Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

INGEGNERIA BIOMEDICA, ELETTRICA E DEI SISTEMI

Ciclo XXXI

Settore Concorsuale: 09 / G1 - AUTOMATICA

Settore Scientifico Disciplinare: ING-INF / 04 - AUTOMATICA

ADAPTIVE OUTPUT REGULATION FOR MULTIVARIABLE NONLINEAR SYSTEMS VIA HYBRID IDENTIFICATION TECHNIQUES

Presentata da: Michelangelo Bin

Coordinatore Dottorato | **Supervisore**

Prof. Daniele Vigo | **Prof. Lorenzo Marconi**

Esame finale anno 2019

Abstract

Output regulation refers to the class of control problems in which some outputs of the controlled system must be steered to some desired references, while maintaining closed-loop stability and in spite of the presence of unmeasured disturbances and model uncertainties. While for linear systems the problem has been elegantly solved in the 70s, output regulation for nonlinear systems is still a challenging research field, and 30 years of active research left open many fundamental problems. In particular, all the regulators proposed so far are limited to very specific classes of nonlinear systems and, even in those cases, they fail in extending in their full generality the celebrated properties of the linear regulator. The aim of this thesis is to make a decisive step towards the systematic extension of the output regulation theory to embrace more general multivariable problems. To this end, we touch here three fundamental pillars of regulation theory: the structure of regulators, the robustness issue, and the adaptation of the control system. Regarding the structural aspects, we pursue here a design paradigm that is complementary to canonical nonlinear regulators and that trades a conceptually more suitable structure with a strong internal intertwining between the different parts of the regulator. For what concerns robustness, we introduce a new framework to characterize robustness of regulators relative to steady-state properties more general than the usual requirement asking a zero asymptotic error. We characterize in this unifying terms a large part of the existing approaches, and we end conjecturing that general nonlinear regulation

admits no robust solution. Regarding the evolution of regulators, we propose an adaptive regulation framework in which adaptation is used online to tune the internal models embedded in the control system. Adaptation is cast as a general system identification problem, allowing for different well-known algorithms to be used.

Sommario

Il termine "regolazione dall'uscita" (output regulation) si riferisce alla classe di problemi della teoria del controllo in cui ad alcune uscite del sistema controllato devono essere fatte inseguire delle traiettorie di riferimento desiderate, in presenza di incertezze sul modello e disturbi esterni non misurabili e mantendendo la stabilità del sistema complessivo. Nonostante per la classe dei sistemi lineari il problema sia stato elegantemente risolto negli anni 70, la regolazione dall'uscita per sistemi nonlineari rappresenta ancora un campo di ricerca alquanto ostico, in cui oltre trent'anni di ricerca attiva hanno lasciato aperti molti problemi fondamentali. Tutte le soluzioni proposte fin'ora, infatti, si limitano a classi specifiche di sistemi nonlineari, ed anche in tali casi falliscono nell'estendere, nella loro interezza, le rinomate proprietà del regolatore lineare. Lo scopo principale di questa tesi è fare un decisivo passo avanti verso l'estensione sistematica della teoria della regolazione verso classi più generali di sistemi nonlineari, sia dal punto di vista applicativo, sia da quello teorico. A tal fine in questa tesi vengono toccati tre pilastri fondamentali della regolazione: la struttura del regolatore, il problema della robustezza e la questione dell'adattamento ed evoluzione del regolatore. Per quanto riguarda gli aspetti strutturali, viene proposto un paradigma di progetto del regolatore complementare a quelli canonici, che presenta una struttura più consona ad eventuali estensioni della teoria, al prezzo però dell'introduzione di un forte legame tra le varie parti del regolatore, che rende impossibile il loro progetto sequenziale e

separato. Per quanto riguarda la robustezza, viene introdotto un nuovo "framework" in cui è possibile formalizzare e caratterizzare concetti di robustezza legati alle performance dei regolatori relativamente a proprietà asintotiche più generali della condizione canonica richiedente un errore di regolazione nullo a regime. Vengono dunque caratterizzati un questo framework diversi tra i regolatori sviluppati negli ultimi vent'anni, e viene proposta una congettura "negativa" che afferma che nel caso nonlineare generale nessun regolatore è robusto. Per quanto riguarda il progetto di regolatori che si auto-adattano, viene proposto un framework teorico in cui il modello interno presente nel regolatore viene adattato online in autonomia sulla base delle uscite misurabili. Il problema dell'adattamento è posto come un probema di identificazione dinamica, permettendo l'utilizzo di diverse tecniche esistenti.

Acknowledgments

Thanks to my advisor and mentor Lorenzo. Thank you for your constant support and for your very active role in my research. Thank you for your illuminating suggestions and clues, for having continuously proposed new directions, and for having always included me in new collaborations and kept me in the loop. I could not imagine a better mentor.

Thanks to my girlfriend Giulia. Thank you for your unconditional love and support and to have always been by my side, even when I was far away. Thank you for distracting me and for all the lovely moments we have spent in these years. Without you I would have gone crazy a long time ago. I love you.

Thanks to my parents, Chiara and Roberto. Thank you for supporting me without having the slightest idea of what I was doing. Thank you for all the discussions we had and to have convinced me that taking life seriously is not worth it. You guys rock.

Thanks to my lab mates, those that have gone, those that have just arrived, and those that have always been there. Thank you guys for all the funny lunches and for the delirious discussions. CASY is the best.

Contents

Notation

$\mathbb{R}$	set of real numbers
$\mathbb{R}_+$	set of non negative real numbers
$\mathbb{R}_+^*$	set of strictly positive real numbers
$\mathbb{N}$	set of natural numbers zero included
$\mathbb{N}^*$	set of natural numbers zero excluded
$\mathbb{C}$	set of complex numbers
$\mathbb{Z}$	set of integer numbers
$\mathbb{Q}$	set of rational numbers
$\in$	belongs to
$\subset$	subset
$\supset$	superset
$:=$	defined as
$\forall$	for all
$\exists$	there exists
$A \cap B$	intersection of sets
$A \cup B$	union of sets
$A \setminus B$	difference of sets
$A \pm B$	set $\{a \pm b \, : \, a \in A, \, b \in B\}$
αA	with $\alpha \in \mathbb{R}$, set $\{\alpha a \, : \, a \in A\}$
$\emptyset$	the empty set
$A \times B$	Cartesian product of sets
A^n	n-fold product of the set A

$\overline{A}$	closure of A
∂A	boundary of A
$\operatorname{int} A$	interior of A
$\lvert S\rvert$	when S is a set $\lvert S\rvert := \sup_{s\in S} \lvert s\rvert$
$\mathbb{B}$	open ball of radius 1
$\alpha\mathbb{B}$	open ball of radius $\alpha > 0$
$\overline{\mathbb{B}}$	closed ball of radius 1
$\alpha\overline{\mathbb{B}}$	closed ball of radius $\alpha > 0$
$E^{n\times m}$	set of matrices with n rows and m columns and coefficients in E
$\lvert\cdot\rvert$	vector or matrix induced norm
$\lvert\cdot\rvert_A$	$\inf\limits_{a\in A} \lvert\cdot - a\rvert$, distance from the set A
M^T	transpose matrix
M^{-1}	inverse matrix
$M^{\dagger}$	Moore-Penrose generalized inverse matrix
M^{-T}	$(M^{-1})^T$
$M \geq 0$	positive semi-definite matrix
$M > 0$	positive definite matrix
$\det(M)$	determinant of M
$\operatorname{rank}(M)$	rank of M
$\sigma(M)$	spectrum of M
$A \otimes B$	Kronecker product of matrices
$\operatorname{Im} A$	image of A
$\operatorname{Ker} A$	kernel of A
$0_{n\times m}$	matrix of dimension $n \times m$ whose entries are all zeros. When $n = m$ we write 0_n and when the dimension is clear from the context the subscript is omitted and we write simply 0
I_n	n-dimensional identity matrix. When the dimension is clear from the context the subscript is omitted and we write simply I
$\operatorname{diag}(A_1, \ldots, A_n)$	block-diagonal matrix block diagonal elements the square matrices $A_1, \ldots, A_n$

$\mathrm{col}(A_1, \ldots, A_n)$	column concatenation of the elements A_i
$\mathrm{col}(A : A \in \mathcal{A})$	column concatenation of the elements $A \in \mathcal{A}$. If $\mathcal{A}$ is indexed by the set N we also write $\mathrm{col}(A_n : n \in N)$
Hurwitz matrix	matrix with all eigenvalues having strictly negative real part
Schur matrix	matrix with all eigenvalues having modulus strictly lessthan 1
simply stable matrix	matrix with all eigenvalues with zero real part and algebraic multiplicity 1
$\mathbb{HC}(n)$	subset of $\mathbb{R}^n$ of all the coefficients $(c_1, \ldots, c_n)$ of a Hurwitz monic polynomial of dimension n, i.e. such that $p(\lambda) := \lambda^n + c_n\lambda^{n-1} + \cdots + c_2\lambda + c_1$ has only roots with strictly negative real part
(n,m)-prime form	a triplet (A, B, C) in $\mathbb{R}^{nm \times nm} \times \mathbb{R}^{nm \times m} \times \mathbb{R}^{m \times nm}$ defined as $A = \begin{pmatrix} 0_{m(n-1)\times m} & I_{m(n-1)} \\ 0_m & 0_{m\times m(n-1)} \end{pmatrix}$ $B = \begin{pmatrix} 0_{m(n-1)\times m} \\ I_m \end{pmatrix}$ $C = \begin{pmatrix} I_m & 0_{m\times m(n-1)} \end{pmatrix}$ If $m = 1$ we say that (A, B, C) is a triplet in prime form of dimension n
$f : A \to B$	a function from A to B
$f\|_C$	with $f : A \to B$ and $C \subset A$, $f\|_C$ is the restriction of f to C
$F : A \rightrightarrows B$	a set-valued map from A to B
$\mathrm{dom}\, F$	the domain of F
$\mathrm{ran}\, F$	the range of F
$\mathrm{supp}\, F$	the support of f
$f \in \mathcal{K}$	f is a class-K function, i.e. $f : [0, a) \to \mathbb{R}_+$ ($a \in \mathbb{R}_+^*$) is continuous, strictly increasing and $f(0) = 0$
$f \in \mathcal{K}_\infty$	f is a class-K_∞ function, i.e. $f \in \mathcal{K}$ and $f(x) \to_{x\to a} \infty$
$f \in \mathcal{L}$	f is a class-L function, i.e. $f : \mathbb{R}_+ \to \mathbb{R}_+$ is continuous, strictly decreasing and $f(x) \to_{x\to\infty} = 0$

$\beta \in \mathcal{KL}$	β is a class-KL function, i.e. $\beta(\cdot, t) \in \mathcal{K}$ for each t and $\beta(s, \cdot) \in \mathcal{L}$ for each s
$\beta \in \mathcal{KLL}$	β is a class-KLL function, i.e. $\beta(\cdot, \cdot, t) \in \mathcal{KL}$ for each t and $\beta(s_1, s_2, \cdot) \in \mathcal{L}$ for each s_1, s_2
C^n	The set of n-time continuously differentiable functions (C^0 is the set of continuous functions)
L_p	Lebesgue space of functions for which the p-th power of the absolute value is Lebesgue integrable
$D^+V(t)$	If $V : \mathbb{R} \to \mathbb{R}$, $D^+V(t)$ denotes the Dini derivative $$D^+V(t) := \limsup_{h \to 0^+} \frac{1}{h}\big(V(t+h) - V(t)\big).$$ If V is obtained by evaluating a function $V : \mathbb{R}^n \to \mathbb{R}$ along a trajectory $x(t, x_0)$, by extension we let $$D^+V(x_0) := \limsup_{h \to 0^+} \frac{1}{h}\big(V(x(t+h, x_0)) - V(x(t, x_0))\big)$$
$L_{f(x)}h(x)$	Lie derivative of $h(x)$ along the vector field $f(x)$, i.e. $L_{f(x)}h(x) := \dfrac{\partial h(x)}{\partial x} f(x)$. If no ambiguity is present we omit the argument of f, thus writing $L_f h(x)$
$L_f^n h(x)$	n-fold Lie derivative of $h(x)$ along the vector field $f(x)$, i.e. $L_f^n h(x) := L_f(L_f^{n-1} h(x))$, with $L_f^0 h(x) = L_f h(x)$
$L_{g_i}^{(x_i)} M(x_1, \ldots, x_k)$	With $M : \mathbb{R}^{n_1 + \cdots + n_k} \to \mathbb{R}^{m \times p}$, with $n_i, m, p \in \mathbb{N}$, is a matrix-valued function and, for $i = 1, \ldots, k$, $g_i : \mathbb{R}^{n_1 + \cdots + n_k} \to \mathbb{R}^{n_i}$ is a vector field, $L_{g_i}^{(x_i)} M(x_1, \ldots, x_k)$ denotes the matrix whose (ℓ, j)-th element is given by: $$\frac{\partial M_{\ell j}(x_1, \ldots, x_k)}{\partial x_i} g_i(x_1, \ldots, x_k).$$ When there is no ambiguity, we just write $L_{g_i} M(x_1, \ldots, x_k)$. If $k = 1$, $n_1 = n \in \mathbb{N}$ and $m = p = 1$, then, with $x = x_1$, $f = M$ and $g = g_1$ we have $L_{g_1}^{(x_1)} M(x_1) = L_g f(x)$

$\lvert x\rvert_{[t_1,t_2)}$	when $x(\cdot)$ is a locally essentially bounded function defined on $\mathbb{R}$, for $t_1, t_2 \in \mathbb{R}$ we let $\lvert x\rvert_{[t_1,t_2)} := \operatorname{ess.\,sup}_{t\in[t_1,t_2)} \lvert x(t)\rvert$
$\lvert x\rvert_\infty$	when $x(\cdot)$ is a locally essentially bounded function defined on $\mathbb{R}$, we let $\lvert x\rvert_{[t_1,t_2]} := \operatorname{ess.\,sup}_{t\in\mathbb{R}} \lvert x(t)\rvert$
$x_{[i,j]}$	when $i, j \in \mathbb{N}$, $i \le j$, and $x \in \mathbb{R}^n$, $n \ge j$, we let $x_{[i,j]} = (x_i, x_{i+1}, \dots, x_j)$
$x^{(i,j)}$	when $i, j \in \mathbb{N}$, $i \le j$, and x is a j-times differentiable function on $\mathbb{R}$, we let $x^{(i,j)} := (x^{(i)}, x^{(i+1)}, \dots, x^{(j)})$
$(x_n)_{n=a}^b$	sequence $x_a, \dots, x_b$
$(x_n)_n$	a sequence of elements x_n indexed by $n \in \mathbb{N}$
$x_n \to x$	short for $\lim_{n\to\infty} x_n = x$
$\limsup x$	if $x : \mathbb{R} \to \mathbb{R}^n$, short for $\limsup_{t\to\infty} x(t)$
$\mathcal{S}_\bigstar(X)$	when $\bigstar$ denotes a system and X a set, $\mathcal{S}_\bigstar(X)$ is the set of all the solutions to $\bigstar$ originating in X. If X is a singleton, i.e. $X = \{x\}$, we write $\mathcal{S}_\bigstar(x)$ in place of $\mathcal{S}_\bigstar(\{x\})$. If $\bigstar$ is a system with state x, we will also use x in place of $\bigstar$ if no confusion arises, i.e. we let $\mathcal{S}_\bigstar = \mathcal{S}_x$
$\mathcal{R}^\tau_\bigstar(X)$, $\Omega_\bigstar(X)$	when $\bigstar$ denotes a system, $X \subset \mathbb{R}^n$ a set, and $\tau > 0$, $\mathcal{R}^\tau_\bigstar(X)$ is the τ-reachable set of $\bigstar$ from X, i.e. $$\mathcal{R}^\tau_\bigstar(X) := \{x \in \mathbb{R}^n \,:\, x = \varphi(t),\, \varphi \in \mathcal{S}_\bigstar(X),\, t \ge \tau\}.$$ As clearly $\mathcal{R}^\tau_\bigstar(X)$ is decreasing in τ (in the sense of inclusion), then we can define the set $$\Omega_\bigstar(X) := \lim_{\tau\to\infty} \mathcal{R}^\tau_\bigstar(X) = \bigcap_{\tau>0} \mathcal{R}^\tau_\bigstar(X),$$ which is called the Ω-limit set. If X is a singleton, i.e. $X = \{x\}$, we write $\mathcal{R}_\bigstar(x)$ and $\Omega_\bigstar(x)$ in place of $\mathcal{R}_\bigstar(\{x\})$ and $\Omega_\bigstar(\{x\})$. If $\bigstar$ is a system with state x, we will also use x in place of $\bigstar$ if no confusion arises, i.e. we let $\mathcal{R}_\bigstar = \mathcal{R}_x$ and $\Omega_\bigstar = \Omega_x$

Introduction

IN 1857, a 21 years old Mark Twain was beginning his training as a steamboat pilot on the Mississippi river, under the command of Mr. Horace Ezra Bixby. Several years later, while writing in the memoir *"Life on the Mississippi"* (Twain, 1883) about how discouraging was to realize how much he had to learn, he reported his mentor as saying:

> *"You only learn the shape of the river, and you learn it with such absolute certainty that you can always steer by the shape that's in your head, and never mind the one that's before your eyes."*

Mr. Bixby was arguing that any good pilot needs to perfectly know the shape of the river, to avoid being fooled by darkness, mist or moonlight shadows during nocturnal navigation. In its essence, Mr. Bixby's intuition hides the *Internal Model Principle*, informally stating that *an internal representation of the "outside world", of the task being executed, and of the agent itself is necessary for a smooth and robust operation*. Not surprisingly, the concept of "internal model" and the related principles pervade many fields of science (Huang et al., 2018), ranging from biology (Sontag, 2003), cognitive science (Grush, 2004) and neuroscience (Wolpert et al., 1998) to control theory and robotics (Isidori, 2017), and they are intimately related with the concept of "knowledge" that lies at the base of any adapting and learning process.

Many studies in neuroscience support the idea that the behaviors of animals and humans are regulated by internal models, refined day after day in the recurrent execution of "similar" tasks in "familiar" environments (Miall et al., 1993; Wolpert et al., 1998; Schubotz, 2007). *Sensorimotor integration*, the transformation of sensory stimuli into motor actions, is perhaps the most well studied function of the nervous system with an internal model-based perspective (Miall et al., 2000; Kawato et al., 2003; Jeannerod, 2006; Schubotz, 2007). Advanced motor gestures, such as hitting a baseball or skiing on moguls, require an exquisite spatio-temporal precision, that is simply not achievable by just sequencing fast "reflexive" corrections (i.e. by pure *feedback* control) due to sensorimotor delays and noise and limited resolution of the sensory apparatus (Wolpert et al., 1998). A combination of sensory-driven feedback and predictive *"feedforward"* actions, incorporating internal models of the environment, of the task, and of the sensorimotor dynamics, is what makes all human-interest operations possible. The internal model principle, moreover, is not confined to the sensorimotor domain, yet it is thought as unifying concept to study higher cognitive and social abilities, including for instance planning, reasoning, imitation and cooperation (Frith et al., 2000; Grush, 2004; Schubotz, 2007).

The branch of control theory that developed around the concept of internal model is known as *"Output Regulation"* (Isidori, 2017), and how internal models can be constructed, adapted and exploited, in the formal context of nonlinear output regulation is, in a nutshell, the subject of this thesis. The thesis is divided in three parts: the first part, subdivided in three chapters, is dedicated to the theory of output regulation; the second part, subdivided in two chapters, concerns the relation between system identification and control, with an accent to observer design; the last part, subdivided in two chapters, is dedicated to the design of adaptive regulators.

