertainty set $\hat{U} \supset U$ we 1. provide stability for plants that are not in U, 2. reduce the (global) robust stability margin since $\hat{\gamma}(U, H, \nu, p) \leq \hat{\gamma}(\hat{U}, \hat{H}, \hat{\nu}, p),\ p \in \mathcal{P}^U$ and 3. introduce unnecessary computational complexity since G is overly large. It is therefore important to be as precise as possible about the uncertainty specification.

The obvious improvement is to work with the uncertainty set U directly. We therefore construct a cover (H, ν) for U, satisfying inequality (5.43), compute $\bar{\gamma}$ and then construct a reasonably sparse G such that the (global) robustness margins provide a cover for U (Figure 6.8 with $G = G_2$).

A different approach to construct G with the objective to optimise the expected performance would be to consider the probability distribution $q(m)$ of m imposed by the manufacturing process. We would then distribute the plants within G such that the grid is more 'dense' where the $p(m)$ is large (Figure 6.8 with $G = G_3$). This will on average reduce the distance between the true plant $p_* \in U$ and a plant $p \in G$. Since $|G_2| = |G_3|$ the computational complexity is equivalent to the choice $G = G_2$. However, we have at present no means of showing that this construction actually leads to an on average lower closed loop gain. Furthermore observe that the gain bound in Theorem 6.4 is weaker for $G = G_3$ than for $G = G_2$ since the maximum distance between $p \in U$ and the closest $p \in G$ is larger. We will show in the next section how probabilistic information may be utilised to explicitly improve the gain bound on average.

Note that although the cover (H, ν) is a powerful tool to deal with infinitely many plant models in G, the underlying principles that make the algorithm conservative remain in place.

8.2 Dynamic EMMSAC - refinement of G

Assume that the tolerance t is known and finite. Define U to be a constant plant generating operator where

$$\mathcal{P}^U = \bigcup_{0\leq j\leq t}\ \bigcup_{m\in\{m_1,m_2,m_3\}} \{(1\pm j)m\},\ j\in\mathbb{R}.$$

Also let the cover (H,ν) for U be constant.

A dynamic on-line refinement strategy for G, inspired by the schemes introduced in Chapter 4, Section 3, is given as follows.

Assume we determined a sufficiently dense, stabilising plant model set $G(0)$ such that the neighbourhoods $b_{P,C}$ around $m\in G(0)$ cover U. For example let $G(0)=G_2$ from Section 8.1. Note that such $G(0)$ can be constructed off-line. Then start the algorithm and construct further $G(k)\supset G(0),\ k>0$ on-line by interpolating new plant models in G with respect to the two smallest residuals; however only if the corresponding plants are adjacent. This is depicted in Figure 6.9.

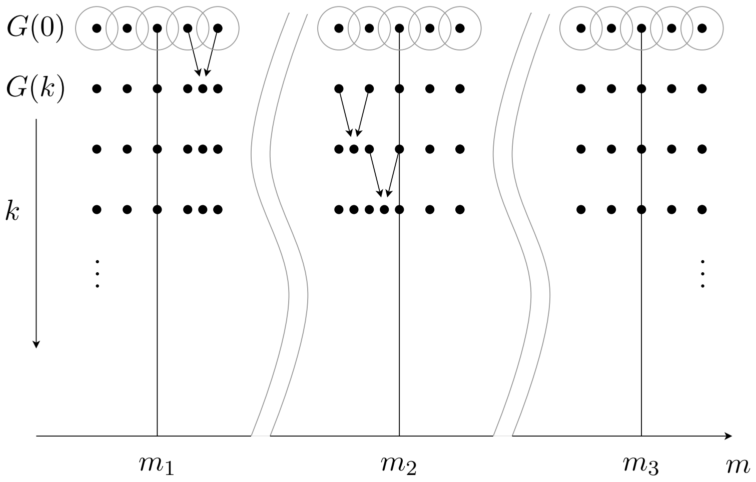

FIGURE 6.9: On-line refinement of G in respect to the size of residuals

A concrete algorithm could be given as follows. Let $G(0)=G_2$ be constructed as in Section 8.1. Let $a_i,b_i\in G(i),\ i\geq 0$ be such that the corresponding residuals r_{a_i} and r_{b_i} satisfy:

$$r_{a_i}[i]\leq r_m[i],\ \forall m\in G(i),\ r_{b_i}[i]\leq r_m[i],\ \forall m\in G(i)\setminus a_i.$$

If there are multiple minimal r_{a_i} or r_{b_i} choose the ones that minimise $|a_i|$ or $|b_i|$. Then let for $j\geq 0$ and some $x\in\mathbb{N}$

$$G((j+1)x)=G(jx)\cup m_{jx},\ G(i)=G(jx),\ jx\leq i<(j+1)x$$

where

$$m_{jx} = \begin{cases} \frac{a_{jx}+b_{jx}}{2} & \text{if } \nexists m \in G(jx) \text{ s.t. } a_{jx} < m < b_{jx} \text{ and } \delta(m_{jx}, m) > \epsilon,\ \forall m \in G(jx) \\ \emptyset & \text{otherwise} \end{cases}.$$

Note that new elements are introduced to G as long as the distance to previously introduced elements is above a certain pre-determined threshold ϵ, i.e. as long as $\delta(m_{jx}, m) > \epsilon,\ \forall m \in G(jx)$. Such a G is monotonic by construction hence all $G(k),\ k \in \mathbb{N}$ are sufficiently dense to provide a stabilising controller since $G(0)$ is sufficiently dense.

Although it seems intuitive that (on-line) refinement does improve performance — plant models are potentially closer to the true plant — it does not follow from the present analysis. To give analytic proof that (on-line) refinement does indeed improve performance remains open. The advantage of using on-line refinement in favour of a static EMMSAC algorithm based on a constant, highly refined plant model set G, is that it has the potential to utilise only a sufficient amount of computational resource (mainly determined by the number of plant models and corresponding estimators). This may, for example, be interesting from a power consumption point of view. However, in the worst case, the given on-line refinement scheme will introduce as many plant models as there is computational resource.

To suppress such behaviour, one could modify the scheme such that only a finite number n of refined plant models is allowed. When the scheme requests more than n plant models, one could, for example, remove the 'oldest' plant model (from a time of introduction point of view) in G which is not in $G(0)$. This would imply that the required amount of computational resource is bounded. Many other algorithms are thinkable.

8.3 Dynamic EMMSAC - expansion of G

For the purpose of this example we assume that the tolerance t is unknown. Furthermore we assume that the uncertainty set is finite and given by

$$\bigcup_{0 \le j \le 50t} \bigcup_{m \in \{m_1, m_2, m_3\}} \{(1 \pm 0.02j)m\},\ j \in \mathbb{N}$$

where the weight increase 0.02 is some small, physically meaningful number.

Finiteness of the uncertainty set is necessary since only then we can let $G = U$ and obtain a feasible EMMSAC design. This allows the direct application of Theorem 5.14 and Theorem 6.7 and makes the results comparable.

We have shown in Theorem 6.6 that for a constant plant-generating operator G the EMMSAC algorithm is conservative under an increasingly large parametric uncertainty. This was due to the fact that we can make the algorithm switch to the controller with the highest gain. We then employed a dynamically expanding plant model set to overcome this issue in Theorem 6.7. We now give two design strategies for a dynamically expanding G, based on the idea to use probabilistic information for ordering the level set $\mathcal{P}_i$.

Strategy 1:
Consider Figure 6.10. We construct G such that less probable parametrisations, indicated by $q(m)$, are introduced later. For that purpose let

$$\mathcal{P}_i^I = \bigcup_{0 \le j < i} \bigcup_{m \in \{m_1, m_2, m_3\}} \{(1 \pm 0.02j)m\}, \ i, j \in \mathbb{N}.$$

With the expansion rule $i(k)$, given by equation (6.10), this defines $G(k) = \mathcal{P}_{i(k)}^I, \ k \in \mathbb{N}$ where we let $G(k) = U(k) = H(k), \ k \in \mathbb{N}, \ \nu = 0$. We then have $N = 3$ and $\mathcal{P}_N^I = \mathcal{P}_3^I$