The first chapter of the thesis aims at introducing the reader to the current state of the art of output regulation, with an accent on the approaches that have influenced more strongly our work. The second chapter focuses on the limits of classical approaches and on the structural problems that necessarily arise in the design of a regulator when we switch from linear to nonlinear systems. In particular, we point out as a *"chicken-egg dilemma"* arises in the solution of general nonlinear regulation problems, stating that if we insist in separating a regulator in an "internal model unit" and a "stabilizer" there is no way, in general, to tell which one has to be fixed first. We thus reinterpret previous approaches as attempts to avoid dealing with the dilemma, trading feasibility with a consequent loss of generality. In the last part of the chapter we give conditions under which a so-called "post-processing" regulator exists that can deal with the chicken-egg dilemma for classes of systems going beyond those treatable by the current literature. The third chapter concerns instead the robustness issue. We analyze the necessary conditions (such as the internal model principle) of the existence of a regulator in case the exosystem is described by a differential inclusion, thus extending the "non-equilibrium theory" of (Byrnes and Isidori, 2003). We then propose a nonlinear regulator based on low-power high-gain observers and on immersion arguments that can guarantee a certain degree of robustness for particular classes of nonlinear problems, extending the "structural robustness" framework of (Byrnes et al., 1997a). Lastly, we build a new formal framework in which robustness of regulators can be characterized in topological terms and relatively to arbitrary steady-state properties. We re-frame in this context many well-known regulators and we point out how robustness of asymptotic regulation is, in a general nonlinear case, idealistic.

The second part of the thesis is dedicated in presenting our approach to adaptation, with applications that, for simplicity, are directed towards the adaptive observation theory. Adaptation and learning are approached as *system identification* problems, in a deterministic setting tailored on control. To fix a common playground in which control and identification can coexist, in the first chapter of this second part we propose a framework where to reformulate the "recursive" system identification problems in system theoretical terms, with a consequent characterization of the identification algorithms in terms of stability theory. In particular, fundamental notions and requirements are defined and then proved to hold for some relevant classes of identification schemes, such as continuous and discrete least-squares, nonlinear "mini-batch" algorithms and recursive wavelet expansions. Emphasis is put on "universal approximators" and, by leveraging Wavelet theory, on the *multiresolution* aspect of learning: coarse traits represent solid knowledge with slow learning dynamics, while details are more volatile and subject to quicker change. What is interesting in multiresolution itself, is that it also impacts in terms of *analogy and generalization* as, roughly, *the same "coarse" skills that allow Mark Twain to drive a boat on a river will be useful to track a hiking trail*. In the second chapter, the theory is applied to high-gain observers, where the problem of observing an unknown nonlinear system in canonical form is taken on in the proposed framework.

The third part is dedicated to the design of adaptive solutions to output regulation problems. Here the results of control theory and identification presented in the previous chapters merge into control systems that learn, adapt and exploit an internal model of the world to achieve at best the regulation goal. In the first chapter of this third part a general framework is proposed to deal with nonlinear continuous-time regulation problems, and the concept of *"class-type" internal models* is introduced to address the structural issues necessarily concealing behind the design of nonlinear internal models and feedback control. Some cases studies that cover relevant classes of systems are then presented, showing how the proposed framework can embrace state-of-art regulation problems and more. The second chapter of this last part concerns, instead, a different approach to output regulation for linear systems, where discrete-time identifiers are used on top of a continuous-time internal model. In this framework the problem of general multivariable adaptive linear regulation is solved, with a design that, up to the author's knowledge, is the first boasting such level of generality.

Overall, the approach to adaptive regulation pursued in this thesis is based on an exquisite mixture of control and identification, all framed in the formal framework of control theory and where the key to success has to be sought in the synergistic design of the different components.

Part I

Output Regulation

1

Output Regulation of Linear and Nonlinear Systems

In control theory the world is described in terms of systems of differential equations, whose solutions model the time evolution of the different phenomena that take place. The systems that populate the world are defined by an internal *state*, containing all the information sufficient to describe the system. The time evolution of the state may be affected by system's *inputs* and may be observed by the rest of the world throughout system's *outputs*. Output regulation is the branch of control theory that studies how, given a system, some of its inputs (the *control inputs*) can be chosen to make some of its outputs (the *regulated outputs*) to follow given *reference* behaviors, despite the presence of exogenous antagonist inputs (the *disturbances*) affecting the system and without the perfect knowledge of the controlled system itself. The act of making the regulated outputs to follow the reference behaviors is called *tracking*, eliminating the effect of disturbances from the regulated outputs is called *disturbance rejection*. Output regulation refers thus to the simultaneous ability of tracking references while rejecting disturbances. the ability to achieve that goal *"without the perfect knowledge of the controlled system"* is a fundamental property (real-world systems cannot be known perfectly), which is referred to as *robustness*.

In line with the control theoretical vision of the world, the most interesting class of output regulation problems is those where the references and the disturbances are generate by an autonomous (i.e. without inputs) system, which is generally referred to as the *exosystem*. The exosystem thus describes the *structure* of the outside world the controlled system interacts with. Locating the exact birth of output regulation is not an easy task, though the first significant example can be attributed to the famous PID (Proportional Integral Derivative) controller, developed in the 30s and celebrated for its robustness property. The PID is currently used to cope with *constant* references and disturbances, and its robustness is a consequence of the fact that the integral term embeds an *internal model* of the process that generates all the possible constant signals[1], that is, an integrator. It can be shown, indeed, that when the PID is applied, while the proportional and derivative terms vanish with the regulation error, the integral term converges to the ideal constant input (called the *error-zeroing input*) that makes invariant the set in which perfect tracking takes place, and this convergence is not affected by the plant's parameters as far as stability is not broken (and this is, in essence, robustness). This fact knew a rigorous generalization to arbitrary linear controlled systems and exosystems (thus to arbitrary references and disturbances described by finite combinations of harmonics) in the mid 70s, in the seminal works of Francis and Wonham (Francis and Wonham, 1975, 1976; Francis, 1977). This result, known under the name of *Internal Model Principle*, is one of the most popular principles in control theory, and informally states that *every linear regulator that solves the problem of linear output regulation robustly, necessarily includes an internal model of the system that generates the ideal error-zeroing input*.

Interesting enough, for linear systems the aforementioned system that produces the error-zeroing input coincides with the exosystem (or, more precisely,

[1] As a matter of fact, if we define the system

$$\dot{\eta}(t) = e(t),$$

where $e(t)$ denotes the regulation error, then the integral term of the PID can be written as $k_I\eta(t)$, for some $k_I \in \mathbb{R}$. When perfect tracking holds (i.e. $e(t) = 0$), then η fulfills $\dot{\eta}(t) = 0$, whose solutions are constants ranging in the whole $\mathbb{R}$, as so does $\eta(0)$.

with its largest cyclic component). Although incredibly powerful and elegant, this perfect matching between exosystem and internal model inexorably breaks down as soon as nonlinearities come into play. As we will see in the next chapters, the knowledge of the exosystem is far to be sufficient to fit into a nonlinear internal model principle even for very simple systems. The goal of this chapter is to introduce the reader to the state of the art of output regulation of linear and nonlinear systems, presenting the main results and control designs currently available in the literature.

1.1 Output Regulation of Linear Systems

This section is dedicated to a brief recap of the output regulation for linear systems, as it is instrumental to understand and interpret the nonlinear framework. For brevity, we give here a simple reinterpretation of the main results of Francis, Wonham and Davison (Francis and Wonham, 1975, 1976; Francis, 1977; Davison, 1976), sometimes sacrificing rigor and generality to underline and magnify the features of interest.

1.1.1 The Steady State of a Linear System

As a preliminary step we present here a characterization of the concept of *steady state* for linear systems. We consider a cascade of the form (we omit here and everywhere else the time dependency when not strictly necessary)

$$\begin{aligned} \dot{w} &= Sw \\ \dot{z} &= Fz + \Sigma w, \end{aligned} \tag{1.1}$$

of the system w onto the system z, where w takes values in $\mathbb{R}^{n_w}$, $z \in \mathbb{R}^{n_z}$, and where S, F and Σ are matrices of appropriate dimensions with elements in $\mathbb{R}$. The cascade (1.1) is characterized by the following proposition.

Proposition 1.1. *If and only if $\sigma(S) \cap \sigma(F) = \emptyset$, there exists $\Pi \in \mathbb{R}^{n_z \times n_w}$ such that the set*

$$\operatorname{graph} \Pi = \left\{ (w, z) \in \mathbb{R}^{n_w + n_z} \,:\, z = \Pi w \right\} = \operatorname{Im} \begin{pmatrix} I_{n_w} \\ \Pi \end{pmatrix} \tag{1.2}$$

is forward invariant[2] *for* (1.1).

Proof. If and only if $\sigma(S) \cap \sigma(F) = \emptyset$, the Sylvester equation

$$\Pi S - F\Pi = \Sigma, \tag{1.3}$$

admits a unique solution $\Pi \in \mathbb{R}^{n_z \times n_w}$. Let Π in (1.2) be the solution to (1.3). To prove sufficiency, let $v \in \operatorname{Im} \operatorname{col}(I_{n_w}, \Pi)$. Then, for some $w \in \mathbb{R}^{n_w}$, $v = \operatorname{col}(w, \Pi w)$, and using (1.3) yields

$$\begin{pmatrix} S & 0 \\ \Sigma & F \end{pmatrix} v = \begin{pmatrix} Sw \\ \Sigma w + F\Pi w \end{pmatrix} = \begin{pmatrix} Sw \\ \Pi S w \end{pmatrix} = \begin{pmatrix} I_{n_w} \\ \Pi \end{pmatrix} Sw,$$

which, for the arbitrariness of v, proves that

$$\begin{pmatrix} S & 0 \\ \Sigma & F \end{pmatrix} \begin{pmatrix} I_{n_w} \\ \Pi \end{pmatrix} \subset \operatorname{Im} \begin{pmatrix} I_{n_w} \\ \Pi \end{pmatrix}.$$

Hence, forward invariance of (1.2) follows by (Basile and Marro, 1992, Thm. 3.2.4). To prove necessity, suppose that $\operatorname{graph} \Pi$ is forward invariant, and pick an initial condition $(w_0, z_0) \in \operatorname{graph} \Pi$. Then $z_0 = \Pi w_0$, and the unique solution (w, z) to (1.1) originating at (w_0, z_0) necessarily satisfies $z(t) = \Pi w(t)$. Hence, z also fulfills

$$\dot{z} = \Pi \dot{w} = \Pi S w$$

and

$$\dot{z} = Fz + \Sigma w = (F\Pi + \Sigma) w.$$

From the arbitrariness of w, the latter equations yield (1.3), which implies $\sigma(S) \cap \sigma(F) = \emptyset$. ■

The result of Proposition 1.1 is particularly interesting when S is simply stable and F Hurwitz (see the notation section for the terminology). As a matter of fact, if we define the error variable

$$\tilde{z} := z - \Pi w,$$

[2]See Section A.2.

then, in view of (1.3), we obtain

$$\dot{\tilde{z}} = Fz + \Sigma w - \Pi S w = F\tilde{z} + (F\Pi + \Sigma - \Pi S)w = F\tilde{z},$$

i.e. the set (1.2) is also globally exponentially stable. As a consequence, asymptotically the state $z(t)$ will approach the signal $\Pi w(t)$, and we write, symbolically,

$$z(t) \to \Pi w(t). \tag{1.4}$$

If S is simply stable, then $w(t)$ is a linear combination of a finite number of harmonics and Π always exists since $\sigma(S) \cap \sigma(F) = \emptyset$ trivially holds. In this case the interpretation of (1.4) thus coincides with the usual notion of the *steady state* of a linear system: *a stable linear system excited by a linear combination of harmonics, asymptotically oscillates as a linear combination of the same harmonics*.

The function $t \mapsto \Pi w(t)$ represents, indeed, the *forced response* of the linear system z and, in this respect, it is worth noting that Π might exist as well also if F is not Hurwitz, as in view of Proposition 1.1 it suffices to have $\sigma(S) \cap \sigma(F) = \emptyset$ (thus, in particular, if F has all the eigenvalues with positive real part, Π will always exist). This means that, if (1.1) is properly initialized in the set (1.2), the z subsystem can have non-trivial bounded trajectories even if arbitrarily unstable.

Finally, we also underline how the necessity of the condition $\sigma(S) \cap \sigma(F) = \emptyset$ for the existence of a steady state and for its attractiveness has the nice interpretation of a *non-resonance* condition: if S and F share some eigenvalue, then, even if both matrices are stable, there exist solutions that oscillate with an amplitude increasing with time, and such trajectories explode to infinity, thus violating invariance of $\operatorname{graph} \Pi$.

1.1.2 The Internal Model Principle

We present in this section a slightly informal adaptation of the linear internal model principle. For further details, and a full rigorous treatise, the reader is referred to (Francis and Wonham, 1975, 1976; Francis, 1977).

In this section we will consider a plant[3] described by linear equations of the

[3] We will often use the historical term "plant" to refer to the controlled system.

form

$$\begin{array}{rcl}\dot{x} & = & Ax + Bu + Pw \\ y & = & Cx + Qw\end{array} \tag{1.5}$$

with state $x \in \mathbb{R}^n$, control input $u \in \mathbb{R}^m$, measured output $y \in \mathbb{R}^{n_y}$ and with $w \in \mathbb{R}^{n_w}$ that represent the exogenous signals acting on the system, such as references and disturbances. The matrices A, B, P, C, Q are real matrices of appropriate dimension. We associate to (1.5) the *regulation errors*

$$e := C_e x + Q_e w \ \in \mathbb{R}^{n_e} \tag{1.6}$$

defined as the difference of the regulated outputs $C_e x$ and the references $-Q_e w$, being $C_e \in \mathbb{R}^{n_e \times n}$ and $Q_e \in \mathbb{R}^{n_e \times n_w}$ such that $\operatorname{Im} Q_e \subseteq \operatorname{Im} C_e$. Finally, we make the assumption that the exogenous input w is generated by a linear exosystem of the form

$$\dot{w} = Sw. \tag{1.7}$$

In this framework, the (linear) output regulation problem reads as follows: find a regulator of the form

$$\begin{array}{rcl}\dot{x}_c & = & A_c x_c + H_c y \\ u & = & K_c x_c + K_y y,\end{array} \tag{1.8}$$

with $x_c \in \mathbb{R}^{n_c}$ for some $n_c \in \mathbb{N}$ and A_c, H_c, K_c, K_y matrices of appropriate dimension, such that the closed-loop system (1.5), (1.7), (1.8) satisfies:

1. The origin of the subsystem (x, x_c) with $w = 0$ is asymptotically stable.

2. Each solution to the closed-loop system (with w any solution to (1.7)) satisfies

$$\lim_{t \to \infty} e(t) = 0. \tag{1.9}$$

Let $z := (x, x_c)$ and $n_z := n + n_c$. Then, by defining

$$F := \begin{pmatrix} A + BK_y C & BK_c \\ H_c C & A_c \end{pmatrix} \qquad \Sigma := \begin{pmatrix} P + BK_y Q \\ H_c Q \end{pmatrix}$$

we obtain a cascade of the form (1.1). Let us assume that S is simply stable, and suppose that the regulator x_c solves the output regulation problem. Then condition 1 implies that F is Hurwitz, so as there exits $\Pi \in \mathbb{R}^{n_z \times n_w}$ satisfying the Sylvester equation (1.3), whose graph is asymptotically stable for the closed-loop system (w,z). Let us partition Π as $\Pi = \mathrm{col}(\Pi_x, \Pi_c)$, with $\Pi_x \in \mathbb{R}^{n\times n_w}$ and $\Pi_c \in \mathbb{R}^{n_c\times n_w}$. The Sylvester equation (1.3) implies that, by letting $\Gamma := K_c\Pi_c + K_yC\Pi_x + K_yQ$, then necessarily

$$\Pi_x S = A\Pi_x + B\Gamma + P \tag{1.10}$$

holds. Condition 2, namely $e(t) \to 0$, also implies that

$$C_e\Pi_x + Q_e = 0. \tag{1.11}$$

Equations (1.10)-(1.11) are called the *regulator equations*, and what said until now can be rephrased as: *if x_c solves the problem of output regulation then necessarily there exists* $(\Pi_x, \Gamma) \in \mathbb{R}^{n_x\times n_w} \times \mathbb{R}^{m\times n_w}$ *solving the regulator equations* (1.10)-(1.11).

Now, the signal $\Gamma w(t)$ has to be interpreted as the *ideal error-zeroing input*, i.e. the feedforward action that makes the set in which e vanishes invariant. By definition of Γ, and by invariance of $\mathrm{graph}\,\Pi$, if the regulator x_c solves the output regulation problem, then necessarily it must be able to generate all the possible outputs $u^\star$ of the system

$$\begin{array}{rcl} \dot{w} &=& Sw \\ u^\star &=& \Gamma w. \end{array} \tag{1.12}$$

Although from the definition of Γ we have that, in principle, $u^\star$ could be generated by only using the static component $K_y y$, this property would be lost at front of any slight variation of any of the plant's matrices from the nominal value used to tune K_y (for further detail the reader is referred to (Francis, 1977)). Thus, if a *robust* design is sought (i.e. a design that is still valid if some of the matrices slightly deviate from the nominal value), then necessarily $u^\star$ must be given by the term $K_c x_c$, and $K_y y$ is rather to be compensated. As a consequence, the regulator x_c must embed a subsystem that generates all the solutions to (1.12), and this property is essentially what is known as the *internal model principle*.

1.1.3 The Linear Regulator

We present here a re-adaptation of the linear regulator originally proposed by Davison in (Davison, 1976). The reader is referred to that paper for further details. We consider the same class of systems (1.5), (1.6), (1.7), and from now on we will assume the following:

Assumption 1.1. *The following holds:*

1. *S is simply stable.*

2. *The pair (A, B) is stabilizable, the pair (C, A) is detectable and the following* non-resonance condition *holds:*

$$\operatorname{rank}\begin{pmatrix} A-\lambda I & B \\ C_e & 0 \end{pmatrix} = n + n_x, \qquad \forall \lambda \in \sigma(S).$$

3. *The regulation error e belongs to the measured outputs y, i.e. we can write*

$$C = \begin{pmatrix} C_e \\ C_a \end{pmatrix}, \qquad Q = \begin{pmatrix} Q_e \\ Q_a \end{pmatrix}$$

   for some $C_a \in \mathbb{R}^{(n_y-n_e)\times n}$ and $Q_a \in \mathbb{R}^{(n_y-n_e)\times n}$.

Remark 1.1. The assumption of S being stable is not necessary, unless boundedness of the closed-loop trajectories is explicitly included in the problem statement. As a matter of fact, the linear regulation theory would equally work also if S contains unstable modes, with the constraint (1.9) possibly implying the existence of unbounded closed-loop trajectories. Nevertheless, to be consistent with the nonlinear regulation theory, and with the forthcoming adaptive results, we decided to keep boundedness of $w(t)$ as a standing assumption from the beginning. △

We augment the plant with the system

$$\dot{\eta} = \Phi\eta + Ge, \tag{1.13}$$

with state $\eta \in \mathbb{R}^{n_\eta}$, being $n_\eta := n_e n_w$, and with

$$\Phi := \begin{pmatrix} 0 & I_{n_e} & 0 & 0 & \cdots & 0 \\ 0 & 0 & I_{n_e} & 0 & \cdots & 0 \\ \vdots & & & \ddots & & \vdots \\ 0 & & & & & I_{n_e} \\ -c_0 I_{n_e} & -c_1 I_{n_e} & & \cdots & & -c_{n_w-1} I_{n_e} \end{pmatrix} \qquad G := \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ I_{n_e} \end{pmatrix}$$

where all the 0 block are of dimension n_e and the coefficients c_i are such that the characteristic polynomial of S reads as

$$p_S(\lambda) = \lambda^{n_w} + c_{n_w-1}\lambda^{n_w-1} + \cdots + c_1\lambda + c_0. \tag{1.14}$$

Remark 1.2. We observe that equation (1.14), expressing the fact that the characteristic polynomial of the exosystem matrix S coincides with those of Φ, is what confer the *internal model property* on the system (1.13). In fact, (1.14) implies that any mode of the exosystem is also a mode of the system (1.13) when $e = 0$. △

We refer to the system η as the *internal model unit*, as up to a change of coordinates, and with $e = 0$, it coincides with n_e copies of the exosystem (1.7).

Directly from point 2 of Assumption 1.1 it follows that the cascade (x, η) is stabilizable. We then define a second system whose role is to stabilize the cascade (x, η), when $w = 0$. For, we let $n_\xi \in \mathbb{N}$ and we define the system

$$\begin{aligned} \dot{\xi} &= A_\xi \xi + H_\eta \eta + H_y y \\ u &= K_\xi \xi + K_\eta \eta + K_y y, \end{aligned} \tag{1.15}$$

with state $\xi \in \mathbb{R}^{n_\xi}$, and where $A_\xi, H_\eta, H_y, K_\xi, K_\eta, K_y$ are real matrices of appropriate dimension such that the matrix

$$F := \begin{pmatrix} A + BK_yC & BK_\eta & BK_\xi \\ GC_e & \Phi & 0_{n_\eta \times n_\xi} \\ H_yC & H_\eta & A_\xi \end{pmatrix} \tag{1.16}$$

is Hurwitz.

The regulator $x_c := (\eta, \xi)$, depicted in Figure 1.1, has the form (1.8), and the following proposition shows that it solves the output regulation problem for the plant (1.5).

Proposition 1.2. *Let Assumption 1.1 be fulfilled, then the regulator* (1.13), (1.15) *solves the output regulation problem relative to the system* (1.5), (1.6), (1.7).

Proof. By letting $z := (x, \eta, \xi)$ and $n_z := n + n_\eta + n_\xi$, the closed-loop system has the form (1.1), with F given as in (1.16) and $\Sigma = \mathrm{col}(P + BK_yQ,\ GQ_e,\ H_yQ)$. As F is Hurwitz and S is simply stable, Proposition 1.1 yields the existence of a matrix $\Pi \in \mathbb{R}^{n_z \times n_w}$ such that $\mathrm{graph}\,\Pi$ is forward invariant for (w, z). Let partition

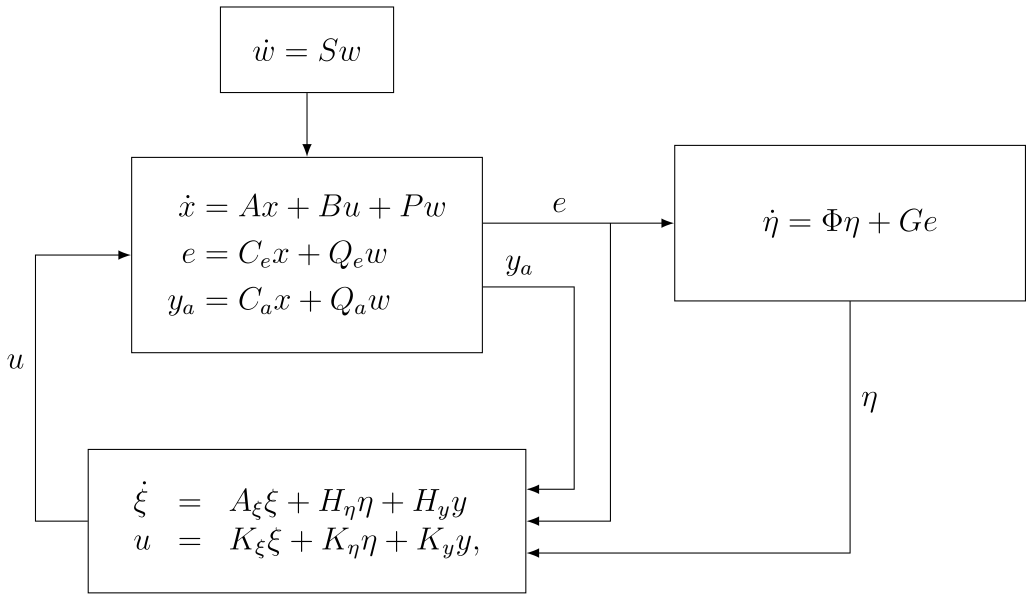

Figure 1.1: Block-diagram of the closed-loop system.

Π as $\Pi = \mathrm{col}(\Pi_x, \Pi_\eta, \Pi_\xi)$, with $\Pi_x \in \mathbb{R}^{n \times n_w}$, $\Pi_\eta \in \mathbb{R}^{n_\eta \times n_w}$ and $\Pi_\xi \in \mathbb{R}^{n_\xi \times n_w}$, and let $\Pi_e := C_e \Pi_x + Q_e$. Af F is Hurwitz, $\mathrm{graph}\,\Pi$ is also globally attractive and, hence, it suffices to show that $\Pi_e = 0$. Equation (1.3), in particular, gives

$$\Pi_\eta S = \Phi \Pi_\eta + G \Pi_e.$$

Let us further partition Π_η as $\Pi_\eta = \mathrm{col}(\Pi_{\eta_1}, \Pi_{\eta_2}, \dots, \Pi_{\eta_{n_w}})$, with the matrices Π_{η_i} of appropriate dimension. From the structure of Φ, we obtain

$$\Pi_{\eta_i} S = \Pi_{\eta_{i+1}}, \qquad \forall i = 1, \dots, n_w - 1, \tag{1.17}$$

and

$$\Pi_{\eta_{n_w}} S = \sum_{i=1}^{n_w} c_{i-1} \Pi_{\eta_i} + \Pi_e. \tag{1.18}$$

Further developing (1.17) yields

$$\Pi_{\eta_i} = \Pi_{\eta_1} S^{i-1}, \qquad \forall i = 2, \dots, n_w,$$

and using (1.18) we obtain

$$\Pi_{\eta_1} S^{n_w} = \Pi_{\eta_{n_w}} S = \sum_{i=1}^{n_w} c_{i-1} \Pi_{\eta_1} S^{i-1} + \Pi_e = \Pi_{\eta_1} \left(\sum_{i=1}^{n_w} c_{i-1} S^{i-1} \right) + \Pi_e.$$

Solving for Π_e yields

$$\Pi_e = \Pi_{\eta_1} \left(S^{n_w} + c_{n_w-1} S^{n_w-1} + \cdots + c_1 S + c_0 I_{n_w} \right). \tag{1.19}$$

In view of (1.14) and the Cayley-Hamilton Theorem, (1.19) implies

$$\Pi_e = 0,$$

and the claim follows. ∎

We close the section, by briefly recalling the main properties of the linear regulator (1.13), (1.15):

P1 **Robustness:** the linear regulator is *structurally robust*, namely output regulation is achieved with the *same* regulator for any choice of P and Q and for any perturbation of A, B and C that does not destroy closed-loop stability and linearity. In particular, if A, B, and C are subject to parameter uncertainties, then if closed-loop stability holds for a *nominal* triple (A, B, C), it also holds for sufficiently small perturbations of it. Moreover, we observe that P, Q, the solutions (Π_x, Γ) of the corresponding regulator equations (1.10)-(1.11), and the steady-state matrix Π, in general play no role in the regulator synthesis, and the internal model unit depends exclusively on the exosystem. This, in turn, allows us to conclude that, as long as closed-loop stability is preserved, no perturbation of (A, B, C) can break the asymptotic property of $e = 0$. For a more formal and in-dept discussion of robustness, the reader is referred to Chapter 3.

P2 **Necessity of the exosystem:** on the other hand, the perfect knowledge of the exosystem is a key requirement to ensure output regulation. The linear regulator gives, indeed, no robustness with respect to perturbations of the exosystem, in the sense that any arbitrarily small perturbation of S will reflect in a non-zero asymptotic error. We also observe that knowing the exosystem does not mean knowing $w(t)$, yet only knowing the *class of signals to which the ideal error-zeroing input* $u^\star = \Gamma w(t)$ *will belong to*. This fact will be a key observation in the adaptive framework presented in Chapter 6.

P3 Independence of η from ξ: while the design of the stabiliser depends on the internal model unit, as it is supposed to stabilize the cascade (x, η), the design of the internal model unit η can be done beforehand and independently on the stabiliser ξ (indeed η only depends on the exosystem). As we will see in a while, this property that permits a sequential design of the regulator is inexorably lost if a general nonlinear regulator is sought. Linearity is of course decisive in guaranteeing this independence; as a matter of fact, while the matrices Φ and G do not depend on the definition of (1.15), the matrices Π and Γ do. Linearity of the exosystem and of the plant, though, imply that regardless what stabilizer is chosen, the ideal error-zeroing input Γw will anyway satisfy (1.7), thus making the dependency on ξ fading away.

P4 Multivariableness: The approach is structurally multivariable with the only requirement (implicit in the non-resonance condition of Assumption 1.1) that the number of inputs is larger or equal than the number of regulation errors. The multivariable case naturally motivates a regulator structure in which the internal model *post-processes the error* (see Figure 1.2), namely internal models are put in cascade to the plant with the errors as input.

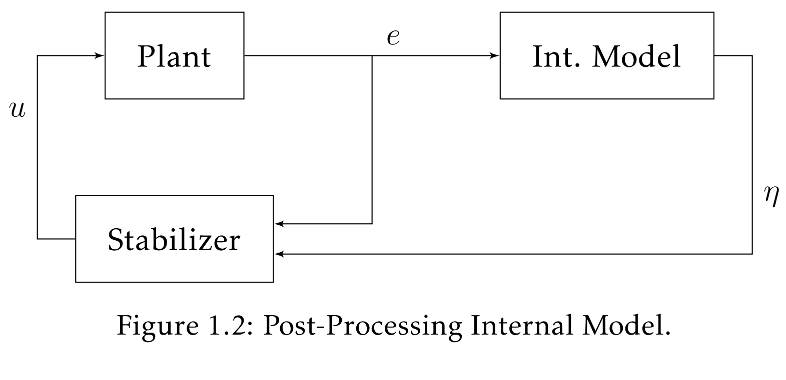

Figure 1.2: Post-Processing Internal Model.

1.2 Output Regulation of Nonlinear Systems

Output regulation of nonlinear systems is nowadays an active and definitely open research field. This section is devoted to present, in simplest possible terms, a state of the art of the most consolidated approaches to nonlinear output regulation.

1.2.1 The Framework of Output Regulation

We consider continuous-time nonlinear systems described by differential equations of the form

$$\begin{aligned} \dot{x} &= f(w, x, u) \\ y &= h(w, x), \end{aligned} \tag{1.20}$$

with state $x \in \mathbb{R}^n$, control input $u \in \mathbb{R}^m$, measured outputs $y \in \mathbb{R}^{n_y}$ and with $w \in \mathbb{R}^{n_w}$ that represents exogenous signals, such as references to be tracked and disturbances to be rejected. As a standing assumption, in the whole chapter we suppose that w is generated by an exosystem of the form

$$\dot{w} = s(w). \tag{1.21}$$

Associated to (1.20), there is a set of $n_e > 0$ *regulation errors* defined as

$$e = h_e(w, x), \tag{1.22}$$

with $h_e : \mathbb{R}^{n_w} \times \mathbb{R}^n \to \mathbb{R}^p$, that represent the errors between the regulated variables and the corresponding references, or selected state variables on which the steady state effect of the exogenous variables should be eliminated. As a second standing assumption, we assume e to belong to the set of measurable outputs. Namely, we suppose that $h(w, x) = \mathrm{col}(h_e(w, x), h_a(w, x))$, where $y_a = h_a(w, x)$ represents some additional measurements that are not required to vanish in steady state but that can be useful for stabilization or other purposes. All the previous functions are assumed to be sufficiently smooth with a degree of smoothness that will be clear from the context.

In this framework we define the problem of ϵ-approximate output regulation as follows: find an output feedback regulator of the form

$$\begin{aligned} \dot{x}_c &= \varphi(x_c, y) \\ u &= \gamma(x_c, y), \end{aligned} \tag{1.23}$$

possibly ϵ-dependent, with state $x_c \in \mathbb{R}^{n_c}$, such that:

P1 **Stability:** The origin of the interconnection (1.20), (1.23) with $w = 0$ is asymptotically stable with a domain of attraction $X \times X_c \subset \mathbb{R}^n \times \mathbb{R}^{n_c}$ that is an open neighborhood of the origin.

P2 **Boundedness:** There exists $W \subset \mathbb{R}^{n_w}$ such that the closed-loop system (1.20), (1.21), (1.23) is uniformly bounded from[4] $W \times X \times X_c$.

P3 **Regulation:** Each solution to the closed-loop system (1.20), (1.21), (1.23) originating in $W \times X \times X_c$ satisfies

$$\limsup_{t \to \infty} |e(t)| \leq \epsilon.$$

If X coincides with $\mathbb{R}^n$, we say that the problem is solved *globally*, otherwise we say that the problem is solved *locally*. If given each $X \subset \mathbb{R}^n$ it is possible to find a possibly X-dependent regulator of the form (1.23) that solves the problem in X, we say that the problem is solved *semi-globally*. If $\epsilon = 0$, we refer to the problem as the *asymptotic* output regulation problem, and a regulator that solves it is called an *asymptotic regulator* or it is said to achieve *asymptotic regulation* for (1.20), (1.21). Finally, we talk about the *practical* regulation problem whenever, given any $\epsilon > 0$, there exists a regulator of the form (1.23) that solves the ϵ-approximate output regulation problem. Practical regulation is subject to the same taxonomy in terms of local, global and semi-global terms.

1.2.2 A Brief Overview of Nonlinear Regulation: From Local to (Semi-)Global

Nonlinear versions of the internal model principle have at first appeared in a local setting in the seminal papers (Isidori and Byrnes, 1990; Huang and Rugh,

[4]Namely there exists a compact set $M \subset \mathbb{R}^{n_w+n+n_c}$ and a $\tau > 0$ such that $\mathcal{R}^{\tau}_{(w,x,x_c)}(W \times X \times X_c) \subset M$.

1990; Huang and Lin, 1994). In (Isidori and Byrnes, 1990), in particular, the authors presented necessary and sufficient condition for local regulation and provided an extension of the concept of steady state to nonlinear systems based on the center manifold theory (Khalil, 2002). The early 90s knew a quite large proliferation of local approaches to nonlinear output regulation, mainly based on linearization (Huang and Lin, 1994; Huang, 1995; Byrnes et al., 1997a) or feedback-linearization (Khalil, 1994), and with an ideal error-zeroing input that, however, was still supposed to be generated by a linear system (see also (Byrnes et al., 1997b) for a complete treatise of structurally stable local regulation in this setting and (Marconi and Isidori, 2000) for a geometric perspective and for a mixed design of feedforward actions and internal models). The late 90s and early 2000s, instead, knew a considerable trend toward the extension of the local approaches to global (Khalil, 1998; Serrani and Isidori, 2000) and semi-global (Serrani et al., 2001; Isidori et al., 2002; Ding, 2003; Huang and Chen, 2004) settings. All these designs, though, were based essentially on the same linearity assumption of the internal model (Huang, 2001). A purely nonlinear theory for non-local output regulation appeared only in 2003, in the pioneering papers (Byrnes and Isidori, 2003, 2004), where the concept of nonlinear steady state and a purely nonlinear internal model principle have been re-framed in the context of non-equilibrium theory (Byrnes and Isidori, 2002; Byrnes et al., 2003). The "Byrnes-Isidori" high-gain design proposed in (Byrnes and Isidori, 2004), in particular, is one of the most celebrated regulator, which knew several further developments (Delli Priscoli et al., 2006; McGregor et al., 2006; Isidori et al., 2012; Forte et al., 2017), and which is still nowadays playing a key part in recent advances. Few years later another milestone design, the Marconi-Praly-Isidori regulator, appeared in (Marconi et al., 2007), leveraging the theory of nonlinear Luenberger observers (Andrieu and Praly, 2006). This latter approach, less constructive yet more general than the Byrnes-Isidori one, has been complemented in (Marconi and Praly, 2008a) in a constructive *practical* regulation framework, and has been the main subject to recent extensions to some classes of multivariable nonlinear systems (Astolfi et al., 2013; Wang et al., 2016, 2017; Pyrkin and Isidori, 2017). Another construction (that is referred here as the Chen-Lu-Huang regulator), perhaps more general than the Byrnes-Isidori regulator, even though applied to more restrictive class of plants, was proposed in (Lu and Huang, 2015), based on the concept of *steady-state generator* (Huang and Chen, 2004; Chen and Huang, 2005). Finally it is worth mentioning the quite recent approach pursued in (Astolfi et al., 2015; Astolfi and Praly, 2017), where the regulator follows a (local) design paradigm closer to the linear perspective. We will get back to these regulators in the chapters 2 and 3, when talking about post-processing regulators and robustness, and we will instead give more technical details about the works (Byrnes and Isidori, 2003; Marconi et al., 2007) in the following section.

1.2.3 The Byrnes-Isidori and the Marconi-Praly-Isidori Regulators

In this section we briefly present the two main approaches to nonlinear regulation designs that have influenced this thesis most strongly. Contrary to the linear setting, the nonlinear regulation theory has mainly developed around Single-Input-Single-Output (SISO) systems, and designs for multivariable plants usually consist in direct extensions of results originally given for the SISO case. As done in the original works, in this section we thus present the results for SISO nonlinear systems. For the related extensions to multivariable systems the reader is referred to Section 1.2.4 thereafter, while for a more detailed treatise, the reader is referred to the original papers (Byrnes and Isidori, 2003, 2004; Marconi et al., 2007; Marconi and Praly, 2008a; Isidori, 2017).

The Framework

We restrict the focus on a subclass of nonlinear systems (1.20),(1.21) obtained with $u \in \mathbb{R}$ and by partitioning the state x as $x = (z, e)$, with $z \in \mathbb{R}^{n_z}$, $n_z := n - 1$ and $e \in \mathbb{R}$, where z and e satisfy following equations

$$\begin{array}{rcl} \dot{w} & = & s(w) \\ \dot{z} & = & f(w, z, e) \\ \dot{e} & = & q(w, z, e) + b(w, z, e)u \\ y & = & e, \end{array} \tag{1.24}$$

for some smooth functions $f : \mathbb{R}^{n_w+n_z+1} \to \mathbb{R}^{n_z}$ and $q, b : \mathbb{R}^{n_w+n_z+1} \to \mathbb{R}$. The state e coincides with the regulation error, and it is the only output of the system. Systems having the form (1.24) are called *(SISO) normal forms*. We assume in the following that b is bounded away from zero and it is sign-definite, i.e. there exists $\underline{b} > 0$ such that $\underline{b} \leq b(w, z, e)$ in the whole state space. We shall also assume that the initial conditions $w(0)$ of the exosystem range in a compact invariant set $W \subset \mathbb{R}^{n_w}$. It can be shown (see Byrnes and Isidori, 2003, Sec. V) that the system (1.24) possesses a well-defined *zero dynamics* given by the system

$$\begin{array}{rcl} \dot{w} & = & s(w) \\ \dot{z} & = & f(w, z, 0). \end{array} \tag{1.25}$$

Common to all the three approaches there is the standing assumption of *minimum phase*. This assumption comes in four main versions detailed below:

Assumption 1.2. *(Weak Minimum Phase) There exists a compact set $\mathcal{A} \subset W \times \mathbb{R}^{n_z}$ which is locally asymptotically stable (LAS) for* (1.25).