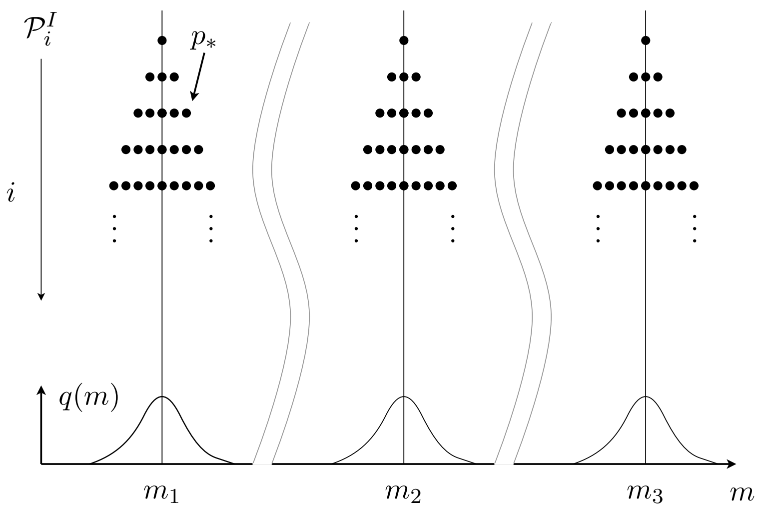

FIGURE 6.10: Strategy for designing the level set $\mathcal{P}_i$ with respect to $q(m)$

since

$$N := \min\{i \ge 1 \mid p_* \in \mathcal{P}_i\}$$

and $p_* = 1.04m_1 \notin \mathcal{P}_2^I$ but $p_* \in \mathcal{P}_3^I$. The gain function bound in Theorem 6.7 is therefore given by

$$\gamma_{mod}(\|w_0\|) = \beta_1 + \beta_2\|w_0\| + \beta_3\|w_0\|^2$$

where

$$\begin{aligned} \beta_1 &= \tilde{\gamma}^{v+2}(\mathcal{P}_3^I) + \tilde{\gamma}(\mathcal{P}_3^I)\tilde{\gamma}^v(\mathcal{P}_1^I) \\ \beta_2 &= 2\tilde{\gamma}^2(\mathcal{P}_3^I) + \tilde{\gamma}^{1-v}(\mathcal{P}_3^I)\tilde{\gamma}^v(\mathcal{P}_1^I) \\ \beta &= \tilde{\gamma}^{2-v}(\mathcal{P}_3^I). \end{aligned}$$

Now observe that $\beta_1, \beta_2, \beta_3$ are invariant to the tolerance t and constant where $\hat{\gamma}$ from the last section scales with t (it is conservative). We therefore conclude that there exists a tolerance t such that the gain function bound γ_{mod} is superior to the gain bound $\hat{\gamma}$.

A further strategy that is making use of even more a priori probabilistic information is given next.

Strategy 2:
We now also utilise the probabilistic information on how likely it is to encounter an item of a certain weight, indicated by $q_2(m)$, to further optimise the (expected value of the) average performance. For that purpose we modify the strategy from the last example and first introduce plants in the neighbourhood of the most likely item, then plants in the neighbourhood of the second most likely item and so on.

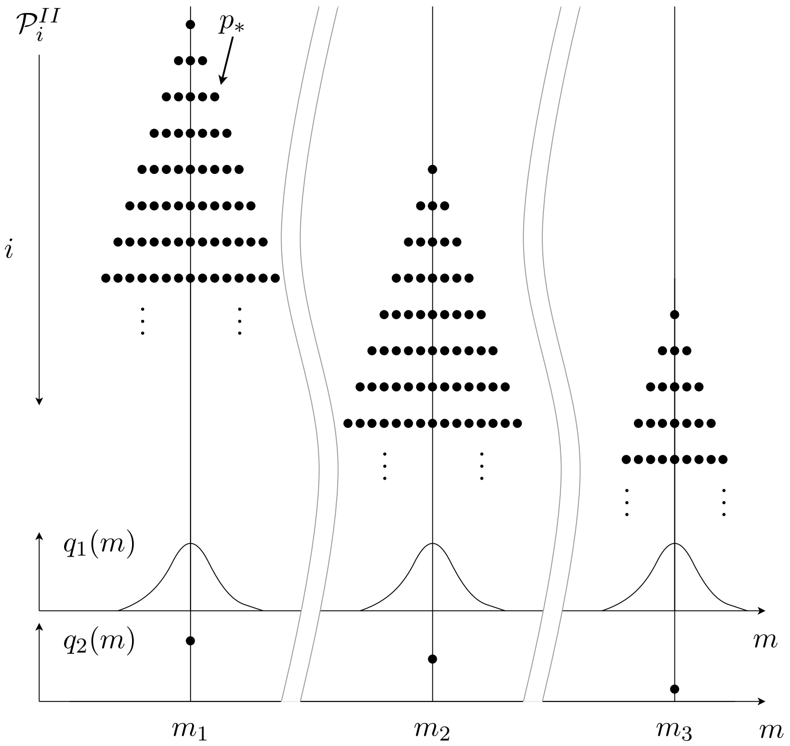

FIGURE 6.11: Strategy for designing a time varying G, minding $q_1(m)$ and $q_2(m)$

Such a strategy is depicted in Figure 6.11 where for $a, i \in \mathbb{N},\ i > 0$

$$\bar{\mathcal{P}}_i(m,a) \quad = \quad \begin{cases} \cup_{a \le j < i}\{(1 \pm 0.02(j-a))m\} & \text{if } a < i \\ \emptyset & \text{otherwise} \end{cases}$$

and

$$\mathcal{P}_i^{II} = \bar{\mathcal{P}}_i(m_1, 0) \cup \bar{\mathcal{P}}_i(m_2, 4) \cup \bar{\mathcal{P}}_i(m_3, 8).$$

With the expansion rule $i(k)$, given by equation (6.10), this defines $G(k) = \mathcal{P}_{i(k)}^{II},\ k \in \mathbb{N}$ where we let $G(k) = U(k) = H(k),\ k \in \mathbb{N},\ \nu = 0$.

For this setup $N = 3$ however, since $\mathcal{P}_3^{II} \subset \mathcal{P}_3^{I}$ and $\beta_1, \beta_2, \beta_3$ are all strictly increasing functions we can conclude that the second strategy yields a tighter bound.

Naturally, if the probabilistic assumptions about p_* are incorrect and p_* is close to its worst case, e.g. $p_* = (1+t)m_3$, the advantage over a constant plant set is lost. To see this, note that for the first strategy $N = 50t + 1$ where for the second strategy $N = 50t + 9$. With $\hat{\gamma} > 1$ from Theorem 5.14 and β_2 from Theorem 6.7 we have in either case

$$
\begin{aligned}
\bar{\gamma} &= \max_{p\in\mathcal{P}_N} \hat{\gamma}(\mathcal{P}_N, \mathcal{P}_N, 0, p) \\
&< \max_{p\in\mathcal{P}_N} (\beta(\mathcal{P}_N, \mathcal{P}_N, 0, p) + \hat{\gamma}(\mathcal{P}_N, \mathcal{P}_N, 0, p)) \\
&\leq \tilde{\gamma}(\mathcal{P}_N) \\
&\leq 2\tilde{\gamma}^2(\mathcal{P}_N) + \tilde{\gamma}^{1-v}(\mathcal{P}_N)\tilde{\gamma}^v(\mathcal{P}_N) \\
&= \beta_2.
\end{aligned}
$$

As discussed in Section 6, there are other sensible choices apart from $\|\mathscr{T}_k w_2\|$ for measuring the current performance at time $k \in \mathbb{N}$. However, the overall objective must be to dynamically expand G as a function of performance (determined by some measure) since only then are we able to overbound the gain in Theorem 5.14 to obtain a constant gain function. We can therefore expect similar tradeoffs for other algorithms that utilise performance information to dynamically expand plant model sets.

We have sketched how a priori information about the plant can be utilised to construct plant model sets and evaluated tradeoffs between the probability distribution on $\|w_0\|$, the probability distribution on the plant, $v > 2$ and the design of the plant model set itself. Furthermore, we have shown how on-line refinement may be conducted. A principled design methodology is within reach; however further research is required.

Chapter 7

Conclusion

This thesis presents comprehensive robustness and performance guarantees for Estimation-based Multiple Model Switched Adaptive Control (EMMSAC) algorithms in terms of $l_r,\ 1 \leq r \leq \infty$ gain (function) bounds on the gain from the external disturbances w_0 to the internal signals w_2. The axiomatic style and abstraction level of the analysis lead to the generality of the results: they apply to the class of minimal MIMO LTI plants but also to non-linear plants showing linear growth. Large classes of estimation algorithms, such as Kalman filters or (matrix) optimisation methods, may be utilised in the estimation process.

Remarkably, the style in which the analysis was conducted led to generalisations almost by accident (e.g. to the MIMO case and to the case of atomic non-linear plants showing linear growth) and makes future generalisations appear inevitable, e.g. to time-varying plants and to non-linear plants with super-linear growth.

It was shown that performance and robustness of the algorithm is guaranteed invariant to the refinement scaling of the plant model set. However it was also shown that if the plant model set is constant then performance and robustness diminish for expansion scaling of the plant model set corresponding to an increasing level of uncertainty — a static EMMSAC algorithm can be conservative. To overcome the conservativeness issue, an extension based on a dynamic (on-line) expansion law for the plant model set was introduced, which lead to the construction of gain function bounds that are invariant to the level of uncertainty — that is a dynamic EMMSAC algorithm is universal.

One particular feature of EMMSAC algorithms is that robustness guarantees can be supplied where LTI controllers fail to perform satisfactory or do not provide stability at all: for plants with large uncertainties and for non-simultaneously stabilisable plants. A qualitative, however completely rigorous discussion was provided, showing when dynamic EMMSAC promises tighter gain bounds than static EMMSAC and vice versa. Also dynamic (on-line) refinement schemes for the plant model set were discussed which seamlessly embed into the EMMSAC framework. Fundamental design questions on how to construct plant model sets, as posed at the beginning of this thesis (Chapter 1), have been addressed and answered from the perspective of prior information and performance.

1 Directions for future research

There are two specific technical questions that follow directly from this thesis: Firstly, it needs to be shown that both estimators can satisfy Assumption 3.4(5), or a modification thereof, with a continuous χ. Secondly, it needs to be investigated how sampling of a time-varying uncertainty set (e.g. a continuum) can lead to the construction of a realisable, stabilising time-varying plant model set (see Chapter 6, Section 4). A positive answer to the first question will allow the unconditional application of the algorithm to compact uncertainty sets, while the answer to the second question may allow the utilisation of sampled, finite plant model sets in dynamic expansion schemes for continuous uncertainty sets.

It is important to investigate the relationship between the distance χ (Assumption 3.4(5)) and the gap metric δ, since then the local cover constructions in terms of (H, ν) and the global cover constructions in terms of G and $b_{P,C}$ may be unified. The question of computing χ in different signal spaces also needs to be addressed.

An interesting question from a performance and design perspective is how the established gain bound may be improved. Superficially, there is plenty of room for such improvements since many simplifications and shortcuts in the analysis are conservative. However, the current bound appears to correctly specify (at least qualitatively) the tradeoffs involved in choosing the algorithm's key variables: the plant-generating operator G, the controller design procedure K as well as the attenuation function l and the delay Δ. We therefore do not expect significant qualitative changes in the bound — perhaps with the exception of (on-line) refinement of the plant model set since the current bound is invariant to refinement and does not reflect expected performance tradeoffs in this respect. This relationship needs to be established in order to make refinement schemes part of a performance-orientated design methodology.

Of great interest are also further schemes that exploit the freedom that G is allowed to be time-varying, although some may require a modified analysis. For example unfalsified control type schemes, where plant models are removed from the plant model set if it is unlikely that they represent the true plant (see Safonov and Tsao (1997)), or safe switching schemes (e.g. see Anderson et al. (2001)), where plant models are excluded if the corresponding controller could be destabilising to the true plant. Covariance information from the Kalman filter may be utilised in this respect to indicate the 'confidence' in a plant model. Schemes that could be implemented within the existing framework directly are, for example, a dynamic expansion/refinement law of the plant model set in relation to residual information from the estimators, or a scheme to momentarily disable estimators corresponding to plant models that will not be considered in the near future e.g. since their residuals are too large in relation to others.

Furthermore it is of great interest to broaden the underlying plant class further, for example to non-linear plants with super-linear growth. As a first step one could consider only local disturbances and overbound super-linear growth by local linear growth. However, in general, non-linear modifications to the controller assumptions and the analysis are required.

The algorithm can already be applied to mildly time-varying plants, where the variation is contained within a small neighbourhood. For larger variations the present estimator assumptions need to be modified to include some kind of 'forgetting factor'. A time-varying generalisation would also potentially allow the application of EMMSAC in the domain of fault detection and fault tolerant control. This link is very interesting since many algorithms in this area are based on Kalman filters.

The investigation of disturbance estimation algorithms which are low in complexity and allow large or even continuous plant model sets is important in order to fully exploit the EMMSAC approach in practice. First steps in this direction could be the use of state sharing ideas, e.g. in the style of Morse, for disturbance estimation. Analogously to a state shared observer, a bank of optimal estimators then shares common information in order to reduce computational complexity. The construction of the estimator necessarily leads to the question of implementation, i.e. to find efficient, numerically stable hardware estimator implementations.

Another open question here is the relationship between optimal estimators and (output error type) observers. This relationship appears to be close (the Kalman filter estimator has observer structure) and it may be possible to treat general (non Kalman filter type) observers as sub-optimal estimators. If this link can be made explicit in terms of bounds between residuals, then the presented theory would encompass the class of observer based multiple model switched adaptive control algorithms.

Further research is needed to conduct a fully Bayesian treatment of the plant model set design problem. I.e. given a signal norm l_r, $1 \leq r \leq \infty$, (user) constraints on the attenuation function l as well as the delay Δ and given probability distributions on the uncertainty and the disturbance signal w_0, a general formalism needs to be constructed that has a (time-varying) plant model set G and the delay Δ as an outcome. For dynamic EMMSAC, such a design flow requires a sensible interpretation of a time varying U in respect to the uncertainty description the control problem. A formalised approach to design would provide a considerable advance over existing theory.

Chapter 8

Appendix

Let P_{p,x_0^p} be defined by

$$P_{p,x_0^p} : \mathcal{U}_e \to \mathcal{Y}_e : u_1^p \mapsto y_1^p,\ p = (A_p, B_p, C_p) \in \bar{\mathcal{P}}_{LTI} \tag{A.1}$$

where

$$x_p(k+1) = A_p x_p(k) + B_p u_1^p(k) \tag{A.2}$$
$$y_1^p(k) = C_p x_p(k) \tag{A.3}$$
$$x_p(0) = x_0^p,\ k \in \mathbb{N}. \tag{A.4}$$

This definition is similar to the one in equations (3.4)–(3.4) however with a possibly non-zero initial condition x_0^p.

Let the Kalman filter to a plant P_{p,x_0^p} with

$x = x_p,\ (w, v)^\top = (u_0^p, y_0^p)^\top, (u, y)^\top = (u_2, y_2)^\top,\ (F, G, B, H) = (A_p, B_p, -B_p, C_p), n = n_p$

and $T \geq 0$, $\Sigma : \mathbb{N} \mapsto \mathbb{R}^{n\times n}$, $\hat{x} : [0, T] \mapsto \mathbb{R}^n$ given by

$$\hat{x}(k+1/2) = \hat{x}(k) + \Sigma(k)H^\top[H\Sigma(k)H^\top + I]^{-1}[y(k) - H\hat{x}(k)] \tag{A.5}$$
$$\Sigma(k+1/2) = \Sigma(k) - \Sigma(k)H^\top[H\Sigma(k)H^\top + I]^{-1}H\Sigma(k) \tag{A.6}$$

$$\hat{x}(k+1) = F\hat{x}(k+1/2) + Bu(k) \tag{A.7}$$
$$\Sigma(k+1) = F\Sigma(k+1/2)F^\top + GG^\top \tag{A.8}$$
$$\tilde{y}_1(k) = H\hat{x}(k) \tag{A.9}$$

where $\Sigma(0) = \Sigma(0)^\top \in \mathbb{R}^{n\times n}$ and $\Sigma(0) = \Sigma(0)^\top \geq 0$.

Define as a notion of the output error between observation and estimation the (scaled) residual $r : \mathbb{N} \to \mathbb{R}^+$ by

$$r(T) = \left[\sum_{k=0}^{T} (\|y(k) - H\hat{x}(k)\|^2_{[H\Sigma(k)H^\top + I]^{-1}}) \right]^{1/2} = \left[\sum_{k=0}^{T} \|y(k) - \tilde{y}_1(k)\|^2_{[H\Sigma(k)H^\top + I]^{-1}} \right]^{1/2}$$

for $T \geq 0$.

The following lemma shows that r is defined, i.e. that $H\Sigma(k)H^\top + I$ is always invertible. It formalises known properties of the discrete-time Riccati equation, which nevertheless appear hard to source in the literature.

Lemma A.1. *Let* $(F, G, H) \in \bar{\mathcal{P}}_{LTI}$ *and suppose* H *is full row rank. Let the Kalman filter equations for* Σ *be given by equations* (A.6),(A.8). *If* $\Sigma(0) = \Sigma^\top(0) > 0$ *then* $\Sigma(k) = \Sigma^\top(k) > 0$ *for all* $k \geq 0$. *If* $\Sigma(0) = \Sigma^\top(0) \geq 0$ *then* $\Sigma(k) = \Sigma^\top(k) \geq 0$ *for all* $k \geq 0$.

Proof Let $k > 0$. We first show that $\Sigma(0) = \Sigma^\top(0) > 0$ implies $\Sigma(k) = \Sigma^\top(k) > 0$. The proof is by induction.

Since $\Sigma(0) = \Sigma(0)^\top > 0$ by assumption, the base step holds trivially.

For the inductive step have to show that $\Sigma(k) = \Sigma^\top(k) > 0$ implies $\Sigma(k+1) = \Sigma^\top(k+1) > 0$.