Assumption 1.3. *(Weak Minimum Phase + LES) Assumption 1.2 holds with $\mathcal{A}$ that is also locally exponentially stable.*

Assumption 1.4. *(Strong Minimum Phase) There exists a compact set $\mathcal{A} \subset W \times \mathbb{R}^{n_z}$ such that the (w, z) subsystem of* (1.24) *is input-to-state stable (ISS) with respect to $\mathcal{A}$ and relatively to the input e, namely there exist $\beta \in \mathcal{LK}$ and $\rho \in \mathcal{K}$ such that each solution to* (1.25) *satisfies*

$$|(w(t), z(t))|_{\mathcal{A}} \leq \beta(|(w(0), z(0))|_{\mathcal{A}}, t) + \rho(|e|_{[0,t)}). \tag{1.26}$$

Assumption 1.5. *(Strong Minimum Phase + Linear ρ) Assumption 1.4 holds with ρ linear.*

Remark 1.3. Clearly, Assumption 1.5 $\implies$ Assumption 1.4 $\implies$ Assumption 1.3 $\implies$ Assumption 1.2. Nevertheless, there is a quite standard method that, if the regulator achieve some given properties, allows to extend results proved under Assumptions 1.5 or 1.4 to cases in which, respectively, only Assumptions 1.3 or 1.2 hold. This "standard machine" is based on the key observation that local asymptotic stability implies local ISS (i.e. that Assumption 1.4 holds locally). Thus, if the regulator can be tuned to ensure that the state e converges to an arbitrarily small neighborhood in an arbitrarily small time, despite the (bounded) value of z, then for initial conditions of (w, z) close enough to $\mathcal{A}$, the fast transitory of e is not able to make (w, z) exit the set in which Assumption 1.4 holds, and the same argument used to prove the result under this latter assumption can be used. Since LES also implies Assumption 1.5 locally, then the same arguments can be translated to Assumptions 1.3 and 1.5. △

Remark 1.4. Together with the SISO limitation, the assumption of minimum phase is the main trait making abrupt the passage from the linear regulation theory to the nonlinear counterpart. For the regulators presented in this sections, the minimum phase assumption is asked to support a stabilization mechanism strongly oriented towards *"high-gain"* techniques, for which a systematic theory to deal with zero dynamics that are unstable relative to a set larger than the origin does not exist yet. △

The Byrnes-Isidori Regulator

The Byrnes-Isidori regulator has originally be presented in (Byrnes and Isidori, 2004). A semi-global practical result was given under Assumption 1.2, which becomes asymptotic whenever Assumption 1.3 holds. Given a compact set $Z \times E \subset \mathbb{R}^{n_z+1}$ of initial conditions, the regulator builds on the following standing assumption:

Assumption 1.6. *With $\mathcal{H}_w$ denoting the exosystem* (1.21), *the following hold:*

1. $\Omega_{\mathcal{H}_w}(W) = \bigcup_{w \in W} \Omega_{\mathcal{H}_w}(w)$,
2. *The positive orbit of $W \times Z$ under the flow of* (1.25) *has compact closure and* $\mathcal{A} := \Omega_{(1.25)}(W \times Z) \subset \text{int}(W \times Z)$.

It is clear from (1.24) that the set of the possible ideal error-zeroing control law are given by the functions[5]:

$$u^\star := -\frac{q(w,z,0)}{b(w,z,0)} \tag{1.27}$$

as (w, z) range in the set of solutions to (1.25) originating in $\mathcal{A}$. A last assumption is required, that reads as follows:

[5]In this respect, we notice that the result in (Byrnes and Isidori, 2004) was given with $b = 1$. Nevertheless the result can be easily shown to apply to the case considered here.

Assumption 1.7. *There exists $d \in \mathbb{N}$ and $\phi : \mathbb{R}^d \to \mathbb{R}$, such that, for each solution*[6] $(w, z) \in \mathcal{S}_{(1.25)}(\mathcal{A})$*, the function $u^\star(t)$ defined in* (1.27) *fulfills*

$$u^{\star(d)} = \phi\big(u^\star, \dot{u}^\star, \dots, u^{\star(d-1)}\big).$$

Remark 1.5. Assumption 1.7 extends the linearity assumption ubiquitous in the previous frameworks (Huang, 2001) in which ϕ were a linear map. It asks that the constrained system

$$\begin{array}{rcl} \dot{w} & = & s(w) \\ \dot{z} & = & f(w, z, 0) \end{array} \qquad (w, z) \in \mathcal{A} \tag{1.28}$$

with output $u^\star = -q(w, z, 0)/b(w, z, 0)$ is *immersed* into a system of the form

$$\begin{array}{rcl} \dot{v}_i & = & v_{i+1}, \qquad i = 1, \dots, d-1 \\ \dot{v}_d & = & \phi(v_1, \dots, v_d) \\ u^\star & = & v_1, \end{array} \tag{1.29}$$

in the sense that each output produced by (1.28) can be reproduced by (1.29). Checking such assumption requires the knowledge of all the solutions to (1.28), and thus it might be impractical. Nevertheless, under the quite common (see Section 1.2.4) further assumption that, for some sufficiently smooth map $\pi : W \to \mathbb{R}^{n_z}$, $\mathcal{A} = \operatorname{graph} \pi$, the above immersion condition is equivalent to ask that the function $c(w) := -q(w, \pi(w), 0)/b(w, \pi(w), 0)$ satisfy

$$L^d_{s(w)} c(w) = \phi(c(w), \dots, L^{d-1}_{s(w)} c(w)),$$

thus making Assumption 1.7 a property of the exosystem, and creating a direct link to the linear case. △

We define the map $\tau : \mathcal{A} \to \mathbb{R}^{n_w} \times \mathbb{R}^{n_z}$ as:

$$\tau_1(w, z) = -\frac{q(w, z, 0)}{b(w, z, 0)}, \qquad \tau_i(w, z) = \frac{\partial \tau_{i-1}(w, z)}{\partial (w, z)} \begin{pmatrix} s(w) \\ f(w, z, 0) \end{pmatrix}, \qquad i = 2, \dots, d,$$

then the Byrnes-Isidori regulator is a system with state $\eta \in \mathbb{R}^d$ satisfying the

[6] We recall that whenever ★ denotes a system and X a set, $\mathcal{S}_\star(S)$ denotes the set of solutions to ★ originating in X (see the notation section).

following equations

$$\begin{array}{rcl}\dot{\eta} & = & A\eta + B\phi_s(\eta) + Gv \\ u & = & C\eta + v \\ v & = & -ke,\end{array} \tag{1.30}$$

where (A, B, C) is a triplet in prime form of dimension d (see the Notation section), $\phi_s : \mathbb{R}^d \to \mathbb{R}$ is any Lipschitz function that agrees with ϕ on the set $\tau(\mathcal{A})$, $k > 1$ is a control parameter, $G := \mathrm{col}(c_1 g,\, c_2 g^2,\, \ldots,\, c_d g^d)$, being $g > 1$ a second control parameter, and $(c_1, \ldots, c_d) \in \mathbb{HC}(d)$. The auxiliary input $v = -ke$ is a high-gain stabilizing component, while $C\eta = \eta_1$ must asymptotically generate $u^\star$. The regulator (1.30) is characterized by the following result, which is adapted from (Byrnes and Isidori, 2004, Prop. 1).

Proposition 1.3. *(Byrnes and Isidori, 2004) Let W and Z be compact, and suppose that Assumptions 1.2, 1.6 and 1.7 hold, the first with a domain of attraction including $W \times Z$. Pick any compact sets $H \subset \mathbb{R}^d$ and $E \subset \mathbb{R}$. Then, there exist $g^\star > 0$ and, for every $g \geq g^\star$ and $\epsilon > 0$, a $k^\star(g, \epsilon) > 0$ such that, if $g \geq g^\star$ and $k \geq k^\star(g, \epsilon)$, the positive orbit of $W \times Z \times E \times H$ under the closed-loop system* (1.24), (1.30) *is bounded and there exists $\bar{t}$ such that $|e(t)| \leq \epsilon$ for all $t \geq \bar{t}$. If, in addition, $\mathcal{A}$ is also locally exponentially stable for* (1.25)*, then $\lim_{t\to\infty} e(t) = 0$.*

The Marconi-Praly-Isidori Regulator

The Marconi-Praly-Isidori regulator was originally introduced in (Marconi et al., 2007), under Assumption 1.2 and with $b(w, z, e) = 1$ (though, the results extend with minor modification). The result is still semi-global, i.e. the initial conditions of (1.24) are supposed to range in a given arbitrary compact set $Z \times E \subset \mathbb{R}^{n_z+1}$. The regulator has state $\eta \in \mathbb{R}^d$, $d \in \mathbb{N}$ and is described by equations of the form

$$\begin{array}{rcl}\dot{\eta} & = & F\eta + Gu, \qquad \eta(0) \in M \\ u & = & \gamma(\eta) + v \\ v & = & \kappa(e)\end{array} \tag{1.31}$$

where $(F, G) \in \mathbb{R}^{d\times d} \times \mathbb{R}^{d\times 1}$, $M \subset \mathbb{R}^d$, $\gamma : \mathbb{R}^d \to \mathbb{R}$ and $\kappa \in \mathbb{R} \to \mathbb{R}$ are all to be fixed. Without any further assumptions (but with the functions s, f, q and b sufficiently regular) the result of (Marconi et al., 2007) reads as follows.

Theorem 1.1. *([Marconi et al., 2007](#)) There exists* $d \in \mathbb{N}$*, a controllable pair* $(F, G) \in \mathbb{R}^{d\times d} \times \mathbb{R}^{d\times 1}$*, a continuous function* $\gamma : \mathbb{R}^d \to \mathbb{R}$ *and, for any compact set* $M \subset \mathbb{R}^d$*, a continuous function* $\kappa \in \mathbb{R} \to \mathbb{R}$*, such that the regulator* (1.31) *solves the problem of asymptotic output regulation for* (1.24)*.*

The stabilizing action $v = \kappa(e)$ is in general nonlinear, and can be taken of the form

$$\kappa(e) = -\text{sign}(e)\alpha(|e|),$$

with $\alpha \in \mathcal{K}_\infty$. Furthermore, if Assumption 1.2 is substituted by Assumption 1.3, then κ can be taken linear (Marconi et al., 2007, Thm. 3), i.e. so that

$$v = -ke$$

for large enough $k > 0$. Moreover, the same linear choice will guarantee practical regulation when only Assumption 1.2 holds (Marconi et al., 2007, Thm. 2).

Remark 1.6. We observe that the structure of the Byrnes-Isidori internal model unit (1.30) has the same form of usual high-gain observers (Khalil and Praly, 2013), where the output injection term is substituted by the stabilizing action v. In the same way, the Marconi-Praly-Isidori regulator (1.31) has the same form of the nonlinear Kazantzis-Kravaris/Luenberger observers (Andrieu and Praly, 2006), where again the output injection term is substituted by the stabilizing action v. This is obviously not a coincidence, and the reason is that in a certain time-scale the two internal model unit act as an observer for the process generating the ideal error-zeroing control law $u^\star$, and v is used as a proxy variable for the output injection term. △

1.2.4 Extensions and Further Developments

In the years after (Marconi et al., 2007) was published, the Marconi-Praly-Isidori was subject to a number refinements: in (Marconi and Praly, 2008a) several exact and approximate expressions of γ have been proposed in a complete framework in which practical regulation can be solved. In (Delli Priscoli et al., 2008) the same design was applied in presence of redundant measurements, in (Marconi and Praly, 2008b) several issues concerning the design of the stabilizing action κ have been considered while in (Isidori et al., 2010) sufficient conditions to be allowed to take γ Lipschitz have been given. In (Isidori and Marconi, 2012) the regulator was "shifted" to the regulation error, i.e. instead of v, the internal model's input was taken as e (we will get back to this issue in Chapter 2). The Marconi-Praly-Isidori regulator was also extended to some classes of multivariable nonlinear systems: a first tiny extension of (Isidori and Marconi, 2012) to multivariable square normal forms was given in (Astolfi et al., 2013), while invertible multivariable systems have been considered in (Wang et al., 2016, 2017; Pyrkin and Isidori, 2017).

The Byrnes-Isidori regulator was the subject to extensions mainly in adaptive and robust contexts. In this respect, it is worth citing (Delli Priscoli et al., 2006), where the regulator has been augmented with a basic adaptation mechanism, (Isidori et al., 2012), where immersion arguments building on Assumption 1.7 have been used to deal with uncertain oscillators without adaptation (an extension of this will be the subject of Section 3.2), and (Forte et al., 2017), where a new adaptive framework based on hybrid identification schemes has been proposed. For what concerns extensions to multivariable systems, up to our knowledge only the "trivial" case of square multivariable normal forms has been considered in (McGregor et al., 2006).

Constructive Designs and the Nonlinear Regulator Equations

Although the Marconi-Praly-Isidori regulator is more general than the Byrnes-Isidori one, in the sense that existence results are given without the immersion Assumption 1.7, the Byrnes-Isidori design is way more "constructive". As a matter of fact, if asymptotic regulation is sough, under Assumption 1.7 it is straightforward to design the regulator (1.30), whereas the is no clue, even under the same assumption, in how to chose (1.31). When practical and approximate regulation problems are considered, moreover, the design of (1.30) is done at the same way as the actual function ϕ of assumption 1.7 were known (Isidori et al., 2012), whereas the constructive procedures for (1.31) proposed in (Marconi and Praly, 2008a) are way more complex to implement, and demanding in terms of computational workload.

Interesting enough, the majority of the constructive extensions of the two regulators require a further common assumption, which is based on a nonlinear version of the regulator equations (1.10)-(1.11), and that reads as follows:

Assumption 1.8. *There exist smooth maps* $\pi : \operatorname{dom}\pi \to \mathbb{R}^{n_z}$ *and* $u^\star : \operatorname{dom} u^\star \to \mathbb{R}$*, defined on open supersets* $\operatorname{dom}\pi$ *and* $\operatorname{dom} u^\star$ *of* W*, solving the following equations*

$$\begin{aligned} L_s\pi(w) &= f(w,\pi(w),0) \\ 0 &= q(w,\pi(w),0) + b(w,\pi(w),0)u^\star(w). \end{aligned} \tag{1.32}$$

Among the papers that require Assumption 1.8 we find almost all the design before the papers (Byrnes and Isidori, 2003, 2004), that in addition ask $u^\star(w)$ to be generated by a linear system (Huang, 2001). Purely nonlinear designs that require Assumption 1.8 are instead: the Chen-Lu-Huang regulator (Chen and Huang, 2005; Lu and Huang, 2015), the (aforementioned) extensions (Isidori and Marconi, 2012; Astolfi et al., 2013; Wang et al., 2016, 2017) of the Marconi-Praly-Isidori regulator and the extensions (Isidori et al., 2012; Forte et al., 2017) of the Byrnes-Isidori regulator.

Equations (1.32) are known as the *nonlinear regulator equations* (Isidori and Byrnes, 1990) and they express the invariance of the set where $e = 0$. In this respect, the function $u^\star$ is the ideal error-zeroing control law, and in the output regulation community it is referred to as the *friend*. The solution $(\pi, u^\star)$ to (1.32) plays a fundamental role and it typically complements one of the assumptions 1.2, 1.3, 1.4 or 1.5 by asking, in addition, that

$$\mathcal{A} = \operatorname{graph}\pi, \tag{1.33}$$

that in turn means that the asymptotic trajectories of the zero dynamics are of the form

$$z(t) = \pi(w(t)).$$

This assumption is usually exploited for the design of the internal model unit, as it is clear from (1.32) that $u^\star$ is given by

$$u^\star(w) = -\frac{q(w,\pi(w),0)}{b(w,\pi(w),0)}, \tag{1.34}$$

i.e. the existence of π permits to express the error-zeroing control law as a function of exclusively w.

Although (1.34) simplifies considerably the problem, because it means that the internal model has to generate signals that are only defined by the exosystem dynamics, it only holds under Assumption 1.8 and if (1.33) holds. These, in turn, are very restrictive assumptions holding in systems that behave almost-linearly. As a matter of fact, the minimum-phase assumptions 1.2-1.4 make reference to an attractor that, in general, is the graph of a set-valued map (Byrnes and Isidori, 2003). The next theorem is the most general sufficient (though not necessary) condition that we found under which we can claim that the steady-state map is single-valued (which is obviously a necessary condition for having Assumption 1.8 and (1.33)). In the forthcoming theorem we make reference to a cascade of the form

$$\Sigma : \begin{cases} \dot{w} &= s(w) \\ \dot{x} &= f(w,x) \end{cases} \tag{1.35}$$

with $w \in \mathbb{R}^{n_w}$, $x \in \mathbb{R}^n$ and with initial conditions that range in a compact set $W \times X \subset \mathbb{R}^{n_w+n}$. We let Σ_w denote the (autonomous) subsystem w of Σ and we make the following structural assumptions

Assumption 1.9. *The following hold:*

1. *Σ is forward complete from $W \times X$.*
2. *W is forward invariant for Σ_w.*
3. *For every $\varepsilon > 0$ there exists $\delta > 0$ such that, for every two solutions $w_1, w_2 \in \mathcal{S}_{\Sigma_w}(W)$, the following holds*

$$|w_1(0) - w_2(0)| \le \delta \qquad \Longrightarrow \qquad |w_1(t) - w_2(t)| \le \varepsilon, \; \forall t \in \mathbb{R}_+.$$

Theorem 1.2. *Consider the system* (1.35) *and suppose the Assumption 1.9 holds. Suppose, moreover, that every two solutions $(w_1, x_1), (w_2, x_2) \in \mathcal{S}_\Sigma(W \times X)$ fulfill*[7]

$$\limsup_{t\to\infty} |x_1(t) - x_2(t)| \le \rho \left(\limsup_{t\to\infty} |w_1(t) - w_2(t)| \right) \tag{1.36}$$

uniformly in $W \times X$. Then there exist a compact set $U \subset W$ and a continuous function $\pi : U \to \mathbb{R}^n$ such that the set graph π *is uniformly attractive for Σ from $W \times X$.*

[7] We observe that the condition (1.36) is a contraction condition typical of systems possessing *incremental stability properties* (see e.g. Lohmiller and Slotine, 1998; Angeli, 2002; Jouffroy and Fossen, 2010).