Substituting equation (A.6) into equation (A.8) leads to

$$\Sigma(k+1) = F\Sigma(k)F^\top - F\Sigma(k)H^\top(H\Sigma(k)H^\top + I)^{-1}H\Sigma(k)F^\top + GG^\top. \tag{A.10}$$

From $\Sigma(k) = \Sigma^\top(k) > 0$, it follows that $H\Sigma(k)H^\top + I$ is symmetric and invertible, hence

$$\left((H\Sigma(k)H^\top + I)^{-1}\right)^\top = \left((H\Sigma(k)H^\top + I)^T\right)^{-1}$$

and we have from equation (A.10) that:

$$\begin{aligned}
\Sigma(k+1)^\top &= \left(F\Sigma(k)F^\top - F\Sigma(k)H^\top(H\Sigma(k)H^\top + I)^{-1}H\Sigma(k)F^\top + GG^\top\right)^\top \\
&= F\Sigma(k)F^\top - F\Sigma(k)H^\top \left((H\Sigma(k)H^\top + I)^{-1}\right)^\top H\Sigma(k)F^\top + GG^\top \\
&= F\Sigma(k)F^\top - F\Sigma(k)H^\top \left((H\Sigma(k)H^\top + I)^T\right)^{-1} H\Sigma(k)F^\top + GG^\top \\
&= F\Sigma(k)F^\top - F\Sigma(k)H^\top(H\Sigma(k)H^\top + I)^{-1}H\Sigma(k)F^\top + GG^\top \\
&= \Sigma(k+1).
\end{aligned}$$

Observe that equation (A.10) can be written as

$$\Sigma(k+1) = \begin{bmatrix} F & G \end{bmatrix} \begin{bmatrix} \Sigma(k) - \Sigma(k)H^\top(H\Sigma(k)H^\top + I)^{-1}H\Sigma(k) & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} F & G \end{bmatrix}^\top. \tag{A.11}$$

Since $(F,G,H) \in \bar{P}_{LTI}$ is minimal, (F,G) is controllable. By the Popov-Belevitch-Hautus (PBH) test this implies that $\begin{bmatrix} Iz-F & G \end{bmatrix}$, $z \in \mathbb{C}$ is full row rank (e.g. see Hendricks et al. (2009), page 141 or Hogben et al. (2007), page 57-8) hence with $z=0$ that $\begin{bmatrix} -F & G \end{bmatrix}$ is full row rank. Since left or right multiplication by a non-singular matrix is a rank preserving operation (e.g. see Hogben et al. (2007), page 2-4) and $\begin{bmatrix} F & G \end{bmatrix} = \begin{bmatrix} -F & G \end{bmatrix} \begin{bmatrix} -I & 0 \\ 0 & I \end{bmatrix}$ we have that also $\begin{bmatrix} F & G \end{bmatrix}$ is full row rank. By equation (A.11) it follows that $\Sigma(k+1)$ is positive definite if

$$\Sigma(k) > \Sigma(k)H^\top (H\Sigma(k)H^\top + I)^{-1} H\Sigma(k). \tag{A.12}$$

Since H is full row rank, inequality (A.12) holds if

$$H\Sigma(k)H^\top > H\Sigma(k)H^\top (H\Sigma(k)H^\top + I)^{-1} H\Sigma(k)H^\top \tag{A.13}$$

holds, where inequality (A.13) is derived by left and right multiplication of inequality (A.12) with H and $H^\top$.

Let $\alpha = H\Sigma(k)H^\top > 0$. Then $\alpha = H\Sigma(k)H^\top + I > 0$, hence $\alpha + I$ is invertible and $(\alpha+I)^{-1} > 0$. Since $\alpha(\alpha+I) = (\alpha+I)\alpha$, it follows that $(\alpha+I)^{-1}\alpha = \alpha(\alpha+I)^{-1}$, and hence $\alpha(\alpha+I)^{-1} > 0$ (see Horn and Johnson (1990)).

Then:

$$0 < \alpha(\alpha+I)^{-1} = \alpha(\alpha+I)^{-1}(I+\alpha-\alpha) = \alpha - \alpha(\alpha+I)^{-1}\alpha$$

hence

$$H\Sigma(k)H^\top > H\Sigma(k)H^\top (H\Sigma(k)H^\top + I)^{-1} H\Sigma(k)H^\top$$

and so $\Sigma(k+1) = \Sigma(k+1)^\top > 0$ if $\Sigma(k) = \Sigma(k)^\top > 0$.

This completes the induction and we conclude that if $\Sigma(0) = \Sigma(0)^\top > 0$ then $\Sigma(k) = \Sigma(k)^\top > 0$ for all $k \in \mathbb{N}$. The same argument holds in the semi-definite case with $\geq$ instead of $>$ in the above inequalities and noting that $\alpha \geq 0 \;\Rightarrow\; (\alpha+I) > 0$ and hence is invertible. Therefore $\Sigma(k+1) \geq 0$ if $\Sigma(k) \geq 0$ as required. □

1 Half-step identities

We now give two key identities that are crucial to subsequent calculations.

Lemma A.2. *Let* $x, m \in \mathbb{R}^n$, $v \in \mathbb{R}^o$, $H \in \mathbb{R}^{o \times x}$ *and* $\Sigma \in \mathbb{R}^{n \times n}$, $\Sigma = \Sigma^\top > 0$ *where* $m, n, o \in \mathbb{N}$. *Then* $H\Sigma H^\top + I$ *is invertible. Define*

$$\begin{aligned} y &= Hx + v \\ \hat{x} &= m + \Sigma H^\top \left[H\Sigma H^\top + I \right]^{-1} (y - Hm) \\ \hat{\Sigma} &= \Sigma - \Sigma H^\top \left[H\Sigma H^\top + I \right]^{-1} H\Sigma. \end{aligned}$$

Then $\Sigma, \hat{\Sigma}$ *are invertible and the following identity holds:*

$$(x - \hat{x})^\top \hat{\Sigma}^{-1} (x - \hat{x}) = v^\top v + (x - m)^\top \Sigma^{-1} (x - m) - (y - Hm)^\top \left[H\Sigma H^\top + I \right]^{-1} (y - Hm)$$

Proof For notational convenience let

$$\begin{aligned} \alpha &= \left[H\Sigma H^\top + I \right]^{-1} \\ \beta &= I - \Sigma H^\top \alpha H \end{aligned}$$

where we note that $H\Sigma H^\top + I$ is invertible since $H\Sigma H^\top$ is positive semi-definite. Then

$$\begin{aligned} \hat{x} &= \beta m + \Sigma H^\top \alpha H x + \Sigma H^\top \alpha v \\ \hat{\Sigma} &= \beta \Sigma. \end{aligned}$$

It now follows that:

$$\begin{aligned} (x - \hat{x})^\top \hat{\Sigma}^{-1} (x - \hat{x}) &= \| x - \Sigma H^\top \alpha H x - \beta m - \Sigma H^\top \alpha v \|^2_{\hat{\Sigma}^{-1}} \\ &= \| \beta x - \beta m - \Sigma H^\top \alpha v \|^2_{\hat{\Sigma}^{-1}} \\ &= (\beta(x - m) - \Sigma H^\top \alpha v)^\top \hat{\Sigma}^{-1} (\beta(x - m) - \Sigma H^\top \alpha v) \\ &= ((x - m)^\top \beta^\top - v^\top \alpha H \Sigma) \hat{\Sigma}^{-1} (\beta(x - m) - \Sigma H^\top \alpha v). \end{aligned}$$

Observe that β is invertible and

$$\beta^{-1} = I + \Sigma H^\top H.$$

Since $\hat{\Sigma}$ is symmetric, $\hat{\Sigma}^{-1} = \Sigma^{-1}\beta^{-1} = (\beta^\top)^{-1}\Sigma^{-1}$, hence

$$\begin{aligned}
(x-\hat{x})^\top \hat{\Sigma}^{-1}(x-\hat{x}) &= ((x-m)^\top \Sigma^{-1} - v^\top \alpha H \beta^{-1})(\beta(x-m) - \Sigma H^\top \alpha v) \\
&= (x-m)^\top \Sigma^{-1}\beta(x-m) - v^\top \alpha (H(x-m)) - (H(x-m))^\top \alpha v \\
&\quad + v^\top \alpha H \beta^{-1} \Sigma H^\top \alpha v \\
&= (x-m)^\top \Sigma^{-1}(x-m) - (H(x-m))^\top \alpha (H(x-m)) \\
&\quad - v^\top \alpha (H(x-m)) - (H(x-m))^\top \alpha v \\
&\quad + v^\top \alpha H \beta^{-1} \Sigma H^\top \alpha v \\
&= (x-m)^\top \Sigma^{-1}(x-m) - (H(x-m)+v)^\top \alpha (H(x-m)+v) \\
&\quad + v^\top \left(\alpha + \alpha H \beta^{-1} \Sigma H^\top \alpha\right) v.
\end{aligned}$$

It remains to show that $\left(\alpha + \alpha H \beta^{-1} \Sigma H^\top \alpha\right) = I$.

To see this observe that

$$\begin{aligned}
\alpha + \alpha H \beta^{-1} \Sigma H^\top \alpha &= \alpha + \alpha H \Sigma H^\top \alpha + \alpha H \Sigma H^\top H \Sigma H^\top \alpha \\
&= \alpha + \alpha H \Sigma H^\top \underbrace{\left[H \Sigma H^\top + I\right]}_{\alpha^{-1}} \alpha \\
&= \alpha \underbrace{\left[H \Sigma H^\top + I\right]}_{\alpha^{-1}} \\
&= I
\end{aligned}$$

as required. □

Note that α is more than a simple notational convenience. It turns out to be the scaling factor in the least-squares calculation below.

Before we state the second key identity we give a preliminary result.

Lemma A.3. *Let $(F, G, H) \in \bar{\mathcal{P}}_{LTI}$ and suppose H is full row rank. Let $\Sigma_1 = \Sigma_1^\top > 0$. Then there exists $\begin{bmatrix} K & L \end{bmatrix}$ such that*

$$\begin{bmatrix} F & G \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} K^\top \\ L^\top \end{bmatrix} = 0 \tag{A.14}$$

$$\begin{bmatrix} K & L \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} F^\top \\ G^\top \end{bmatrix} = 0 \tag{A.15}$$

$$\begin{bmatrix} K & L \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} K^\top \\ L^\top \end{bmatrix} = I \tag{A.16}$$

and $\begin{bmatrix} F & G \\ K & L \end{bmatrix}$ is invertible.

Proof Since $(F, G, \cdot) \in \bar{P}_{LTI}$ are minimal, (F, G) is controllable. This implies that $\begin{bmatrix} F & G \end{bmatrix}$ is full row rank (see the proof of Lemma A.1 above). Let $V = rowspan(\begin{bmatrix} F & G \end{bmatrix})$. After Gram-Schmidt we can construct a orthonormal basis for $V^\perp$ with respect to the weight $\begin{bmatrix} \Sigma_1 & 0 \\ 0 & I \end{bmatrix} = W$ and the weighted scalar product $\langle x, y \rangle_W = x^\top W y$. Let the basis vectors of $V^\perp$ be the rows of $\begin{bmatrix} K & L \end{bmatrix}$.

Equations (A.14)–(A.16) now follow directly from the definition of the weighted scalar product. Since $\begin{bmatrix} F & G \end{bmatrix}$ is full row rank, $\begin{bmatrix} K & L \end{bmatrix}$ and hence $\begin{bmatrix} F & G \\ K & L \end{bmatrix}$ is full rank. Therefore $\begin{bmatrix} F & G \\ K & L \end{bmatrix}$ is invertible as required. □

We now come to the second key identity:

Lemma A.4. *Let $(F, G, H) \in \bar{\mathcal{P}}_{LTI}$ and suppose H is full row rank. Let $\Sigma_1 = \Sigma_1^\top > 0$. Define*

$$\Sigma_2 := F \Sigma_1 F^\top + G G^T. \tag{A.17}$$

Then Σ_1, Σ_2 are invertible and there exist K,L such that

$$\|Fa + Gb\|^2_{\Sigma_2^{-1}} + \|Ka + Lb\|^2 = \|a\|^2_{\Sigma_1^{-1}} + \|b\|^2$$

for all a,b.

Proof Let $\begin{bmatrix} K & L \end{bmatrix}$ be constructed as in Lemma A.3. From equation (A.17) and Lemma A.3 it follows that

$$\begin{bmatrix} \Sigma_2 & 0 \\ 0 & I \end{bmatrix} = \begin{bmatrix} F & G \\ K & L \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} F & G \\ K & L \end{bmatrix}^\top .$$

Since $\begin{bmatrix} F & G \\ K & L \end{bmatrix}$ is invertible by construction and $\Sigma_1 > 0$ it follows that Σ_2 is invertible and

$$\begin{bmatrix} F & G \\ K & L \end{bmatrix}^\top \begin{bmatrix} \Sigma_2^{-1} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} F & G \\ K & L \end{bmatrix} = \begin{bmatrix} \Sigma_1^{-1} & 0 \\ 0 & I \end{bmatrix}. \tag{A.18}$$

Therefore by equation (A.18) for all $\begin{bmatrix} a \\ b \end{bmatrix}$ there holds

$$\begin{aligned} \|Fa + Gb\|^2_{\Sigma_2^{-1}} + \|Ka + Lb\|^2 &= \begin{bmatrix} a \\ b \end{bmatrix}^\top \begin{bmatrix} F & G \\ K & L \end{bmatrix}^\top \begin{bmatrix} \Sigma_2^{-1} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} F & G \\ K & L \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} \\ &= \begin{bmatrix} a \\ b \end{bmatrix}^\top \begin{bmatrix} \Sigma_1^{-1} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} \\ &= \|a\|^2_{\Sigma_1^{-1}} + \|b\|^2. \end{aligned}$$

as required. □

2 Kalman filtering and least squares

We are now in the position to combine the two established key lemmas.

Lemma A.5. *Let $(A_p, B_p, C_p) \in \bar{\mathcal{P}}_{LTI}$ and suppose C_p is full row rank. Let $(F, G, B, H) = (A_p, B_p, -B_p, C_p)$. Let $v, w \in l_2$ and $\Sigma(0) = \Sigma(0)^\top > 0$. Then for all $T \geq 0$ there exists a $K(k),\ L(k),\ k \in [0, T]$ such that*

$$\begin{aligned} \|x(T+1) - \hat{x}(T+1)\|^2_{\Sigma(T)^{-1}} = \ & \|x(0) - \hat{x}(0)\|^2_{\Sigma(0)^{-1}} + \sum_{k=0}^{T} (\|w(k)\|^2 + \|v(k)\|^2) \\ & - \sum_{k=0}^{T} (\|y(k) - H\hat{x}(k)\|^2_{[H\Sigma(k)H^\top + I]^{-1}}) \\ & - \sum_{k=0}^{T} \|K(k)[x(k) - \hat{x}(k+1/2)] + L(k)w(k)\|^2 \end{aligned} \tag{A.19}$$

and $\begin{bmatrix} K(k) & L(k) \end{bmatrix}$ is full rank.

Proof Let $k \in [0, T]$. By Lemma A.1 $\Sigma(i) = \Sigma(i)^\top > 0$ for all $i \geq 0$ hence $\Sigma(k) = \Sigma^\top(k) > 0$ is invertible. From equations (A.5)–(A.9) and Lemma A.2 we have with

$$\hat{\Sigma} = \Sigma(k+1/2),\ \Sigma = \Sigma(k),\ m = \hat{x}(k),\ y = y(k),\ v = v(k),\ x = x(k),\ \hat{x} = \hat{x}(k+1/2)$$

that $\Sigma(k+1/2)$ is invertible and

$$\|x(k) - \hat{x}(k+1/2)\|^2_{\Sigma(k+1/2)^{-1}} - \|x(k) - \hat{x}(k)\|^2_{\Sigma(k)^{-1}} = \|v(k)\|^2 - \|y(k) - H\hat{x}(k)\|^2_{[H\Sigma(k)H^\top + I]^{-1}}.$$

Furthermore since $x(k+1) = Fx(k) + Bu + Gw$ we have by equations (A.2), equations (A.7), (A.8) and Lemma A.4 with

$$\Sigma_1 = \Sigma(k+1/2),\ \Sigma_2 = \Sigma(k+1),\ K = K(k),\ L = L(k),\ a = x(k) - \hat{x}(k+1/2),\ b = w$$

that

$$\begin{aligned}\|x(k+1) - \hat{x}(k+1)\|^2_{\Sigma(k+1)^{-1}} &- \|x(k) - \hat{x}(k+1/2)\|^2_{\Sigma(k+1/2)^{-1}} \\ &= \|w(k)\|^2 - \|K(k)[x(k) - \hat{x}(k+1/2)] + L(k)w(k)\|^2.\end{aligned}$$

Adding these two equalities gives

$$\begin{aligned}\|x(k+1) - \hat{x}(k+1)\|^2_{\Sigma(k+1)^{-1}} &- \|x(k) - \hat{x}(k)\|^2_{\Sigma(k)^{-1}} \\ &= \|v(k)\|^2 + \|w(k)\|^2 - \|y(k) - H\hat{x}(k)\|^2_{[H\Sigma(k)H^\top + I]^{-1}} \\ &\quad - \|K(k)\,[x(k) - \hat{x}(k+1/2)] + L(k)w(k)\|^2 .\end{aligned}$$

Summing from $k = 0$ to $k = T$ leads to equation (A.19) as required. □

Define

$$\mathcal{Z}_p^{[a,b]}(w_2) = \left\{ v \in \mathbb{R}^{m(T+1)} \times \mathbb{R}^{o(T+1)} \times \mathbb{R}^n \;\middle|\; \begin{array}{c} \exists (u_0^p, y_0^p, x_0^p)^\top \in \mathcal{U}_e \times \mathcal{Y}_e \times \mathbb{R}^n \ s.t. \\ \mathscr{R}_{b-a,b} P_{p,x_0^p}(u_0^p - u_2) = \mathscr{R}_{b-a,b}(y_0^p - y_2), \\ v = \left(\mathscr{R}_{b-a,b} u_0^p, \mathscr{R}_{b-a,b} y_0^p, x_p^0\right)^\top \end{array} \right\}$$

which is the set of initial conditions x_0^p and disturbance signals u_0^p, y_0^p that are compatible with a plant P_{p,x_0^p} and the observation u_2, y_2 over the interval $[a, b],\ a \leq b$.

Theorem A.6. [Theorem 3.9] *Let $p = (A_p, B_p, C_p) \in \bar{\mathcal{P}}_{LTI}$ and suppose C_p is full row rank. Let $(F, G, B, H) = (A_p, B_p, -B_p, C_p)$. The Kalman filter equations* (A.5)–(A.9) *with initial condition $\hat{x}(0) = 0$ and $\Sigma(0) = \Sigma(0)^\top > 0$ describe a deterministic least-squares filter:*

$$r^2(T) = \inf_{(u_0^p, y_0^p, x_0^p) \in \mathcal{Z}_p^{[0,T]}(w_2)} (\|x_0^p\|^2_{\Sigma^{-1}(0)} + \|u_0^p\|_2^2 + \|y_0^p\|_2^2).$$

Proof Let $\hat{x}(0) = 0$. We then have from equality (A.19) that:

$$\sum_{k=0}^{T} (\|y(k) - H\hat{x}(k)\|^2_{[H\Sigma(k)H^\top + I]^{-1}}) \leq \|x(0)\|^2_{\Sigma^{-1}(0)} + \sum_{k=0}^{T} (\|w(k)\|^2 + \|v(k)\|^2). \quad \text{(A.20)}$$

where $\hat{x}$ is generated from y by equations (A.5),(A.7) and Σ is from equations (A.6),(A.8).