Proof. Clearly, invariance of W for the subsystem w and (1.36) imply that Σ is uniformly eventually bounded from $W \times X$. Thus (see Proposition 3.8) the Ω-limit set $\Omega_\Sigma(W \times X)$ is compact, non-empty and uniformly attractive from $W \times X$. let $\Pi : \operatorname{dom}\Pi \subset \mathbb{R}^{n_w} \to \mathbb{R}^n$ be the set-valued map

$$\Pi(w) := \{x \in \mathbb{R}^n \,:\, (w,x) \in \Omega_\Sigma(W \times X)\}\,,$$

then $\operatorname{graph}\Pi \subset \mathcal{A}$ and $U := \operatorname{dom}\Pi \subset W$ are not empty and compact, and, hence, Π is upper semicontinuous[8] (see Aubin and Cellina, 1984, Cor. 1, Chap. 1). It remains to show that Π is single-valued. Suppose the opposite, then there exist $\bar{w} \in \mathbb{R}^{n_w}$ and $\bar{x}_1, \bar{x}_2 \in \mathbb{R}^n$ such that $(\bar{w}, \bar{x}_1), (\bar{w}, \bar{x}_2) \in \operatorname{graph}\Pi$ and

$$\bar{x}_1 \neq \bar{x}_2. \tag{1.37}$$

As $\operatorname{graph}\Pi$ coincides with $\Omega_\Sigma(W \times X)$, then, by definition of Ω-limit set, for $i = 1,2$ there exist sequences $((w_i^n, x_i^n))_n$ of $(w_i^n, x_i^n) \in \mathcal{S}_\Sigma(W \times X)$ and $(t_i^n)_n$ of $t_i^n \in \mathbb{R}_+$ such that $t_i^n \to \infty$ and

$$w_i^n(t_i^n) \to \bar{w}, \qquad\qquad x_i^n(t_i^n) \to \bar{x}_i. \tag{1.38}$$

We can write, for each $n \in \mathbb{N}$,

$$\begin{aligned} |\bar{x}_1 - \bar{x}_2| &= |\bar{x}_1 - x_1^n(t_1^n) + x_1^n(t_1^n) - x_2^n(t_2^n) + x_2^n(t_2^n) - \bar{x}_2| \\ &\leq |\bar{x}_1 - x_1^n(t_1^n)| + |\bar{x}_2 - x_2^n(t_2^n)| + |x_1^n(t_1^n) - x_2^n(t_2^n)|. \end{aligned}$$

Equation (1.38) implies that for each $\epsilon > 0$ there exists $\bar{n}_1(\epsilon) \in \mathbb{N}$ such that $n \geq \bar{n}_1(\epsilon)$ implies $|\bar{x}_1 - x_1^n(t_1^n)| + |\bar{x}_2 - x_2^n(t_2^n)| < \epsilon/2$, so that we can write

$$|\bar{x}_1 - \bar{x}_2| < \epsilon/2 + |x_1^n(t_1^n) - x_2^n(t_2^n)| \tag{1.39}$$

for $n \geq \bar{n}_1(\epsilon)$. On the other hand, as $t_i^n \to \infty$, equation (1.36) implies that there exists $\bar{n}_2(\epsilon) \in \mathbb{N}$ such that, for all $n \geq \bar{n}_2(\epsilon)$,

$$|x_1^n(t) - x_2^n(t)| \leq \epsilon/4 + \rho\left(\limsup_{\tau\to\infty} |w_1^n(\tau) - w_2^n(\tau)|\right) \tag{1.40}$$

[8] Namely, for any $w \in \operatorname{dom}\Pi$ and any open set N containing $\Pi(w)$, there exists a neighborhood M of w such that $F(M) \subset N$.

for all $t \geq t_i^n$. Furthermore, for each $n \in \mathbb{N}$,

$$\limsup_{\tau \to \infty} |w_1^n(\tau) - w_2^n(\tau)| = \inf_{t \geq 0} \sup_{\tau \geq t} |w_1^n(\tau) - w_2^n(\tau)| \leq \sup_{\tau \geq \max\{t_1^n, t_2^n\}} |w_1^n(\tau) - w_2^n(\tau)|. \tag{1.41}$$

Let

$$\varepsilon := \rho^{-1}(\epsilon/4),$$

and let δ be the constant for which point 3 of Assumption 1.9 holds with such choice of ε. Then, (1.38) implies that there exists $\bar{n}_3(\epsilon) \in \mathbb{N}$ such that, for all $n \geq \bar{n}_3(\epsilon)$, it holds that

$$|w_1^n(t_1^n) - w_2^n(t_2^n)| \leq \delta,$$

so as point 3 of Assumption 1.9 implies that

$$\sup_{\tau \geq \max\{t_1^n, t_2^n\}} |w_1^n(\tau) - w_2^n(\tau)| \leq \rho^{-1}(\epsilon/4).$$

Hence, by letting $\bar{n}(\epsilon) := \max\{\bar{n}_1(\epsilon), \bar{n}_2(\epsilon), \bar{n}_3(\epsilon)\}$, in view of (1.40)-(1.41) we obtain

$$|x_1^n(t_1^n) - x_2^n(t_2^n)| \leq \epsilon/2$$

for all $n \geq \bar{n}(\epsilon)$, that in turn, in view of (1.39), implies

$$|\bar{x}_1 - \bar{x}_2| < \epsilon.$$

For the arbitrariness of ϵ we thus conclude that $|\bar{x}_1 - \bar{x}_2| = 0$, that contradicts (1.37). Hence we claim that $\bar{x}_1 = \bar{x}_2$ and, for the arbitrariness of $(\bar{w}, \bar{x}_i)$ we conclude that Π is single-valued. As Π is upper semicontinuous, then Π is continuous and the claim of the theorem follows with $\pi := \Pi$. ■

2

Post-Processing Internal Models

The linear regulator, as presented in Section 1.1.3, follows a so-called *post-processing* paradigm, in which the internal model unit is driven by the regulation errors and the stabilizer is designed to ensure the closed-loop stability. Interestingly enough, almost all the approaches to nonlinear regulation show instead a complementary structure (referred to as *pre-processing*), in which the stabilizer is driven by the regulation errors and the internal model unit by the control input. In this chapter we discuss the main properties of the two classes of regulators; we show that, on the first hand, the pre-processing designs are characterized by some strong conceptual limitations preventing their applicability to more general nonlinear systems and that, on the other hand, the post-processing regulators, in principle not affected by such structural drawbacks, introduce an intertwining between the internal model unit and the stabilizer that does not permit a sequential design of the two units. We then present sufficient conditions for the existence of a post-processing regulator for a class of multivariable nonlinear regulation problems that cannot be solved, in their full generality, by existing pre-processing regulators.

2.1 Pre-processing vs Post-processing

Most of the approaches to nonlinear regulation mentioned so far (see Section 1.2.2) are driven by the stabilizing action (v in (1.30) and (1.31)), i.e. by a part of the control input. In fact, the regulators presented in Section 1.2.3 have a common form of the kind

$$\begin{array}{rcl}\dot{\eta} & = & \Phi(\eta) + Gv \\ u & = & \alpha(\eta) + v,\end{array} \tag{2.1}$$

for some functions Φ and α that change for each design, and where v is a stabilizing action depending on e. This is however not the case of the linear regulator (see Section 1.1.3) that instead is directly driven by the unprocessed regulation errors e and, at least in the state-feedback case, can be taken of the form

$$\begin{array}{rcl}\dot{\eta} & = & \Phi\eta + Ge \\ u & = & K_1\eta + K_2x.\end{array} \tag{2.2}$$

We call a regulator of this latter kind a *post-processing* regulator, as it directly processes e. Conversely, we refer to the regulators that are driven by the input as *pre-processing* regulators, as the error is pre-processed by the stabilizer before being accessed by the internal model unit. Figures 2.1 and 2.2 depict the conceptual block diagrams of the two paradigms.

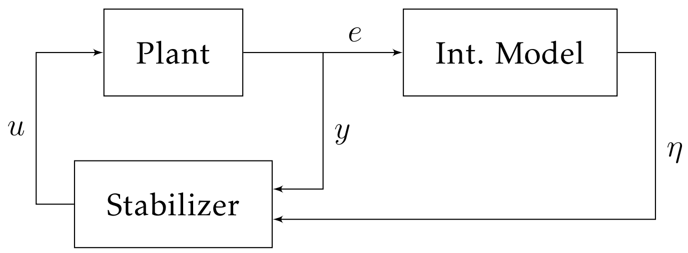

Figure 2.1: Post-Processing Internal Model.

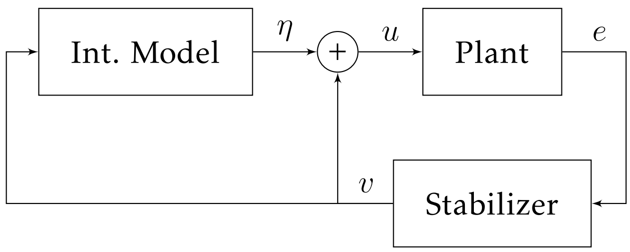

Figure 2.2: Pre-Processing Internal Model.

A part from their structural differences, pre- and post-processing schemes also differ in terms of "design philosophy": in post-processing regulators the plant is augmented with the internal model unit; the stabilizer is designed to stabilize the resulting cascade and to guarantees that the closed-loop system has a well-defined steady state; finally, the properties of the steady-state regulation error are inferred by the structure of the internal model. In pre-processing regulators, instead, the internal model unit is designed to directly generate the ideal error-zeroing control action ($u^\star(w)$ in (1.32)) that makes the ideal steady state in which $e = 0$ ($\pi(w)$ in (1.32)) invariant. The stabilizer is then designed to enforce the desired attractiveness properties of such ideal steady state. Thus, in pre-processing schemes, the ideal steady states of the internal model unit and of the stabilizer are fixed by the plant's data (by the regulator equations). In post-processing regulators, instead, the ideal steady state for the internal model unit and for the stabilizer cannot be fixed a priori. In fact, since η is used for stabilization purposes, and thus it is processed by the stabilizer, its ideal steady state is strongly dependent on the particular instance of the stabilizer.

The pre-processing schemes have the interesting property that the roles of the internal model unit and the stabilizer are neatly separated and the ideal steady state of the closed-loop system is given by the problem statement. *This higher conceptual simplicity is, perhaps, the reason why most of the existing designs are of the pre-processing schemes*. Nevertheless, pre-processing regulators have some structural limitations that prevent their application to larger classes of systems of those treated until now, thus making their applicability to be restricted essentially to minimum-phase square normal forms, where the only outputs usable for feedback are the regulation errors themselves. In particular, it is not clear, at a conceptual level, how a pre-processing regulator could deal in a systematic way with additional outputs not vanishing at the ideal steady sate in which $e = 0$, or with an input dimension larger than those of the errors. If more inputs than errors are present, indeed, it is not clear how to extend the role of the stabilizing action (v in (2.1)) to larger dimensions. If the stabilizing action is taken of dimension equal to $\dim(e)$, then we must find a suitable selection of the inputs to implement it, and this requires the adoption of necessarily non-robust squaring down strategies. Conversely, if the stabilizing action is taken of dimension $\dim(u)$, then it is not clear how to chose the dimension of the internal model, which is fed by the stabilizing action: if the dimension is kept equal to

$\dim(e)$, then a squaring down is required; if, instead, the dimension of the internal model is taken to be equal to $\dim(u)$, then we need to add some redundant internal models, leading to a system that is not stabilizable by error feedback (as a simple linear example would show). On the other hand, there is not even a clear road map to handle additional measured outputs that are necessary to obtain closed-loop stability (or even minimum-phase) but that need not to vanish at the steady state. As a matter of fact, if they contribute to the stabilizing action, then they must be filtered out by the stabilizer at the steady state for the regulator to be consistent with the steady-state specifications.

These conceptual problems, in principle not present in regulators of the post-processing type, recently motivated the community to look for post-processing alternatives to the existing regulators. In (Isidori and Marconi, 2012) the authors tried to "shift" the pre-processing Marconi-Praly-Isidori regulator to an equivalent pre-processing design. The same regulator has been then subject to the minor extension to multivariable square normal forms in (Astolfi et al., 2013). In (Bin and Marconi, 2017b), the Byrnes-Isidori regulator has been shifted to a post-processing version as well. However, no conceptual progress has been made in terms of extension to larger classes of systems compared to pre-processing schemes, and the obtained design are equivalent to the pre-processing counterparts. A different approach to the design of post-processing regulators is the one adopted in (Astolfi and Praly, 2017) and in (Astolfi et al., 2015), where the linear regulator is attached to a class of nonlinear systems. In particular, in (Astolfi and Praly, 2017) the authors showed that the output regulation problem can be solved robustly by a post-processing regulator (an integral action) whenever the steady state is made of equilibria. In (Astolfi et al., 2015), the authors extended the results to the case in which the steady-state signals are periodic, obtaining, however, only an approximate result stating that the Fourier coefficients in the regulation errors corresponding to the frequencies embedded in the internal model vanish at the steady state (this result is treated in a more general envelope in Section 3.4).

In conclusion, almost the totality of nonlinear regulator designs presented so far are of the pre-processing type, and the few existing post-processing examples are either constructed by "shifting" an existing pre-processing regulator (thus leading to no advantage) or are unable to give an asymptotic result if not under very restrictive assumptions on the steady-state trajectories. The reason behind this fact has to be sought in the internal intertwining between the internal model unit and the stabilizer that is structurally present in post-processing schemes and that arises necessarily as far as nonlinear systems are concerned. We explore this property in the next section.

2.2 The Chicken-egg Dilemma of Output Regulation

Post-processing design paradigms do not present, in principle, the main conceptual obstructions of pre-processing schemes. Nevertheless they come with other structural features, not present in the pre-processing case, that make the boundary between the roles of the internal model unit and the stabilizers to fade away, thus invalidating the nice conceptual separation of the two subsystems and leading to a more challenging synthesis phase.

In post-processing regulators the stabilizer is designed to stabilize the cascade of the plant and the internal model unit, by working on the available information given by the plant output y and the internal model state η. It is thus easy to see that the stabilizer strongly depends on the choice of the internal model unit. On the other hand, at the steady state, the internal model unit has to generate the input $\eta^\star$ that, processed by the stabilizer together with the other steady-state plant output $y^\star$, produces the ideal error-zeroing control action $u^\star$. This can be seen by means of a simple example.

Example 2.1. Consider the system

$$\begin{aligned}
\dot{w}_1 &= w_2 \\
\dot{w}_2 &= \phi(w) \\
\dot{e} &= u - w_1
\end{aligned}$$

with $w \in \mathbb{R}^2$ and $e, u \in \mathbb{R}$. By following the linear intuition, we define a candidate internal model unit of the kind

$$\begin{aligned}
\dot{\eta}_1 &= \eta_2 + G_1 e \\
\dot{\eta}_2 &= \hat{\phi}(\eta) + G_2 e,
\end{aligned}$$

with G_1, G_2 and $\hat{\phi}$ to be fixed, and we choose a stabilizing action of the form

$$u = k_1 e + k_2 \eta, \tag{2.3}$$

for some $k_1 \in \mathbb{R}$ and $k_2 \in \mathbb{R}^{1\times 2}$. It is clear from the equation of e that in order to keep e to zero, asymptotically u must equal w_1. Hence, (2.3) implies that the ideal error-zeroing steady-state for η must satisfy

$$k_{21}\eta_1^\star + k_{22}\eta_2^\star = w_1, \qquad \eta_2^\star = \dot{\eta}_1^\star, \qquad \dot{\eta}_2^\star = \hat{\phi}(\eta^\star). \tag{2.4}$$

These three constraints can be condensed in the following equation

$$k_{21}\hat{\phi}(\eta^\star) + k_{22}\frac{\partial\hat{\phi}(\eta^\star)}{\partial\eta_1}\eta_2^\star + k_{22}\frac{\partial\hat{\phi}(\eta^\star)}{\partial\eta_2}\hat{\phi}(\eta^\star) = \phi(w), \tag{2.5}$$

which expresses the fact that the correct map $\hat{\phi}$ and the ideal steady state $\eta^\star$ must necessarily fulfill a condition strongly dependent on w and on the stabilizer gains k_{21} and k_{22}. If ϕ were linear, we could in principle take $\hat{\phi} = \phi$ and play with $\eta^\star$ to have (2.5) fulfilled, by leveraging the fact that a weighted sum of sinusoids is again a sinusoid at the same frequency. If ϕ is not linear, however, a general solution in which $\hat{\phi}$ is not dependent on k_2 is hard to imagine.

A possible way to proceed is to take $k_{22} = 0$ and $k_{21} \neq 0$, thus obtaining from (2.4)-(2.5)

$$\hat{\phi}(\eta^\star) = \frac{1}{k_{21}}\phi(k_{21}\eta^\star), \qquad \eta^\star = \frac{w}{k_{21}}. \tag{2.6}$$

We now notice that, by taking $k_{21} = k_1$, and changing variables as $\eta \mapsto \tilde{\eta} := \eta - \eta^\star$ and $e \mapsto \varepsilon := e + \tilde{\eta}_1$ yields

$$\begin{aligned} \dot{\tilde{\eta}} &= M\tilde{\eta} + B\big(\hat{\phi}(\tilde{\eta} + \eta^\star) - \phi(w)/k_1\big) + G\varepsilon \\ \dot{\varepsilon} &= (G_1 + k_1)\varepsilon + \tilde{\eta}_2 - G_1\tilde{\eta}_1, \end{aligned}$$

with $M := \mathrm{col}((-G_1 1), (-G_2 0))$, $B := \mathrm{col}(0,1)$ and $G := \mathrm{col}(G_1, G_2)$. Assuming for simplicity that ϕ is Lipschitz, (2.6) and $k_{21} = k_1$ yield

$$|\hat{\phi}(\tilde{\eta} + \eta^\star) - \phi(w)/k_1| = \frac{1}{k_1}|\phi(k_1\tilde{\eta} + w) - \phi(w)| \leq L_\phi|\tilde{\eta}|$$

for some $L_\phi > 0$. Hence, the gains G_1 and G_2 can be designed (for instance by high-gain arguments as in (Byrnes and Isidori, 2004)) to make the subsystem $\tilde{\eta}$ ISS relative to the origin and with respect to the input ε, and the gain k_1 can be taken sufficiently negative to induce a contraction in the closed-loop system, thus solving the problem at hand. With this post-processing solution the intertwining between the internal model and the stabilizer is clear: on the first hand, the internal model must satisfy equation (2.6) with $k_{21} = k_1$. On the other hand, the stabilizer gain k_1 must be taken large enough to stabilize the closed-loop system, and how large depends on G_1, G_2 and L_ϕ, i.e. by the structure of the internal model.

Interestingly enough, we also observe how choosing $k_{21} = 1$ instead makes $\hat{\phi}$ and $\eta^\star$ in (2.6) to be independent on k_1. Nevertheless (we leave the computations to the reader and we refer to (Bin and Marconi, 2017b) for further details), if we insist with a similar stabilization approach, the gains G_1 and G_2 turn out to be necessarily proportional to k_1, i.e. we obtain an internal model unit of the form

$$\begin{aligned} \dot{\eta}_1 &= \eta_2 + G_1' v \\ \dot{\eta}_2 &= \phi(\eta) + G_2' v \\ v &= k_1 e, \end{aligned}$$

for some G_1' and G_2', which is exactly the *pre-processing* Byrnes-Isidori regulator. $\triangle$

More in general, the intertwining between the internal model and the stabilizer can be seen by considering a stabilizer of the generic form

$$\begin{aligned} \dot{\xi} &= \varphi(\xi, y, \eta) \\ u &= \gamma(\xi, y, \eta). \end{aligned}$$

The ideal error-zeroing control action $u^\star$, as well as the ideal steady-state value $y^\star$ of the output y, are given by the plant's data (by the regulator equations) and, at the steady state, the stabilizer must necessarily fulfill a *right-invertibility* condition[1] of the kind

$$\begin{array}{rcl} \dot{\xi}^\star & = & \varphi(\xi^\star, y^\star, \eta^\star) \\ u^\star & = & \gamma(\xi^\star, y^\star, \eta^\star) \end{array} \tag{2.7}$$

a.e. for some ideal steady state trajectories $\xi^\star$ and $\eta^\star$. The condition (2.7) clearly underlines that the ideal steady state $\eta^\star$ of the internal model, the ideal steady state $\xi^\star$ of the stabilizer, and the functions φ and γ, that are all unknowns of the same equation, necessarily have to be fixed together at the same time, possibly relying on the knowledge of $y^\star$. As $\eta^\star$ must be a solution of the internal model unit, we then see that the structure of the internal model becomes dependent on the stabilizer. In (Bin and Marconi, 2018a,b) we called this intertwining between the internal model unit and the stabilizer the *"chicken-egg dilemma"* of output regulation to underline that, *if we insist to separate a regulator in an internal model unit and a stabilizer, then the two units cannot be designed by means of a sequential strategy as in pre-processing schemes but, rather, they have to be co-designed.*

In linear systems the chicken-egg dilemma is broken by linearity, as it implies that, no matter how the stabilizer is chosen in the class of linear systems, the steady-state closed-loop signals will have the same modes of the driving exosystem. Therefore, choosing the internal model unit to embed such modes permits to bypass the difficulties introduced by (2.7) as linearity ensures that, for any possible $\xi^\star$ and $y^\star$, the corresponding $\eta^\star$ exists and is producible by a system of appropriate dimension that contains the same frequencies of the exosystem. Thus, all the possible uncertainties in the particular value of $\eta^\star$ coming from the chicken-egg dilemma will just reflect into the *right initialization* of the internal model unit, by leaving its structure untouched. This last fact is, in a nutshell, the only reason why the linear regulator is "robust": *all the considered uncertainties, coming from plant's uncertain parameters and from the chicken-egg dilemma, do not change the structure of the "right" internal model unit to be implemented, but only its correct initialization.*

In the case of nonlinear system this fortunate conditions are far to be possible, and the chicken-egg dilemma has to be faced or avoided in some other way. The most common way to avoid dealing with it is to go for pre-processing

[1]Calling (2.7) a right-invertibility condition is motivate by the fact that finding the set $\{(\xi^\star, \eta^\star) : u^\star = \gamma(\xi^\star, y^\star, \eta^\star)\}$ coincides with finding the graph of the right-pseudoinverse of γ. This is better show in the case in which $y^\star$ is not present in the second equation, i.e. when we can write $u^\star = \gamma(\xi^\star, \eta^\star)$. In such case, indeed, we can take $(\xi^\star, \eta^\star) = \gamma^r(u^\star)$, with γ^r such that $\gamma \circ \gamma^r$ is the identity.

schemes, with the drawback, however, that only limited classes of systems can be considered. In most of the primordial nonlinear output regulation literature (see e.g. (Byrnes et al., 1997a; Huang and Chen, 2004; Huang, 2001) and all the successive designs based on the same idea) the chicken-egg dilemma was avoided by assuming that the steady-state signals are defined by an algebraic function of the exosystem trajectory $w(t)$ (i.e. essentially Assumption 1.8) and the ideal error-zeroing control law $u^\star$ can be written as $u^\star(t) = c(w(t))$ for some polynomial functions c. In that case, for specific classes of systems (Byrnes et al., 1997b; Huang, 2001) and with the exosystem that is linear, it can be shown that any $\eta^\star$ coming from the inversion (2.7) can be generated by a linear system (of dimension in general larger than those of the exosystem) and, thus, simple arguments can be used to ensure asymptotic regulation. On the same line, in the framework of (Astolfi and Praly, 2017) the chicken-egg dilemma is avoided thanks to the assumption that all the possible steady states are equilibria (and thus (2.7) has the easy solution of the integrator). In (Astolfi and Marconi, 2015) the authors tried to extend the result of (Astolfi and Praly, 2017) to the case of periodic steady states, and the construction of the regulator was possible only by sacrificing asymptotic regulation, by giving a result that, in general, is only approximate and not even practical.