Observe that $\hat{x}$ depends on y but not on the disturbances v, w and the initial condition $x(0)$ that generated y. Hence

$$\|x(0)\|^2_{\Sigma^{-1}(0)} + \sum_{k=0}^{T}(\|w(k)\|^2 + \|v(k)\|^2)$$

is minimised if equality holds in inequality (A.20).

With the sufficient conditions $\hat{x}(0) = 0$, $x(T+1) = \hat{x}(T+1)$ and

$$\begin{aligned} &K(k)\left[x(k) - \hat{x}(k+1/2)\right] + L(k)w(k) = \\ &\quad K(k)\Big[x(k) - \hat{x}(k) - \Sigma(k)H^\top[H\Sigma(k)H^\top + I]^{-1}[y(k) - H\hat{x}(k)]\Big] + L(k)w(k) = 0, \end{aligned}$$

for $k \in [0, T]$ we have from (A.19) that

$$\|x(0)\|^2_{\Sigma^{-1}(0)} + \sum_{k=0}^{T}(\|w(k)\|^2 + \|v(k)\|^2) = \sum_{k=0}^{T}(\|y(k) - H\hat{x}(k)\|^2_{[H\Sigma(k)H^\top + I]^{-1}}).$$

In the following we show that these sufficient conditions can be met. From equation (A.2), describing P_{p,x_0^p}, we have with

$$w = u_0^p, u = u_1^p,\ (F, G, B) = (A_p, B_p, -B_p),\ x = x_p$$

that

$$x(k+1) = Fx(k) + Gw(k) + Bu(k),$$

hence we obtain for $k \in [0, T]$

$$\begin{bmatrix} x(k+1) - Bu(k) \\ K(k)\left[\hat{x}(k) + \Sigma(k)H^\top[H\Sigma(k)H^\top - I]^{-1}[y(k) - H\hat{x}(k)]\right] \end{bmatrix} = \begin{bmatrix} F & G \\ K(k) & L(k) \end{bmatrix} \begin{bmatrix} x(k) \\ w(k) \end{bmatrix}.$$

Since $\begin{bmatrix} F & G \\ K(k) & L(k) \end{bmatrix}$ is invertible and $u(k), y(k),\ \hat{x}(k),\ \Sigma(k),\ k \in [0, T]$ are known, this can be solved backwards for $x(k),\ w(k),\ k \in [0, T+1]$. Therefore there exist solutions $x = \tilde{x}$ and $w = \tilde{w}$ for $\hat{x}(0) = 0$ such that $x(T+1) = \hat{x}(T+1)$. Hence

$$\begin{aligned} &K(k)\Big[x(k) - \hat{x}(k) - \Sigma(k)H^\top[H\Sigma(k)H^\top + I]^{-1}[y(k) - H\hat{x}(k)\Big] \\ &\qquad\qquad\qquad\qquad + L(k)w(k) = 0,\ k \in [0, T]. \end{aligned}$$

Recall that $(x(0), w, v)^\top = (x_0^p, u_0^p, y_0^p)^\top$ and $(u, y)^\top = (u_2, y_2)^\top$. To see that the Kalman filter is a least-squares filter observe that if

$$x_p(k+1) = Ax_p(k) + Bu_1^p(k) = Ax_p(k) + B(u_2(k) - u_0^p(k))$$

is initialised with $x_p(0) = x_0^p = \tilde{x}(0)$ and driven by $u_0^p(k) = \tilde{w}(k),\ k \in [0,T]$ then $x_p(T+1) = \hat{x}(T+1)$. Hence

$$\begin{aligned}\|x(0)\|^2_{\Sigma^{-1}(0)} + \sum_{k=0}^{T}(\|w(k)\|^2 + \|v(k)\|^2) \\ = \inf_{(u_0^p, y_0^p, x_0^p) \in \mathcal{Z}_p^{[0,T]}(w_2)} (\|x_0^p\|^2_{\Sigma^{-1}(0)} + \|u_0^p\|_2^2 + \|y_0^p\|_2^2) \\ = \sum_{k=0}^{T}(\|y(k) - H\hat{x}(k)\|^2_{[H\Sigma(k)H^\top + I]^{-1}}) = r^2(T)\end{aligned}$$

as required. □

At this point we emphasise that $(\tilde{w}, \tilde{v}, \tilde{x})^\top$ are generated by the least-squares filter in a non-recursive way. This, however, does not matter since more importantly $\hat{x}(k), \Sigma(k),\ k \in \mathbb{N}$ are recursively generated via the Kalman filter equations and so is the residual $r(k)$.

Before we come to our last Theorem, showing the relation of the Kalman filter to a least-squares filter for the initial condition $\Sigma(0) = 0$, we establish the following lemma:

Lemma A.7. *Let L be a closed subset of $\mathbb{R}^n$. Let*

$$(\tilde{x}^n, \tilde{y}^n) = \operatorname*{argmin}_{(x,y)\in L}(n\|x\|^2 + \|y\|^2). \tag{A.21}$$

Suppose $(\tilde{x}^n, \tilde{y}^n) \to (0, \tilde{y})$ as $n \to \infty$. Then

$$(0, \tilde{y}) = \operatorname*{argmin}_{(0,y)\in L} \|y\|^2.$$

Proof Suppose

$$(0, y) = \operatorname*{argmin}_{(0,y)\in L} \|y\|^2. \tag{A.22}$$

Since L is closed, $(\tilde{x}^n, \tilde{y}^n) \in L$ implies $(0, \tilde{y}) \in L$. Therefore we have

$$\|(0, y)\|^2 \leq \|(0, \tilde{y})\|^2 \tag{A.23}$$

since $(0, y)$ is the minimiser, but $(0, \tilde{y}) \in L$ is not necessarily the minimiser.

For all $n \geq 1$ we also have from equation (A.21) that

$$n\|\tilde{x}^n\|^2 + \|\tilde{y}^n\|^2 \leq n\|\bar{x}\|^2 + \|\bar{y}\|^2$$

for any $(\bar{x}, \bar{y}) \in L$, in particular if $(0, y) \in L$ then

$$n\|\tilde{x}\|^2 + \|\tilde{y}\|^2 \leq n\|0\|^2 + \|y\|^2 = \|y\|^2.$$

We therefore arrive at

$$\|y\|^2 \geq n\|\tilde{x}^n\|^2 + \|\tilde{y}^n\|^2 \geq \|\tilde{y}^n\|^2 \geq \|\tilde{y} - \tilde{y} + \tilde{y}^n\|^2 \geq (\|\tilde{y}\| - \|\tilde{y}^n - \tilde{y}\|)^2 .$$

Since $\tilde{y}^n \to \tilde{y}$, $\|\tilde{y}_n - \tilde{y}\| \to 0$ hence $\|y\|^2 \geq \|\tilde{y}\|^2$ and therefore

$$\|(0, y)\|^2 \geq \|0, \tilde{y}\|^2. \tag{A.24}$$

Inequalities (A.23),(A.24) now lead to

$$\|(0, \tilde{y})\| = \|(0, y)\| = \underset{(0,y)\in L}{\operatorname{argmin}} \|y\|^2$$

as required. □

Recall the definition of $\mathcal{N}_p^{[a,b]}$ from Chapter 3:

$$\mathcal{N}_p^{[a,b]}(w_2) := \left\{ v \in \mathcal{W}|_{[a,b]} \;\middle|\; \begin{array}{c} \exists (u_0^p, y_0^p)^\top \in \mathcal{W}_e \ s.t. \\ \mathscr{R}_{b-a,b} P_p \,(u_0^p - u_2) = \mathscr{R}_{b-a,b}(y_0^p - y_2), \\ v = (\mathscr{R}_{b-a,b} u_0^p, \mathscr{R}_{b-a,b} y_0^p) \end{array} \right\} \subset \mathcal{W}|_{[a,b]}.$$

Hence $\mathcal{N}_p^{[0,T]}(w_2) = \mathcal{Z}_p^{[0,T]}(w_2)$ when $x_p^0 = 0$. The next theorem is to handle this case ($x_p^0 = 0$) in contrast to the previous theorem where this is not enforced.

Theorem A.8. [Theorem 3.10] *Let* $p = (A_p, B_p, C_p) \in \bar{\mathcal{P}}_{LTI}$ *and suppose* C_p *is full row rank. Let* $(F, G, B, H) = (A_p, B_p, -B_p, C_p)$. *The Kalman filter equations* (A.5)–(A.9) *with initial condition* $\hat{x}(0) = 0$ *and* $\Sigma(0) = \Sigma(0)^\top = 0$ *describe a deterministic least-squares filter initialised to zero:*

$$r^2(T) = \inf_{(u_0^p, y_0^p) \in \mathcal{N}_p^{[0,T]}(w_2)} (\|u_0^p\|_2^2 + \|y_0^p\|_2^2).$$

Proof Let $T \in \mathbb{N}$. For $n \in \mathbb{N}$ define $\Sigma_n(0) = \frac{1}{n} I$. So $\Sigma_n(0) = \Sigma_n(0)^\top > 0$. Let

$$(\tilde{u}_0^n, \tilde{y}_0^n, \tilde{x}_0^n) = \underset{(u_0, y_0, x_0) \in \mathcal{Z}_p^{[0,T]}(w_2)}{\operatorname{argmin}} (\|x_0\|^2_{\Sigma_n^{-1}(0)} + \|u_0\|_2^2 + \|y_0\|_2^2)$$

that is the least-squares estimate from the Kalman filter initialised with

$$\Sigma_n(0) = \frac{1}{n} I, \ \hat{x}(0) = 0$$

at time $T \in \mathbb{N}$.