In the next section we present some sufficient conditions under which a post-processing regulator can be constructed in the context of (partial) normal forms that can deal with *non-square systems* and can manage *additional non-vanishing outputs* in a systematic way. These conditions presented are quite not constructive in practice, as they require a level of detail about the system that is hard to achieve. However the result show how in principle asymptotic post-processing regulators can be constructed for classes of system that cannot be considered in pre-processing schemes.

2.3 A Post-Processing Regulator for Multivariable Nonlinear Systems

This section contains unpublished original results, except for a very preliminary idea appeared in (Bin and Marconi, 2017b). Motivated by the previous discussion about the conceptual limitation of pre-processing approaches, we present here the construction of a regulator of the post-processing kind. We give sufficient conditions for asymptotic regulation and we prove that, even if they are not met, a practical regulation result of the same kind of those proved in (Isidori et al., 2012) for the Byrnes-Isidori regulator holds. As we will further comment in Section 2.3.3, the chicken-egg dilemma clearly manifests when the different degrees of freedom of the regulator have to be chosen to fulfill the conditions for asymptotic regulation. As opposite to regulators of the pre-processing type, here we deal with non-square systems having more inputs than measured outputs, and we can handle additional measured outputs that are not vanishing at the steady state. Even if conceptually interesting, the result however still limits to a design procedure strongly based on a high-gain perspective, that remains the major conceptual limitation of this approach.

2.3.1 The Framework

We consider controlled systems of the form

$$\begin{array}{rcl} \dot{w} & = & s(w) \\ \dot{x} & = & f(w,x) + b(w,x)u \\ y & = & \begin{pmatrix} e \\ y_a \end{pmatrix} = \begin{pmatrix} h_e(w,x) \\ h_a(w,x) \end{pmatrix}, \end{array} \tag{2.8}$$

where $w \in \mathbb{R}^{n_w}$, $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, $e \in \mathbb{R}^{n_e}$, $y_a \in \mathbb{R}^{n_a}$, $y \in \mathbb{R}^{n_y}$ ($n_y = n_e + n_a$) and with s, f and b sufficiently smooth functions with the property that there exist $r > 0$, a set of integers $p_1, \dots, p_r > 0$ satisfying $p_1 + \cdots + p_r = n_y$, a set of integers $N_1, \dots, N_r > 0$ satisfying $p_1 N_1 + \cdots + p_r N_r =: N \le n$, and, for $i = 1, \dots, r$, a set of $\mathbb{R}^{p_i}$-valued smooth functions[2] $\{\xi_1^i(w,x), \dots, \xi_{N_i-1}^i(w,x), \zeta^i(w,x)\}$ with linearly independent differentials fulfilling

$$\frac{\partial \xi(w,x)}{\partial x} b(w,x) = 0, \qquad \forall (w,x) \in \mathbb{R}^{n_w} \times \mathbb{R}^n \tag{2.9}$$

and that satisfy

$$\dot{\xi} = F\xi + H\zeta \tag{2.10a}$$

[2]With slight abuse of notation, in the following we will call with the same symbols ξ and ζ both the functions $\xi(w,x)$ and $\zeta(w,x)$ and the functions $t \mapsto \xi(w(t),x(t))$ and $t \mapsto \zeta(w(t),x(t))$.

$$\dot{\zeta} = q(w,x) + B(w,x)u \tag{2.10b}$$
$$y = TC\xi \tag{2.10c}$$

for some continuous functions $q : \mathbb{R}^{n_w} \times \mathbb{R}^n \to \mathbb{R}^{n_y}$ and $B : \mathbb{R}^{n_w} \times \mathbb{R}^n \to \mathbb{R}^{n_y \times m}$, and where we let $\xi := \mathrm{col}(\xi^1, \dots, \xi^r) \in \mathbb{R}^{N-n_y}$, $\xi^i := \mathrm{col}(\xi^i_1, \dots, \xi^i_{N_i-1}) \in \mathbb{R}^{p_i(N_i-1)}$, $\zeta := \mathrm{col}(\zeta^1, \dots, \zeta^r) \in \mathbb{R}^{n_y}$ and where $T \in \mathbb{R}^{n_y \times n_y}$ is a known permutation matrix, $C := \mathrm{diag}(C_1, \dots, C_r)$ have the form

$$C_i := \begin{pmatrix} I_{p_i} & 0_{p_i \times p_i(N_i-2)} \end{pmatrix}$$

and $F \in \mathbb{R}^{(N-n_y)\times(N-n_y)}$ and $H \in \mathbb{R}^{(N-n_y)\times n_y}$ are block lower-triangular matrices whose diagonal blocks are given respectively by

$$F_{ii} := \begin{pmatrix} 0_{p_i(N_i-2)} & I_{p_i(N_i-2)} \\ 0_{p_i} & 0_{p_i \times p_i(N_i-2)} \end{pmatrix}, \qquad H_{ii} := \begin{pmatrix} 0_{p_i \times (N_i-2)} \\ I_{p_i} \end{pmatrix}.$$

According to the partition $y = \mathrm{col}(e, y_a)$ we let $C_e \in \mathbb{R}^{n_e \times (N-n_y)}$ and $C_a \in \mathbb{R}^{n_a \times (N-n_y)}$ be such that

$$TC = \begin{pmatrix} C_e \\ C_a \end{pmatrix}.$$

For simplicity, we develop here the case in which $T = I_{n_y}$, i.e. we assume that

$$e = C_e\xi = \mathrm{col}(\xi^i_1 \ : \ i = 1, \dots, r_e)$$
$$y_a = C_a\xi = \mathrm{col}(\xi^i_1 \ : \ i = r_e + 1, \dots, r)$$

where r_e is such that $p_1 + \cdots + p_{r_e} = n_e$ and $r_a := r - r_e$ (we also let $N_e = p_1 N_1 + \cdots + p_{r_e} N_{r_e}$ and $N_a = N - N_e$). We note though, that the result can be easily extended to arbitrary T by means of a simple change of coordinates.

Remark 2.1. The class of systems considered includes multivariable normal forms and partial normal forms (Isidori, 1999), with the latter that always exist locally whenever (possibly after a preliminary feedback) the system (2.8) is *a)* strongly invertible in the sense of (Hirschorn, 1979; Singh, 1981), and *b)* input-output linearizable[3] (Isidori, 1995). For what concerns the computation of the functions

[3]That is, there exists a state feedback control of the form $u = \alpha(w,x) + G(w,x)v$, with $v \in \mathbb{R}^m$ an auxiliary input and G full rank, such that the resulting system has linear input-output behavior from v to y.

$q(w,x)$ and $B(w,x)$ in (2.10a)-(2.10c) in the context of partial normal forms, the reader is referred to (Isidori, 1999; Wang et al., 2015a). △

Remark 2.2. Differently from almost all the previous literature (see e.g McGregor et al., 2006; Astolfi et al., 2013; Wang et al., 2016, 2017), we do not constraint $m = \dim(u)$ to be equal to n_e or n_y, i.e. we consider a *non-square* system in which the number of inputs can be larger than that of the outputs. Furthermore, we structurally handle the feedback of additional outputs y_a that do not need to vanish at the steady state but that might be necessary to obtain the form (2.10a), (2.10b), (2.10c) or to fulfil all the assumptions below. △

We will construct the regulator based on a number of assumptions introduced below:

Assumption 2.1. *There exists a compact set $\mathcal{A} \subset \mathbb{R}^{n_w+n}$, $\beta \in \mathcal{KL}$, and a locally Lipschitz $\rho \in \mathcal{K}$, such that all the solution pairs (w,x,u) to* (2.8) *satisfy*

$$|(w(t),x(t))|_{\mathcal{A}} \leq \beta(|(w(0),x(0))|_{\mathcal{A}},t) + \rho\big(|\xi^e|_{[0,t)} + |\zeta^e|_{[0,t)}\big), \tag{2.11}$$

for all $t \in \mathbb{R}_+$ for which they are defined and with $\xi^e := \mathrm{col}(\xi^i \, : \, i = 1,\dots,r_e)$ and $\zeta^e := \mathrm{col}(\zeta^i \, : \, i = 1,\dots,r_e)$.

Assumption 2.2. *There exists a C^1 map $\mathcal{P} : \mathbb{R}^{n_w+n_x} \to \mathbb{R}^{n_y \times n_y}$ and, for each compact set $X \subset \mathbb{R}^n$, a full-rank matrix $\mathcal{L} \in \mathbb{R}^{m \times n_y}$ such that:*

a. $\mathcal{P}(w,x) > 0$ *in* $W \times \mathbb{R}^n$.

b. For all $(w,x) \in W \times \mathbb{R}^n$ *and all* $u \in \mathbb{R}^m$,

$$L^{(x)}_{b(w,x)u}\mathcal{P}(w,x) = 0.$$

c. for all $(w,z) \in W \times X$

$$\mathcal{L}^T B(w,x)^T \mathcal{P}(w,x) + \mathcal{P}(w,x) B(w,x) \mathcal{L} \geq I. \tag{2.12}$$

Remark 2.3. As in the context of normal forms and partial normal forms, ξ and ζ are combinations of derivatives of the output y, then Assumption 2.1 can be seen as a uniform (in u) "output-input stability" (OIS) property, in the sense of

(Liberzon et al., 2002; Liberzon, 2004), of (w, x) relatively to the set $\mathcal{A}$, that here plays the role of a *strong minimum-phase* assumption. The same minimum-phase assumption appeared for instance in (Wang et al., 2015a, 2016, 2017). However, we stress that here the minimum phase is asked with respect to the *whole set of outputs* (included those that do not need to vanish at the steady state) and, thus, Assumption 2.1 is milder than usual minimum-phase assumptions, and it can be possibly obtained by adding further measurements. We also observe that the result presented here can be extended, in view of Remark 1.3, to the case in which Assumption 2.1 holds only locally provided, however, that the IOS property holds only with respect to $e = \xi_1^e$. △

Remark 2.4. Assumption 2.2 is a controllability assumption. As it will be clarified later below the proof of Proposition 2.1 this assumption is implicated by many customary assumptions made in the context of regulation and stabilization of partial normal forms. △

With $d \in \mathbb{N}$, the regulator is a system with state $\eta \in \mathbb{R}^{n_e d}$ described by the following equations

$$\begin{array}{rcl} \dot{\eta} &=& \Phi(\eta) + Ge \\ u &=& \mathcal{L}\big(\mathcal{K}_\xi \xi + \mathcal{K}_\zeta \zeta + \mathcal{K}_\eta \eta_1\big) \end{array} \tag{2.13}$$

with Φ and G having the form

$$\Phi(\eta) := \begin{pmatrix} 0 & I_{n_e} & 0 & \cdots & 0 \\ 0 & 0 & I_{n_e} & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & I_{n_e} \\ & & \phi(\eta) & & \end{pmatrix}, \qquad G := \begin{pmatrix} G_1 \\ G_2 \\ \vdots \\ G_d \end{pmatrix}$$

where $\phi : \mathbb{R}^{n_e d} \to \mathbb{R}^{n_e}$ is a bounded function satisfying

$$|\phi(\mathbb{R}^{d n_e})| \le C_\phi$$

or some $C_\phi > 0$, and where $G_i \in \mathbb{R}^{n_e \times n_e}$, $\mathcal{K}_\xi \in \mathbb{R}^{n_y \times (N - n_y)}$, $\mathcal{K}_\zeta \in \mathbb{R}^{n_y \times n_y}$ and with $\mathcal{K}_\eta \in \mathbb{R}^{n_y \times n_\eta}$ that has the form

$$\mathcal{K}_\eta = \begin{pmatrix} \mathcal{K}'_\eta \\ 0_{n_a \times n_e} \end{pmatrix}, \tag{2.14}$$

for some $\mathcal{K}'_\eta \in \mathbb{R}^{n_e \times n_e}$. All these degrees of freedom will be fixed later according to Proposition 2.1.

Remark 2.5. We give here a *partial state feedback* result, that employs the auxiliary variables ξ and ζ (i.e. combinations of derivatives of the measured output y). We notice though that a purely output-feedback regulator can be easily obtained by augmenting (2.13) with a partial-state observer of the kind proposed in (Wang et al., 2015a) (see also Teel and Praly, 1995). △

Substituting the expression of u into the ζ subsystem of (2.8) yields

$$\dot{\zeta} = q(w,x) + B(w,x)\mathcal{L}(\mathcal{K}_\xi \xi + \mathcal{K}_\zeta \zeta + \mathcal{K}_\eta \eta_1). \tag{2.15}$$

Let us define the error-zeroing set

$$\mathcal{O} := \Big\{(w,x) \in \mathbb{R}^{n_w \times n_z} \ : \ (w,x) \in \mathcal{A},\ \xi^e(w,x) = 0,\ \zeta^e(w,x) = 0\Big\}.$$

Let q^e and q^a be functions with values in $\mathbb{R}^{n_e}$ and $\mathbb{R}^{n_a}$ respectively such that $q(w,x) = \mathrm{col}(q^e(w,x), q^a(w,x))$, and let denote for brevity

$$D(w,x) := B(w,x)\mathcal{L} \in \mathbb{R}^{n_y \times n_y}.$$

Let partition $D(w,x)$, $\mathcal{K}_\xi$ and $\mathcal{K}_\zeta$ as:

$$D := \begin{pmatrix} D^{e,e} & D^{e,a} \\ D^{a,e} & D^{a,a} \end{pmatrix}, \quad \mathcal{K}_\xi := \begin{pmatrix} \mathcal{K}_\xi^{e,e} & \mathcal{K}_\xi^{e,a} \\ \mathcal{K}_\xi^{a,e} & \mathcal{K}_\xi^{a,a} \end{pmatrix}, \quad \mathcal{K}_\zeta := \begin{pmatrix} \mathcal{K}_\zeta^{e,e} & \mathcal{K}_\zeta^{e,a} \\ \mathcal{K}_\zeta^{a,e} & \mathcal{K}_\zeta^{a,a} \end{pmatrix} \tag{2.16}$$

for some $D^{i,j}(w,x), \mathcal{K}_\xi^{i,j}, \mathcal{K}_\zeta^{i,j} \in \mathbb{R}^{n_i \times n_j}$, $i,j \in \{e,a\}$. Then, in view of (2.14), equation (2.15) gives

$$\begin{aligned} \dot{\zeta}^e = q^e(w,x) + D^{e,e}(w,x)\Big(\mathcal{K}_\xi^{e,e}\xi^e + \mathcal{K}_\xi^{e,a}\xi^a + \mathcal{K}_\zeta^{e,e}\zeta^e + \mathcal{K}_\zeta^{e,a}\zeta^a + \mathcal{K}'_\eta \eta_1\Big) \\ + D^{e,a}(w,x)\Big(\mathcal{K}_\xi^{a,e}\xi^e + \mathcal{K}_\xi^{a,a}\xi^a + \mathcal{K}_\zeta^{a,e}\zeta^e + \mathcal{K}_\zeta^{a,a}\zeta^a\Big). \end{aligned} \tag{2.17}$$

In view of Assumption 2.1, by considering the restriction of this latter equation on the error-zeroing set $\mathcal{O}$, we obtain that invariance of $\mathcal{O}$ is ensured only if there exists a function $\eta_1^\star : \mathcal{A} \to \mathbb{R}^{dn_e}$ satisfying

$$\begin{aligned} D^{e,e}(w,x)\mathcal{K}'_\eta \eta_1^\star(w,x) := -q^e(w,x) - D^{e,e}(w,x)\Big(\mathcal{K}_\xi^{e,a}\xi^a(w,x) + \mathcal{K}_\zeta^{e,a}\zeta^a(w,x)\Big) \\ - D^{e,a}(w,x)\Big(\mathcal{K}_\xi^{a,a}\xi^a(w,x) + \mathcal{K}_\zeta^{a,a}\zeta^a(w,x)\Big) \end{aligned} \tag{2.18}$$

for all $(w, x) \in \mathcal{A}$. The existence of a unique solution to (2.18) in a neighborhood of $\mathcal{A}$ is ensured by the following assumption.

Assumption 2.3. *There exists $\epsilon > 0$ and an open superset $\mathcal{N}_1$ of $\mathcal{A}$ such that*

$$|\min \sigma(D^{e,e}(w,x))| \geq \epsilon$$

for all $(w, x) \in \mathcal{N}_1$.

Under this assumption, if $\mathcal{K}'_\eta$ is invertible, then (2.18) admits a unique solution given by

$$\begin{aligned} \eta_1^\star(w,x) := (\mathcal{K}'_\eta)^{-1} D^{e,e}|_\mathcal{A}(w,x)^{-1}\bigg(& - q^e|_\mathcal{A}(w,x) \\ & - D^{e,e}|_\mathcal{A}(w,x)\Big(\mathcal{K}_\xi^{e,a}\xi^a|_\mathcal{A}(w,x) + \mathcal{K}_\zeta^{e,a}|_\mathcal{A}\zeta^a|_\mathcal{A}(w,x)\Big) \\ & - D^{e,a}|_\mathcal{A}(w,x)\Big(\mathcal{K}_\xi^{a,a}\xi^a|_\mathcal{A}(w,x) + \mathcal{K}_\zeta^{a,a}|_\mathcal{A}\zeta^a|_\mathcal{A}(w,x)\Big)\bigg), \end{aligned}$$

and that solution is bounded with a bound that depends on $\mathcal{A}$, $\mathcal{N}_1$, ϵ and $\mathcal{K}'_\eta$.

The error-zeroing set $\mathcal{O}$ is a subset of $\mathcal{A}$, and $\eta_1^\star$ is only defined on $\mathcal{A}$. Hence, in order to exploit the existence of $\eta_1^\star$, we need to introduce a further structural assumption. With $\varphi(w, x) \in \mathbb{R}^{n_a}$, we consider now the following equations in the unknowns $\lambda_\xi(w,x) \in \mathbb{R}^{N_a-n_a}$, $\lambda_\zeta(w,x) \in \mathbb{R}^{n_a}$ and $\upsilon_1(w,x) \in \mathbb{R}^{n_e}$:

$$\begin{aligned}
&\lambda_\xi(x,w) = \begin{pmatrix} 0 \\ \lambda_\xi^a(w,x) \end{pmatrix}, \qquad \lambda_\xi^a(w,x) \in \mathbb{R}^{N_a-n_a} \\
&\lambda_\zeta(x,w) = \begin{pmatrix} 0 \\ \lambda_\zeta^a(w,x) \end{pmatrix}, \qquad \lambda_\zeta^a(w,x) \in \mathbb{R}^{N_a-n_a} \\
&\frac{\partial \lambda_\xi(w,x)}{\partial w} s(w) + \frac{\partial \lambda_\xi(w,x)}{\partial x} f(w,x) = F\lambda_\xi(w,x) + H\lambda_\zeta(w,x) \\
&\frac{\partial \lambda_\zeta(w,x)}{\partial w} s(w) + \frac{\partial \lambda_\zeta(w,x)}{\partial x} f(w,x) = q(w,z) + B(w,z)\mathcal{L}\Big(\mathcal{K}_\xi \lambda_\xi(w,x) \\
&\qquad\qquad\qquad\qquad + \mathcal{K}_\zeta \lambda_\zeta(w,x) + \mathcal{K}_\eta \upsilon_1(w,x)\Big) + \varphi(w,x)
\end{aligned} \tag{2.19}$$

In view of the first two equations, the same arguments used above show that, whenever $D^{e,e}(w,x)$ and $\mathcal{K}'_\eta$ are invertible, υ_1 reads as

$$\begin{aligned}
\upsilon_1(w,x) := (\mathcal{K}'_\eta)^{-1} D^{e,e}(w,x)^{-1} \Bigg(& - q^e(w,x) - \varphi^e(w,x) \\
& - D^{e,e}(w,x)\Big(\mathcal{K}_\xi^{e,a}\lambda_\xi^a(w,x) + \mathcal{K}_\zeta^{e,a}\lambda_\zeta^a(w,x)\Big) \\
& - D^{e,a}(w,x)\Big(\mathcal{K}_\xi^{a,a}\lambda_\xi^a(w,x) + \mathcal{K}_\zeta^{a,a}\lambda_\zeta^a(w,x)\Big)\Bigg),
\end{aligned} \tag{2.20}$$

for all (w,x) for which it is defined and with φ^a such that $\varphi = \mathrm{col}(\varphi^e, \varphi^a)$ for some suitable φ^e. We observe that, for $\varphi = 0$, equations (2.19), (2.20) have a solution in $\mathcal{O}$, and the solution $(\lambda_\xi, \lambda_\zeta, \upsilon_1)$ is such that $\upsilon_1(w,x) = \eta_1^\star(w,x)$ on $\mathcal{O}$, and the functions λ_ξ and λ_ζ equal the ideal steady state of the variables ξ and ζ in which $\xi^e = 0$, $\zeta^e = 0$ (i.e. the regulation error vanishes). In other words, the solution $(\lambda_\xi, \lambda_\zeta, \upsilon_1)$ to (2.19) represents an *extension* to an open superset of $\mathcal{O}$ of the ideal steady state values that $(\xi, \zeta, \eta_1^\star)$ assumes on $\mathcal{O}$. Hence, in particular, υ_1 is a function defined also outside $\mathcal{O}$ but that coincides with the ideal steady-state control law $\eta_1^\star$ on $\mathcal{O}$. This in turn motivates the following Assumption.

Assumption 2.4. *There exist an open superset $\mathcal{N}_2$ of $\mathcal{A}$, a $L_\varphi > 0$ and a function φ satisfying*

$$|\varphi(w,x)| \le L_\varphi |(w,x)|_{\mathcal{A}}, \qquad (w,x) \in \mathcal{N}_2,$$

such that (2.19) *have a solution* $(\lambda_\xi, \lambda_\zeta, \upsilon_1)$ *defined in* $\mathcal{N}_2$.

2.3.2 The Asymptotic Properties of the Regulator

With $\mathcal{N}_1$ and $\mathcal{N}_2$ the sets given respectively by Assumption 2.3 and 2.4, we let $\mathcal{N}$ be any open set included in $\mathcal{N}_1 \cap \mathcal{N}_2$. We observe that, under such assumptions, the function υ_1 is well-defined on $\mathcal{N}$ and unique. Thus, with d the same integer appearing in the definition of (2.13), we can recursively define the functions:

$$\begin{aligned} \upsilon_i(w,x) &= \frac{\partial \upsilon_{i-1}(w,x)}{\partial w} s(w) + \frac{\partial \upsilon_{i-1}(w,x)}{\partial x}\big(f(w,x) + b(w,x)\upsilon_1(w,x)\big), \\ &\qquad\qquad i = 2, \dots, d+1 \\ \upsilon(w,x) &= \mathrm{col}(\upsilon_i(w,x) \, : \, i = 1, \dots, d). \end{aligned} \tag{2.21}$$

We stress that, in view of (2.19)-(2.20), the functions υ_i all depend on the stabilizer. We further observe that it follows from (2.17) and (2.18) that, in order to ensure asymptotic regulation, the output η_1 of the regulator (2.13) must be able to generate all the signals in the set

$$\mathcal{H}^\star := \Big\{\eta_1 : \mathbb{R} \to \mathbb{R}^{n_e} \, : \, \eta_1(t) = \eta_1^\star(w(t), x(t)), \, (w(0), x(0)) \in \mathcal{A}\Big\}.$$

Assumptions 2.3 and 2.4, in turn, imply that $\eta_1^\star$ coincides with the restriction of the function υ_1 on $\mathcal{A}$, with υ_1 that is defined in a neighborhood $\mathcal{N}$ of $\mathcal{A}$. With the definition of υ given by (2.21), we thus can conclude that υ_1 (and hence $\eta_1^\star$) could be generated by the output η_1 of the regulator (2.13) whenever

$$\phi(\upsilon(w,x)) = \upsilon_{d+1}(w,x), \qquad \forall (w,x) \in \mathcal{A}. \tag{2.22}$$

The condition (2.22) expresses the fact that the regulator (2.13) has the *internal model property*, as it guarantees that the set $\mathcal{H}^\star$ is a subset of the outputs η_1 producible by (2.13). In general, however, υ and υ_{d+1} are uncertain whenever the plant functions f and g are, as indeed these quantities enter explicitly in (2.21). This motivates introducing the following quantity

$$\delta(w,x) := \phi(\upsilon(w,x)) - \upsilon_{d+1}(w,x), \tag{2.23}$$

which represents the *internal model mismatch*, i.e. the modeling error that the system η of (2.13) attains on $\mathcal{N}$ in representing the process that generates $v_1(w, x)$.

The closed-loop system reads as follows

$$\Sigma_{cl}: \begin{cases} \dot{w} &= s(w) \\ \dot{x} &= f(w,x) + b(w,x)\mathcal{L}\big(\mathcal{K}_\xi \xi(w,x) + \mathcal{K}_\zeta \zeta(w,x) + \mathcal{K}_\eta \eta_1\big) \\ \dot{\eta} &= \Phi(\eta) + G h_e(w,x). \end{cases} \tag{2.24}$$

The following proposition states the main properties of the regulator.

Proposition 2.1. *Let $W \subset \mathbb{R}^{n_w}$, $X \subset \mathbb{R}^n$ be arbitrary compact sets and suppose that Assumptions 2.1, 2.2, 2.3 and 2.4 hold. Then there exist a compact set $H \subset \mathbb{R}^{n_e d}$, a $\alpha > 0$, a $g^\star > 0$ and, for each $g \geq g^\star$ and $\varepsilon > 0$, matrices $\mathcal{L}$, $\mathcal{K}_\xi$, $\mathcal{K}_\zeta$, $\mathcal{K}_\eta$, and G_i, $i = 1, \dots, d$, such that:*

1. *the closed-loop system Σ_{cl} is uniformly bounded from $W \times X \times H$,*

2. *there exists $\bar{t}$ such that $\mathcal{R}^\tau_{\Sigma_{cl}}(W \times X \times H) \subset \big(\mathcal{O} + \varepsilon\mathbb{B}\big) \times \mathbb{R}^{n_e d}$ for all $\tau \geq \bar{t}$,*

3. *each solution to Σ_{cl} satisfies*

$$\limsup_{t\to\infty} |e(t)| \leq \frac{\alpha}{g^d} \sup_{(w,x)\in\mathcal{A}} |\delta(w,x)| \tag{2.25}$$

uniformly in the initial conditions in $W \times X \times H$.

2.3.3 Remarks on the Result

The claim of Proposition 2.1 states that the trajectories of the closed-loop system originating in $W \times X \times H$ are equibounded (item 1), that they converge to an ε-neighborhood of $\mathcal{A}$ *uniformly* in the initial conditions (item 2), and that the asymptotic bound on the regulation error is proportional to the worst-case internal model mismatch. Furthermore, as clarified in the proof, ε and the asymptotic bound on the error can be reduced arbitrarily by adjusting g, and this makes the result of the proposition a *practical output regulation* result. Since X is arbitrary, moreover, the result is *semiglobal* in the plant's initial conditions.

According to the proof of Proposition 2.1, the degrees of freedom that characterize the regulator (2.13) are chosen by following a "high-gain" strategy. The only parameters that are truly arbitrary are the order d of the internal model and the bound C_ϕ on ϕ. Lower values of C_ϕ yield lower values of the high-gain parameters g and ℓ, but reduce the representation capabilities of the internal model unit. As a matter of fact, point 3 of the proposition states that the asymptotic bound on the regulation error is related to the maximum value attained by the mismatch (2.23) on the set $\mathcal{A}$; thus if the bound of C_ϕ is too tight, such error might be anyway non zero, even if υ and υ_{d+1} are perfectly known. The functions υ and υ_{d+1} are obtained by the recursion (2.20), (2.21) and, hence, they are strongly dependent on the plant's and the exosystem's functions s, f and g, and on the control parameters $\mathcal{K}_\xi$, $\mathcal{K}_\zeta$ and $\mathcal{K}_\eta$; therefore their perfect knowledge cannot be assumed while fixing C_ϕ. This dependence between the internal model unit and the stabilizer's parameters is a manifestation of the "chicken-egg" dilemma introduced in Section 2.2. The interplay between feedback and internal model is way more evident when non-vanishing outputs are used for stabilization, as they need to be compensated at the steady state by the output of the internal model. In this respect we also note that, as it is the case of the linear regulator, the feedback of auxiliary outputs might also have a simplifying effect on the internal model.

The intertwining between internal model and the stabilizer, though, makes the result of Proposition 2.1 not very constructive, as finding the "right" ϕ to put in (2.13) might be very complicated. Though, Proposition 2.1 individuates a clear sufficient condition for asymptotic regulation, given by equation (2.22), which expresses in this setting the nonlinear version of the *internal model principle*: the regulator must embed a copy of the process that generates the ideal steady state control action (i.e. $\eta_1^\star = \upsilon_1|_{\mathcal{A}}$ above) that makes the set in which the error vanishes invariant.

We also observe that, even if the mismatch δ is not zero on $\mathcal{A}$, the regulator ensures practical regulation. As a matter of fact, for each $\epsilon > 0$, choosing

$$g > \sqrt[d]{\frac{\alpha \sup_{(w,x)\in\mathcal{A}} |\delta(w,x)|}{\epsilon}},$$

yields

$$\limsup_{t\to\infty} |e(t)| < \epsilon.$$

The difficulty of finding the right ϕ and the robustness issues related to its dependency on the plant's data motivated the research toward adaptive designs of regulators of the kind (2.13) presented in Chapter 6.

2.3.4 Proof of Proposition 2.1

We develop the proof in 3 sections. In the first we prove uniform boundedness of the closed-loop system (i.e. point 1 of the proposition), in the second we prove that the trajectories are uniformly attracted by an ε neighborhood of $\mathcal{O}$ (i.e. point 2 of the proposition) and in the third section we prove the bound (2.25) (that is point 3 of the proposition).

I. Uniform boundedness:
Let partition η as $\eta = \operatorname{col}(\eta_1, \dots, \eta_d)$ and, for each $\eta_i \in \mathbb{R}^{n_e}$, let partition η_i as $\eta_i = \operatorname{col}(\eta_i^1, \dots, \eta_i^{r_e})$ with $\eta_i^j \in \mathbb{R}^{p_j}$. Consider the change of coordinates:

$$\begin{aligned} &\forall i = 1, \dots, r_e : \quad && \begin{cases} \xi_1^i \mapsto \chi_1^i := \xi_1^i + \eta_1^i, \\ \xi_j^i \mapsto \chi_j^i := \xi_j^i, \qquad j = 2, \dots, N_i - 1 \end{cases} \\ &\forall i = r_e, \dots, r : && \xi^i \mapsto \chi^i := \xi^i. \\ & && e \mapsto \bar{e} := e + \eta_1. \end{aligned} \tag{2.26}$$

By letting $\chi := \operatorname{col}(\chi^1, \dots, \chi^r)$, with $\chi^i := \operatorname{col}(\chi_1^i, \dots, \chi_{N_i-1}^i)$, (2.26) can be compactly rewritten as

$$\begin{aligned} \chi &:= \xi + C_e^T \eta_1 \\ \bar{e} &= C_e \chi. \end{aligned} \tag{2.27}$$

From (2.13), and since by construction $F C_e^T = 0$, we obtain

$$\dot{\chi} = (F + C_e^T G_1 C_e)\chi + H\zeta + C_e^T(\eta_2 - G_1 \eta_1). \tag{2.28}$$

We state now the following result, which is proved below this proof.

Lemma 2.1. *For any $G_1 \in \mathbb{R}^{n_e \times n_e}$ and $\epsilon > 0$, there exists $K \in \mathbb{R}^{p \times (N-p)}$ such that the system* (2.28) *with output $\bar{e}$ and with input $\zeta = \bar{\zeta} + K\chi$, being $\bar{\zeta} \in \mathbb{R}^p$ an auxiliary input, satisfies*

$$|\chi(t)| \le \beta_\chi(|\chi(0)|, t) + a_1\Big(|\bar{\zeta}|_{[0,t)} + |\eta|_{[0,t)}\Big) \tag{2.29}$$

$$|\bar{e}(t)| \le \beta_{\bar{e}}(|\chi(0)|, t) + \epsilon|\bar{\zeta}|_{[0,t)} + \epsilon|\tilde{\eta}|_{[0,t)} \tag{2.30}$$

for some $\beta_\chi, \beta_{\bar e} \in \mathcal{KL}$ and $a_1 > 0$.

Pick $(h_1, \dots, h_d) \in \mathbb{HC}(d)$ and, with $g > 0$ a control parameter, let

$$G_i := g^i h_i I_{n_e}.$$

Let $\Delta(g) := \mathrm{diag}(1, g, \dots, g^{d-1})$ and change variables as

$$\eta \mapsto \mu := \Delta(g)^{-1}\eta.$$

In the new variables we obtain

$$\dot\mu = \Delta(g)^{-1}\Big(A\Delta(g)\mu + E\phi(\Delta(g)\mu) + G(\bar e - \Gamma\mu_1)\Big)$$

with (A, E, Γ) a triplet in (d, n_e)-prime form. Noting that:

$$\begin{aligned} &\Delta(g)^{-1}A\Delta(g) = gA \\ &\Delta(g)^{-1}E = g^{1-d}E \\ &\Delta(g)G = gR, \qquad R := \mathrm{col}(h_i I_{n_e} \,:\, i = 1, \dots, d), \end{aligned}$$

then, by letting $M := A - R\Gamma$, we obtain

$$\dot\mu = gM\mu + g^{1-d}E\phi(\Delta(g)\mu) + gRC_e\chi. \tag{2.31}$$

Fix $C_\phi > 0$ arbitrarily. As $(h_1, \dots, h_d) \in \mathbb{HC}(d)$, M is Hurwitz; since ϕ is bounded by C_ϕ, then standard high-gain arguments (Khalil and Praly, 2013) can be used to show that the system μ, seen as a system with input χ, fulfills

$$|\mu(t)| \le \beta_\mu(|\mu(0)|, t) + \frac{a_2}{g^d}C_\phi + a_2|\bar e|_{[0,t)}.$$

for some $\beta_\mu \in \mathcal{KL}$ and some $a_2 > 0$. Thus, by letting $\beta_\eta = g^{d-1}\beta_\mu$, using the fact that $|\eta| \le g^{d-1}|\mu|$ and $|\eta_1| \le |\mu|$, we also obtain that the system η seen as a system with input $\bar e$ and output η_1 fulfills

$$\begin{aligned} |\eta(t)| &\le \beta_\eta(|\eta(0)|, t) + \frac{a_2}{g}C_\phi + a_2 g^{d-1}|\bar e|_{[0,t)} \\ |\eta_1(t)| &\le \beta_\eta(|\eta(0)|, t) + \frac{a_2}{g^d}C_\phi + a_2|\bar e|_{[0,t)} \end{aligned} \tag{2.32}$$

With ϵ be any small number so that

$$\epsilon < \epsilon_1^\star := a_2,$$

being a_2 the same as in (2.32), let K be the corresponding matrix produced by Lemma 2.1, and change variables according to

$$\zeta \mapsto \bar{\zeta} := \zeta - K\chi. \tag{2.33}$$

Then, the bounds (2.29)-(2.30) hold, and Assumption 2.1 also yields

$$|(w(t), x(t))|_{\mathcal{A}} \leq \beta(|w(0), x(0)|_{\mathcal{A}}, t) + \rho\big((1+|K|)|\chi|_{[0,t)} + |\eta_1|_{[0,t)} + |\bar{\zeta}|_{[0,t)}\big) \tag{2.34}$$

Let

$$\mathcal{B} := \Big\{(w, x, \eta) \in \mathbb{R}^{n_w+n} \,:\, (w, x) \in \mathcal{A},\, \chi(w, x) = 0,\, \eta = 0\Big\},$$

then, in view of (2.29), (2.30), (2.32) and (2.34), the small-gain arguments of (Jiang et al., 1994) can be used to show that the subsystem (w, x, η) fulfills

$$|(w(t), x(t), \eta(t))|_{\mathcal{B}} \leq \beta_{\mathcal{B}}(|(w(0), x(0), \eta(0))|_{\mathcal{B}}, t) + \rho_{\mathcal{B}}\left(|\bar{\zeta}|_{[0,t)} + \frac{C_\phi}{g^d}\right), \tag{2.35}$$

for some $\beta_{\mathcal{B}} \in \mathcal{KL}$ and some locally Lipschitz $\rho_{\mathcal{B}} \in \mathcal{K}$.

In view of (2.10b), $\bar{\zeta}$ fulfills

$$\dot{\bar{\zeta}} = \varrho(\eta, \chi, \bar{\zeta}) + q(w, x) + \Omega(w, x)u \tag{2.36}$$

where

$$\varrho(\eta, \chi, \bar{\zeta}) := -K\Big((F + C_e^T G_1 C_e)\chi + H\zeta + C_e^T(\eta_2 - G_1\eta_1)\Big).$$

With $\ell > 0$ a design parameter to be fixed, in (2.13), let

$$\mathcal{K}_\xi := \ell K, \qquad \mathcal{K}_\zeta := -\ell I_{n_e}, \qquad \mathcal{K}_\eta := \ell K C_e^T. \tag{2.37}$$

In the new coordinates we have

$$u = \ell\mathcal{L}\big(K(\xi + C_e^T\eta_1) - \zeta\big) = -\ell\mathcal{L}\bar{\zeta}$$

and thus (2.36) yields

$$\dot{\bar{\zeta}} = \varrho(\eta, \chi, \bar{\zeta}) + q(w, x) - \ell B(w, x)\mathcal{L}\bar{\zeta}. \tag{2.38}$$

We fix ℓ on the basis of the following Lemma, whose proof is postponed at the end of this proof.

Lemma 2.2. *Consider an equation of the form* (2.38). *Under Assumption 2.2, for each compact set $W \times X \subset \mathbb{R}^{n_w+n}$ there exist $\mathcal{L}$ and $\ell_2^\star(g) > \ell_1^\star$ such that, for all $\ell > \ell_2^\star(g)$ and as long as $(w, x) \in W \times X$, the following holds*[4]

$$|\bar{\zeta}(t)| \leq \beta_{\bar{\zeta}}(|\bar{\zeta}(0)|, t) + \frac{a_3}{\ell}\left(\left||(x, w)|_{\mathcal{A}}\right|_{[0,t)} + |\chi|_{[0,t)} + |\eta|_{[0,t)} + \sup_{(w,x)\in\mathcal{A}} |q(w, x)| \right) \tag{2.39}$$

for some $\beta_{\bar{\zeta}} \in \mathcal{KL}$ and $a_3 > 0$.