Since there is a solution $(u_0, y_0) \in \mathcal{N}_p^{[0,T]}(w_2)$, hence a solution$(u_0, y_0, 0) \in \mathcal{Z}_p^{[0,T]}(w_2)$ it follows that

$$\|\tilde{x}_0^n\|_{nI}^2 + \|\tilde{u}_0^n\|_2^2 + \|\tilde{y}_0^n\|_2^2 \leq \|u_0\|_2^2 + \|y_0\|_2^2 \tag{A.25}$$

which implies $\|\tilde{x}_0^n\| \to 0$ for $n \to \infty$ (since if not, $\|\tilde{x}_0^n\|_{nI} \to \infty$ as $n \to \infty$ but the right half side of inequality (A.25) is constant), hence $\tilde{x}_0^n \to 0$. By continuity of the Kalman filter equation solutions with respect to the initial conditions $(\Sigma(0), \hat{x}(0)) = (\Sigma_n(0), 0)$ on the interval $[0, T]$ it follows that $(\tilde{u}_0^n, \tilde{y}_0^n) \to (\tilde{u}_0, \tilde{y}_0)$ as $n \to \infty$ where $(\tilde{u}_0, \tilde{y}_0)$ is the solution of the Kalman filter equations on $[0, T]$ with initial condition $(\Sigma(0), \hat{x}(0)) = (0, 0)$. The desired result then follows from Lemma A.7 with $\tilde{y}^n = (\tilde{u}_0^n, \tilde{y}_0^n)$, $y^n = (\tilde{u}_0, \tilde{y}_0)$ $\tilde{x}_n = \tilde{x}_0^n$ and $L = \mathcal{Z}_p^{[0,T]}(w_2)$:

$$r^2(T) = \lim_{n \to \infty} \inf_{(u_0, y_0, x_0) \in \mathcal{Z}_p^{[0,T]}(w_2)} (\|x_0\|^2_{\Sigma_n^{-1}(0)} + \|u_0\|_2^2 + \|y_0\|_2^2)$$

$$= \inf_{(u_0^p, y_0^p) \in \mathcal{N}_p^{[0,T]}(w_2)} (\|u_0^p\|_2^2 + \|y_0^p\|_2^2)$$

as required. □

Bibliography

D. Abramovitch. The outrigger: a prehistoric feedback mechanism. *IEEE Control Systems Magazine*, 25(4):57–72, 8 2005.

B. D. O. Anderson. Failures of adaptive control theory and their resolution. *Communications in Information and Systems*, 5(1):1–20, 2005.

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. *Int. J. of Robust and Nonlinear Control*, 10:909–929, 2000.

B. D. O. Anderson, T. S. Brinsmead, D. Liberzon, and A. S. Morse. Multiple model adaptive control with safe switching. *Int. J. of Adaptive Contr. and Signal Processing*, 15:445–470, 2001.

M. Athans, K. P. Dunn, C. S. Greene, W. H. Lee, N. R. Sandell, I. Segall, and A. S. Willsky. The stochastic control of the f-8c aircraft using the multiple model adaptive control (mmac) method. In *Proc. of 1975 IEEE Conf. on Decision and Control (CDC) in Houston, TX, USA*, volume 14, pages 217–228, 1975.

H. S. Black. Stabilized feed-back amplifiers. *Electrical Engineering*, 53:114–120, 1934.

V. Blondel, M. Gevers, and A. Lindquist. Survey on the state of systems and control. *European J. of Control*, 1:5–23, 1995.

S. Boyd and L. Vandenberghe. *Convex Optimization.* Cambridge University Press, March 2004.

D. Buchstaller and M. French. Scaling of gain bounds for switched adaptive control with large uncertainties. In *Proc. of the 46th IEEE Conf. on Decision and Control (CDC) in New Orleans, LA, USA*, pages 915–920, 2007.

D. Buchstaller and M. French. Gain bounds for multiple model switched adaptive control of general MIMO LTI systems. In *Proc. of the 47th IEEE Conf. on Decision and Control (CDC) in Cancun, Mexico*, pages 5330–5335, 2008.

D. Buchstaller and M. French. Robust stability and performance analysis for multiple model adaptive controllers. In *Proc. of the 48th IEEE Conf. on Decision and Control (CDC) in Shanghai, China*, 2009. Accepted.

M. W. Cantoni and K. Glover. Robustness of linear periodically-time-varying closed-loop systems. In *Proc. of the 37rd IEEE Conf. on Decision and Control (CDC) in Tampa, FL, USA*, volume 4, pages 3807–3812, 1998.

G. W. Chang, J. P. Hespanha, A. S. Morse, M. S. Netto, and R. Ortega. Supervisory field-oriented control of induction motors with uncertain rotor resistance. *Int. J. of Adaptive Contr. and Signal Processing*, 15:353–375, 2001.

J. G. Deshpande, T. N. Upadhyay, and D. G. Lainiotis. Adaptive control of linear stochastic systems. *Automatica*, 9:107–115, 1973.

J. C. Doyle, B. Francis, and A. Tannenbaum. *Feedback Control Theory*. Macmillan Publishing Co., 1990.

J. C. Doyle, J. Wall, and G. Stein. Performance and robustness analysis for structured uncertainty. In *Proc. of the 21th IEEE Conf. on Decision and Control (CDC) in Honolulu, HI, USA*, volume 21, pages 629–636, 1982.

A. K. El-Sakkary. The gap metric: robustness of stabilization of feedback systems. *IEEE Transaction on Automatic Control*, 30(3):240–247, 1985.

S. Fekri, M. Athans, and A. Pascoal. A new robust adaptive control method using multiple-models. In *Proc. of the 12th IEEE Mediterranean Conf. on Control and Automation (MED) in Kusadasi, Turkey*, 2004.

S. Fekri, M. Athans, and A. Pascoal. Issues, progress and new results in robust adaptive control. *Int. J. of Adaptive Contr. and Signal Processing*, 20(10):519–579, 2006.

T. P. Fisher-Jeffes. *Multiple-Model Switching Control to Achieve Asymptotic Robust Performance*. PhD thesis, University of Cambridge, 2003.

J. A. Fitch and P. S. Maybeck. Multiple model adaptive control of a large flexible space structure with purposeful dither for enhanced identifiability. In *Proc. of the 33rd IEEE Conf. on Decision and Control (CDC) in Lake Buena Vista, FL, USA*, volume 3, pages 2904–2909, 1994.

W. H. Fleming. Deterministic nonlinear filtering. *Ann. Scuola Normale Superiore Pisa, Cl. Sci. Fis. Mat.*, 25:435–454, 1997.

M. French. An analytical comparison between the weighted LQ performance of a robust and an adaptive backstepping design. *IEEE Transaction on Automatic Control*, 47 (4):670–675, 2002.

M. French. Smooth adaptive controllers have discontinuous closed loop operators. *Leuven: Sixteenth International Symposium on Mathematical Theory of Networks and Systems (MTNS2004)*, 2004.

M. French. Adaptive control and robustness in the gap metric. *IEEE Transactions on Automatic Control*, 53(2):461–478, 2008.

M. French, A. Ilchmann, and E. P. Ryan. Robustness in the graph topolgy of a common adaptive controller. *SIAM Journal of Control and Optimization*, 45(5):1736–1757, 2006.

M. French and S. Trenn. l^p gain bounds for switched adaptive controllers. In *Proc. of the 44th IEEE Conf. on Decision and Control (CDC) and the European Control Conf. (ECC) in Seville ,Spain*, pages 2865–2870, 2005.

T. T. Georgiou. On the computation of the gap metric. In *Proc. of the 27th IEEE Conf. on Decision and Control (CDC) in Austin, TX, USA*, volume 2, pages 1360–1361, 1988.

T. T. Georgiou and M. C. Smith. Optimal robustness in the gap metric. *IEEE Transactions on Automatic Control*, 35(6):673–686, June 1990.

T. T. Georgiou and M. C. Smith. Robustness analysis of nonlinear feedback systems: An input-output approach. *IEEE Transactions on Automatic Control*, 42(9):1200–1220, September 1997.

T. T. Georgiou and M. C. Smith. Remarks on robustness analysis of nonlinear feedback systems: An input–output approach. *IEEE Transaction on Automatic Control*, 46(1): 171–172, 2001.

C. Guan and S Pan. Nonlinear adaptive robust control of single-rod electro-hydraulic actuator with unknown nonlinear parameters. *IEEE Transaction on Control Systems Technology*, 16(3):434–445, 2008.

W. H He, H. Kaufman, and R. Roy. Multiple model adaptive control procedure for blood pressure control. *IEEE Transaction on Biomedical Engineering*, BME-33(1): 10–19, 1986.

E. Hendricks, O. Jannerup, and P. H. Sorensen. *Linear Systems Control*. Springer, 2009.

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

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. *Int. J. of Robust and Nonlinear Control*, 11:479–496, 2001.

J. P. Hespanha, D. Liberzon, and E. D. Sontag. Nonlinear observability and an invariance principle for switched systems. In *Proc. of the 41st IEEE Conf. on Decision and Control (CDC) in Las Vegas, NV, USA*, 2002.

O. Hijab. *Minimum Energy Estimation.* PhD thesis, Department of Mathematics, University of California, Berkeley, 1980.

L. Hogben, R. A. Brualdi, A. Greenbaum, and R. Mathias. *Handbook of linear algebra.* CRC Press, 2007.

R. A. Horn and C. R. Johnson. *Matrix analysis.* Cambridge University Press, 1990.

N. V. Q. Hung, H. D. Tuan, T. Narikiyo, and P. Apkarian. Adaptive control for nonlinearly parameterized uncertainties in robot manipulators. *IEEE Transactions on Control Systems Technology*, 16(3), 2008.

F. Ikhouane and M. Krstic. Robustness of the tuning functions adaptive backstepping design for linear systems. *IEEE Transaction on Automatic Control*, 43(3):431–437, 1998.

P. Ioannou and J. Sun. *Robust Adaptive Control.* Prentice Hall, 1996.

R. E. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME–Journal of Basic Engineering*, 82(Series D):35–45, 1960.

R. E. Kalman and R. S. Bucy. New results in linear filtering and prediction theory. *Transactions of the ASME–Journal of Basic Engineering*, 83D:95–108, 1961.

A. N. Kolmogorov. On some asymptotic characteristics of completely bounded spaces. *Dokl. Akad. Nauk SSSR,*, 108(3):385–389, 1956.

D. G. Lainiotis. Optimal adaptive estimation: Structure and parameter adaption. *IEEE Transaction on Automatic Control*, 16(2):160–170, 1971.

D. G. Lainiotis. Partitioning: A unifying framework for adaptive systems. Part 1: Estimation. In *Proc. of the IEEE*, volume 64, pages 1126–1143, 1976a.

D. G. Lainiotis. Partitioning: A unifying framework for adaptive systems. Part 2: Control. In *Proc. of the IEEE*, volume 64, pages 1182–1198, 1976b.

T. H. Lee and K. S. Narendra. Stable discrete adaptive control with unknown high-frequency gain. *IEEE Transaction on Automatic Control*, 31(5):477–479, 1986.

W. S. Lee, B. D. O. Anderson, R. L. Kosut, and I. Mareels. A new approach to adaptive robust control. *Int. J. of Adaptive Contr. and Signal Processing*, 7(3):183–211, 1993.

D. Liberzon. *Switching in Systems and Control.* Birkhauser, 2003.

L. Ljung. *System Identification.* Prentice Hall, 1999.

I. Markovsky, J. C. Willems, S. V. Huffel, B. D. Moor, and R. Pintelon. Application of structured total least squares for system identification and model reduction. *IEEE Transactions on Automatic Control*, 50(10):1490–1500, 10 2005.

J. F. Martin, A. M. Schneider, and N. T. Smith. Multiple-model adaptive control of blood pressure using sodium nitroprusside. *IEEE Transaction on Biomedical Engineering*, BME-34(8):10–19, 1987.

J. C. Maxwell. On governors. In *Proceedings of the Royal Society of London*, volume 16, pages 220–283, 1868.

P. S. Maybeck. *Stochastic Models, Estimation and Control, Volume 1*, volume 1. Academic Press, 1979.

P. S. Maybeck. *Stochastic Models, Estimation and Control, Volume 2*, volume 1. Academic Press, 1982a.

P. S. Maybeck. *Stochastic Models, Estimation and Control, Volume 3*, volume 1. Academic Press, 1982b.

P. S. Maybeck and R. D. Stevens. Reconfigurable flight control via multiple model adaptive control methods. In *Proc. of the 29th IEEE Conf. on Decision and Control (CDC) in Honolulu, HI, USA*, volume 6, pages 3351–3356, 1990a.

P. S. Maybeck and R. D. Stevens. Robustness of a moving-bank multiple model adaptive algorithm for control of a flexible space structure. In *Proc. of the 1990 IEEE Aerospace and Electronics Conf. in Dayton, OH, USA*, volume 1, pages 368–374, 1990b.

O. Mayr. *The Origins of Feedback Control.* MIT Press, Cambridge, MA, USA, 1970. ISBN 026213067X.

W. M. McEneaney. Robust H_∞ filtering of nonlinear systems. *System and Control Letters*, 33:315–325, 1998.

A. S. Morse. Supervisory control of families of linear set-point controllers - Part 1: Exact matching. *IEEE Transactions on Automatic Control*, 41(10):1413–1431, October 1996.

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

A. S. Morse. Analysis of a supervised set-point control system containing a compact continuum of finite dimensional linear controllers. *Leuven: Sixteenth International Symposium on Mathematical Theory of Networks and Systems (MTNS2004)*, 2004.

R. E. Mortenson. Maximum likelihood recursive nonlinear filtering. *Journal of Optimization Theory and Applications*, 2:386–394, 1968.

E. Mosca, F. Capecchi, and A. Casavola. Designing predictors for MIMO switching supervisory control,. *Int. J. of Adaptive Contr. and Signal Processing*, 15:265–286, 2001.

D. R. Mudgett and A. S. Morse. Adaptive stabilization of linear systems with unknown high-frequency gains. *IEEE Transaction on Automatic Control*, 30(6):549–554, 1985.

R. Murray-Smith and T. A. Johansen. *Multiple Model Approaches to Modeling Control*. Systems and Control Book Series, Taylor and Francis, 1997.

K. S. Narendra and A. M. Annaswamy. *Stable adaptive systems.* Prentice Hall, 1989.

K. S. Narendra and J. Balakrishnan. Improving transient response of adaptive. control systems using multiple models and switching. In *Proc. of the 32th IEEE Conf. on Decision and Control (CDC) in San Antonio, TX, USA*, 1993.

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

K. S. Narendra, J. Balakrishnan, and M. K. Ciliz. Adaptation and learning using multiple models, switching and tuning. *IEEE Control Systems Magazine*, 15(3):37–51, 1995.

R. D. Nussbaum. Some remarks on a conjecture in parameter adaptive control. *System and Control Letters*, 3:243–246, 1983.

J. W. Polderman and J. C. Willems. *Introduction to Mathematical Systems Theory: A Behavioural Approach*. Springer, 1 edition, 1997.

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

E. J. Routh. *A Treatise on the Stability of a Given State of Motion*. Macmillan & Co, 1877.

W. Rugh. Analytical framework for gain scheduling. *IEEE Control Systems Magazine*, 11(1), 1991.

R. Saeks and J. Murray. Fractional representation, algebraic geometry, and the simultaneous stabilization problem. *IEEE Transaction on Automatic Control*, 27:895–903, 1982.

M. Safonov and T. Tsao. The unfalsified control concept and learning. *IEEE Transaction on Automatic Control*, 42(6):843–847, 1997.

A. S. Sanei and M. French. A performance comparison of robust adaptive controllers: linear systems. *Mathematics of Control, Signals, and Systems*, 18(4):369–394, 2006.

G. N. Saridis and T. K. Dao. A learning approach to the parameter-adaptive self-organizing control problem. *Automatica*, 8:589–597, 1972.

S. Sastry and M. Bodson. *Adaptive Control: Stability, Convergence and Robustness.* Prentice Hall, 1989.

A. Schrijver. *Theory of Linear and Integer Programming.* Wiley, 1998.

E. D. Sontag. *Mathematical Control Theory.* Springer, 1998.

Staff of the Flight Research Center. Experiences with the X-15 adaptive control system. Nasa Technical Note NASA TN D-6208, 1971.

R. F. Stengel. *Stochastic Optimal Control, Theory and Application.* John Wiley and Sons, 1986.

P. Swerling. Modern state estimation methods from the viewpoint of the method of least squares. *IEEE Transaction on Automatic Control*, AC-16(6):707–719, 1971.

M. Vidyasagar. *Control System Synthesis: A Factorization Approach.* MIT Press, 1985.

G. Vinnicombe. *Uncertainty and Feedback.* Imperial College Press, 2000.

G. Vinnicombe. Examples and counterexamples in finite $\mathcal{L}_2$-gain adaptive control. *Leuven: Sixteenth International Symposium on Mathematical Theory of Networks and Systems (MTNS2004)*, 2004.

G. Welch and G. Bishop. An introduction to the Kalman filter. Siggraph 2001, 2001.

N. Wiener. *Extrapolation, Interpolation, and Smoothing of Stationary Time Series.* Wiley, 1949.

J. C. Willems. Deterministic least squares filtering. *Journal of Econometrics*, 118: 341–370, 2004.

J. C. Willems. Note on the deterministic discrete-time Kalman filter. Private Communication, 2006.

D. C. Youla, J. J. Bongiorno Jr., and C. N. Lu. Single-loop feedback stabilization of linear multivariable dynamical plants. *Automatica*, 10:159–173, 1974.

G. Zames. Adaptive control: Towards a complexity-based general theory. *Automatica*, 34(10), 1998.

G. Zames and A. K. El-Sakkary. Unstable systems and feedback: The gap metric. In *Proc. of the Allerton Conf.*, pages 380–385, 1980.