In view of Lemma 2.2, by quite standard arguments (in this respect see for instance (Byrnes et al., 2003; Isidori, 1995, 1999)) based again on the small-gain arguments of (Jiang et al., 1994) can be used to show that for each compact set $W \times X$ of initial conditions there exist $\beta_{\mathcal{C}} \in \mathcal{KL}$, $\rho_{\mathcal{C}} > 0$, $\mathcal{L}$ and $\ell_3^\star(g) > \ell_2^\star(g)$ such that, for all $\ell > \ell_3^\star(g)$, and with

$$\begin{aligned} \mathcal{C} &:= \mathcal{O} \times \{0\} \\ &= \Big\{ (w, x, \eta) \in \mathbb{R}^{n_w+n+n_e d} \; : \; (w, x) \in \mathcal{A}, \; \chi(w, x) = 0, \; \zeta(w, x) = 0\,, \eta = 0 \Big\}, \end{aligned}$$

the following bound holds

$$|(w(t), x(t), \eta(t))|_{\mathcal{C}} \leq \beta_{\mathcal{C}}(|(w(0), x(0), \eta(0))|_{\mathcal{C}}, t) + \frac{\rho_{\mathcal{C}}}{\ell}\left(1 + \frac{C_\phi}{g^d}\right). \tag{2.40}$$

and thus the closed-loop system is forward complete and uniformly bounded from $W \times X \times \mathbb{R}^{d \times n_e}$, and point 1 of the proposition follows.

II. Existence of a Steady State:
The equation (2.40) implies that, for each $H \subset \mathbb{R}^{dn_e}$ compact, the Ω-limit set $\Omega_{cl}(W \times X \times H)$ of the closed-loop system (2.8), (2.13) is compact, non-empty,

[4]We observe that the constant a_ζ and the function β_ζ might depend on ϵ produced by Lemma 2.1.

uniformly attractive from $W \times X \times H$ and invariant. Moreover, we can chose H and ℓ such that $\mathcal{D} := \Omega_{cl}(W\times X\times H) \subset W\times X\times H$, so that $\mathcal{D}$ is also asymptotically stable. Furthermore, in view of (2.40), given any $\varepsilon > 0$, then for large enough ℓ (say $\ell > \ell_4^\star(g) \geq \ell_3^\star(g)$) we have

$$|(w(t), x(t), \eta(t))|_{\mathcal{C}} \leq \beta_{\mathcal{C}}(|(w(0), x(0), \eta(0))|_{\mathcal{C}}, t) + \varepsilon,$$

so that also point 2 of the proposition holds.

III. Asymptotic Bound:
Note that in (2.37) $\mathcal{K}_\eta$ has the form (2.14), with $\mathcal{K}'_\eta = \ell\mathcal{Q}$, for some $\mathcal{Q} \in \mathbb{R}^{n_e\times n_e}$ that, as detailed in the proof of Lemma 2.1, is invertible. Moreover, in view of (2.40), $\ell_4^\star(g)$ can be taken such that, for all $\ell \geq \ell_4^\star(g)$, $\mathcal{D} \subset \mathcal{N}$. Therefore, by assumptions 2.3, 2.4, the functions λ_ξ, λ_ζ and υ are defined on $\mathcal{D}$ and satisfy (2.19), (2.20) and (2.21). Furthermore, as $\mathcal{K}_\xi$ and $\mathcal{K}_\zeta$ depend linearly on ℓ, the function υ can be bounded uniformly on ℓ.

Suppose now that $(w, x) \in \mathcal{N}$ and consider the change of variables:

$$\begin{aligned}
&\bar{e} \mapsto \tilde{e} := \bar{e} - \upsilon_1(w, x)\\
&\eta \mapsto \tilde{\eta} := \eta - \upsilon(w, x)\\
&\chi \mapsto \tilde{\chi} := \chi - C_e^T \upsilon_1(w, x) - \lambda_\xi(w, x)\\
&\bar{\zeta} \mapsto \tilde{\zeta} := \bar{\zeta} + K\big(\lambda_\xi(w, x) + C_e^T \upsilon_1(w, x)\big) - \lambda_\zeta(w, x)
\end{aligned}$$

noting that $\lambda_\xi^e = 0$ and $\lambda_\zeta^e = 0$, then we obtain

$$\tilde{e} = C_e\chi - \upsilon_1(w, x) = C_e\tilde{\chi} \tag{2.41}$$

and, in view of (2.28) and noting that $C_e\lambda_\xi = 0$, $FC_e^T = 0$ and $C_eC_e^T = I_{n_e}$, we also have

$$\dot{\tilde{\chi}} = (F + C_e^T G_1 C_e + HK)\tilde{\chi} + H\tilde{\zeta} + C_e^T(\tilde{\eta}_2 - G_1\tilde{\eta}_1). \tag{2.42}$$

Thus, Lemma 2.1 can be used to show that, for each $g > 0$ and each $\epsilon > 0$, K can be taken so that there exist $\beta_{\tilde{\chi}}, \beta_{\tilde{e}} \in \mathcal{KL}$ and $q_1 > 0$ such that

$$\begin{aligned}
|\tilde{\chi}(t)| &\leq \beta_\chi(|\tilde{\chi}(0)|, t) + q_1\Big(|\tilde{\zeta}|_{[0,t)} + |\eta|_{[0,t)}\Big)\\
|\tilde{e}(t)| &\leq \tilde{\beta}_e(|\tilde{\chi}(0)|, t) + \epsilon|\tilde{\zeta}|_{[0,t)} + \epsilon|\tilde{\eta}|_{[0,t)}.
\end{aligned} \tag{2.43}$$

Let

$$\tilde{\mu} := \Delta(g)^{-1}\tilde{\eta}, \tag{2.44}$$

the same argument used before in dealing with μ show that

$$\dot{\tilde{\mu}} = gM\tilde{\mu} + g^{1-d}E\big(\phi(\Delta(g)\tilde{\mu} + \upsilon(w,x)) - \upsilon_{d+1}(w,x)\big) + gRC_e\tilde{\chi} \tag{2.45}$$

Then, as ϕ is locally Lipschitz and bounded by C_ϕ, for some $L_\phi > 0$ it holds that

$$\begin{aligned}
&|\phi(\Delta(g)\tilde{\mu} + \upsilon(w,x)) - \upsilon_{d+1}(w,x)| \\
&\qquad \leq |\phi(\Delta(g)\tilde{\mu} + \upsilon(w,x)) - \phi(\upsilon(w,x))| + |\phi(\upsilon(w,x)) - \upsilon_{d+1}(w,x)| \\
&\qquad \leq \min\{2C_\phi, L_\phi g^{d-1}|\tilde{\mu}|\} + |\delta(w,x)| \\
&\qquad \leq L_\phi g^{d-1}|\tilde{\mu}| + |\delta(w,x)|.
\end{aligned}$$

where δ is defined as in (2.23). Let $\bar{\delta}$ be a smooth function defined on $\mathcal{N}$ such that

$$\begin{aligned}
\bar{\delta}(w,x) &= \delta|_{\mathcal{A}}(w,x) && \forall (w,x) \in \mathcal{A} \\
|\bar{\delta}(w,x)| &\leq \sup_{(w,x)\in\mathcal{A}} |\delta(w,x)| && \forall (w,x) \in \mathbb{R}^{n_w+n}.
\end{aligned} \tag{2.46}$$

Then the function $\delta(w,x) - \bar{\delta}(w,x)$ vanishes on $\mathcal{A}$, so that we have

$$\begin{aligned}
|\delta(w,x)| &= |\delta(w,x) - \bar{\delta}(w,x) + \bar{\delta}(w,x)| \leq |\delta(w,x) - \bar{\delta}(w,x)| + |\bar{\delta}(w,x)| \\
&\leq \gamma\big(|(w,x)|_{\mathcal{A}}\big) + \sup_{(w,x)\in\mathcal{A}} |\delta(w,x)|
\end{aligned}$$

for some $\gamma \in \mathcal{K}$ that can be taken locally Lipschitz in $\mathcal{N}$.

Standard high-gain arguments thus can be used to prove that there exists $g_1^\star > 0$ dependent on L_ϕ such that, for all $g > g_1^\star$, the following estimate holds

$$|\tilde{\mu}(t)| \leq \beta_{\tilde{\mu}}(|\tilde{\mu}(0)|, t) + \frac{q_2}{g^d}\left(||(w,x)|_{\mathcal{A}}|_{[0,t)} + \sup_{(w,x)\in\mathcal{A}} |\delta(w,x)|\right) + q_2|\tilde{e}|_{[0,t)}.$$

for some $\beta_{\tilde{\mu}} \in \mathcal{KL}$ and some $q_2 > 0$. Thus, by letting $\beta_{\tilde{\eta}} = g^{d-1}\beta_{\tilde{\mu}}$, using the fact that $|\tilde{\eta}| \leq g^{d-1}|\tilde{\mu}|$ and $|\tilde{\eta}_1| \leq |\tilde{\mu}|$, we also obtain that the system $\tilde{\eta}$ seen as a system with input $\tilde{e}$ and output $\tilde{\eta}_1$ fulfills

$$\begin{aligned}
|\tilde{\eta}(t)| &\leq \beta_{\tilde{\eta}}(|\tilde{\eta}(0)|, t) + \frac{q_2}{g^d}\left(||(w,x)|_{\mathcal{A}}|_{[0,t)} + \sup_{(w,x)\in\mathcal{A}} |\delta(w,x)|\right) + q_2 g^{d-1}|\tilde{e}|_{[0,t)} \\
|\tilde{\eta}_1(t)| &\leq \beta_{\tilde{\eta}}(|\tilde{\eta}(0)|, t) + \frac{q_2}{g^d}\left(||(w,x)|_{\mathcal{A}}|_{[0,t)} + \sup_{(w,x)\in\mathcal{A}} |\delta(w,x)|\right) + q_2|\tilde{e}|_{[0,t)}.
\end{aligned} \tag{2.47}$$

Since

$$\begin{aligned}
\xi &= \chi - C_e^T \eta_1 = \tilde{\chi} - C_e^T \tilde{\eta}_1 + \lambda_\xi(w,x) \\
\zeta &= \bar{\zeta} + K\chi = \tilde{\zeta} + K\tilde{\chi} + \lambda_\zeta(w,x)
\end{aligned} \tag{2.48}$$

and $C_e \lambda_\xi = \lambda_\xi^e = 0$, $C_e \lambda_\zeta = \lambda_\zeta^e = 0$, then

$$\begin{aligned}
|\xi^e| &\leq |\tilde{\chi}| + |\tilde{\eta}_1| \\
|\zeta^e| &\leq |\tilde{\zeta}| + |K||\tilde{\chi}|,
\end{aligned}$$

so that, for some $q_3 > 0$, inside $\mathcal{N}$ Assumption 2.1 gives

$$|(w(t),x(t))|_{\mathcal{A}} \leq \beta(|(w(0),x(0))|_{\mathcal{A}}, t) + q_3\big((1+|K|)|\tilde{\chi}|_{[0,t)} + |\tilde{\zeta}|_{[0,t)} + |\tilde{\eta}_1|_{[0,t)}\big). \tag{2.49}$$

Thus, in view of (2.43), (2.47) and (2.49), by choosing $g > g^\star$, where

$$g^\star := \max\{g_1^\star, \sqrt[d]{q_2 q_3}\} \tag{2.50}$$

and choosing K using Lemma 2.1 with such a choice of G_1 and with ϵ that satisfies

$$\epsilon < \epsilon^\star := \min\{\epsilon_1^\star, 1/q_2\},$$

then, as (2.43) does not depends on $|(w,x)|_{\mathcal{A}}$, applying the small-gain arguments of (Jiang et al., 1994) twice yields the existence of a $\tilde{\beta}_1 \in \mathcal{KL}$ and a constant $q_4 > 0$ such that, by letting

$$\nu(t) := \max\{|(w(t),x(t))|_{\mathcal{A}}, |\tilde{\chi}(t)|, |\tilde{\eta}(t)|\} \tag{2.51}$$

then in $\mathcal{N}$ the following bound holds:

$$\nu(t) \leq \tilde{\beta}_1(\nu(0), t) + \frac{q_4}{g^d} \sup_{(w,x)\in\mathcal{A}} |\delta(w,x)| + q_4|\tilde{\zeta}|_{[0,t)}. \tag{2.52}$$

For what concerns $\tilde{\zeta}$, instead, in view of (2.37) and (2.48), we have

$$
\begin{aligned}
u &= \mathcal{L}\Big(\mathcal{K}_\xi \xi + \mathcal{K}_\zeta \zeta + \mathcal{K}_\eta \eta_1\Big) \\
&= \mathcal{L}\Big(\mathcal{K}_\xi(\tilde{\chi} - C_e^T \tilde{\eta}_1 + \lambda_\xi(w,x)) + \mathcal{K}_\zeta(\tilde{\zeta} + K\tilde{\chi} + \lambda_\zeta(w,x)) + \mathcal{K}_\eta(\tilde{\eta}_1 + \upsilon_1(w,x))\Big) \\
&= \mathcal{L}\Big(\mathcal{K}_\xi(\tilde{\chi} - C_e^T \tilde{\eta}_1) + \mathcal{K}_\zeta(\tilde{\zeta} + K\tilde{\chi}) + \mathcal{K}_\eta \tilde{\eta}_1\Big) \\
&\qquad\qquad + \mathcal{L}\Big(\mathcal{K}_\xi \lambda_\xi(w,x) + \mathcal{K}_\zeta \lambda_\zeta(w,x) + \mathcal{K}_\eta \upsilon_1(w,x)\Big) \\
&= \ell\mathcal{L}\Big(K(\tilde{\chi} - C_e^T \tilde{\eta}_1) - I_{n_e}(\tilde{\zeta} + K\tilde{\chi}) + KC_e^T \tilde{\eta}_1\Big) \\
&\qquad\qquad + \mathcal{L}\Big(\mathcal{K}_\xi \lambda_\xi(w,x) + \mathcal{K}_\zeta \lambda_\zeta(w,x) + \mathcal{K}_\eta \upsilon_1(w,x)\Big) \\
&= \ell\mathcal{L}\tilde{\zeta} + \mathcal{L}\Big(\mathcal{K}_\xi \lambda_\xi(w,x) + \mathcal{K}_\zeta \lambda_\zeta(w,x) + \mathcal{K}_\eta \upsilon_1(w,x)\Big).
\end{aligned}
$$

As a consequence, in view of (2.19), with ϱ the same linear map as in (2.36), we have

$$
\begin{aligned}
\dot{\tilde{\zeta}} &= \dot{\zeta} - K\dot{\tilde{\chi}} - \dot{\lambda}_\zeta(w,x) \\
&= \varrho(\tilde{\eta}, \tilde{\chi}, \tilde{\zeta}) + q(w,x) + B(w,x)u \\
&\qquad - \bigg(q(w,x) + B(w,x)\mathcal{L}\Big(\mathcal{K}_\xi \lambda_\xi(w,x) + \mathcal{K}_\zeta \lambda_\zeta(w,x) + \mathcal{K}_\eta \upsilon_1(w,x)\Big)\bigg) + \varphi(w,x) \\
&= \varrho(\tilde{\eta}, \tilde{\chi}, \tilde{\zeta}) + \varphi(w,x) - \ell B(w,x)\mathcal{L}\tilde{\zeta}.
\end{aligned}
$$

Therefore, the same arguments as in Lemma 2.2 applied with $q = \varphi$ (which, in view of Assumption 2.4 is locally Lipschitz and satisfies $\varphi|_{\mathcal{A}} = 0$), show that there exists $\ell^\star(g) \geq \ell_4^\star(g)$ such that, for all $\ell \geq \ell^\star(g)$, inside $\mathcal{N}$ it holds that

$$
|\tilde{\zeta}(t)| \leq \beta_{\tilde{\zeta}}(|\tilde{\zeta}(0)|, t) + \frac{q_5}{\ell}|\nu|_{[0,t)}. \tag{2.53}
$$

and with $q_5 > 0$ and $\beta_{\tilde{\zeta}} \in \mathcal{KL}$ such that

$$
\frac{q_5 q_4}{\ell} < 1.
$$

Therefore, standard small-gain arguments show that inside $\mathcal{N}$, the following bound holds:

$$
\limsup_{t\to\infty} \max\{\nu(t), |\tilde{\zeta}(t)|\} \leq \frac{q_6}{g^d} \sup_{(w,x)\in\mathcal{A}} |\delta(w,x)|. \tag{2.54}
$$

for some $q_6 > 0$.

Now, to establish the bound (2.25), pick a point (w, x, η) in $\mathcal{D}$. Then there exist sequences $((w^n, x^n, \eta^n))_n$ in $\mathcal{S}_{\mathcal{H}_{cl}}(W \times X \times H)$ and $(t_n)_n$ in $\mathbb{R}_+$ such that $t_n \to \infty$ and

$$(w^n(t_n), x^n(t_n), \eta^n(t_n)) \to (w, x, \eta). \tag{2.55}$$

From point 2 of the proposition, proved above, for sufficiently large n we have $(w^n(t_n), x^n(t_n), \eta^n(t_n)) \in \mathcal{N}$, and, hence, (2.54) holds. For each n, let $\tilde{\chi}^n := \tilde{\chi}(w^n, x^n)$, $\tilde{\zeta}^n := \tilde{\zeta}(w^n, x^n)$ and $\tilde{\eta}^n := \eta^n - \upsilon(w^n, x^n)$. Then, the error at (w, x) satisfies

$$e = h_e(w, x) = C_e\xi(w, x) = C_e\tilde{\chi}(w, x) - C_e C_e^T \tilde{\eta}_1,$$

so as

$$|e| \le |\tilde{\chi}| + |\tilde{\eta}_1| \le |\tilde{\chi} - \tilde{\chi}^n(t_n)| + |\tilde{\eta}_1 - \tilde{\eta}_1^n(t_n)| + |\tilde{\eta}_1^n(t_n)| + |\tilde{\chi}^n(t_n)|$$

As $\tilde{\chi}(w, x)$ is continuous in (w, x), then, in view of (2.54) and (2.55), for each $\varepsilon > 0$, there exists $n^\star(\varepsilon) \in \mathbb{N}$ such that $n \ge n^\star(\varepsilon)$ implies

$$\begin{aligned} |\tilde{\chi} - \tilde{\chi}^n(t_n)| &= |\tilde{\chi}(w, x) - \tilde{\chi}(w^n(t_n), x^n(t_n))| \le \varepsilon/2 \\ |\tilde{\eta}_1^n(t_n)| + |\tilde{\chi}^n(t_n)| &\le \varepsilon/2 + \frac{q_6}{g^d} \sup_{(w,x)\in\mathcal{A}} |\delta(w, x)|. \end{aligned}$$

By arbitrariness of ε and $(w, x, \eta) \in \mathcal{D}$, we then conclude that

$$\sup_{(w,x,\eta)\in\mathcal{D}} |h_e(w, x)| \le \frac{q_6}{g^d} \sup_{(w,x)\in\mathcal{A}} |\delta(w, x)|, \tag{2.56}$$

that, in view of the uniform attractiveness of $\mathcal{D}$, implies (2.25). ■

Proof of Lemma 2.1. Pick $i \in \{1, \dots, r\}$ and, with $k_i > 0$, define the matrix $\Lambda_i(k_i) := \mathrm{diag}(k_i^{N_i-2} I_{p_i}, k_i^{N_i-3} I_{p_i}, \dots, k_i I_{p_i}, I_{p_i})$ and change coordinates as

$$\chi^i \mapsto z^i := \Lambda_i(k_i)\chi^i\,.$$

With reference to the matrices defined in (2.10a)-(2.10c), noting that

$$\Lambda_i(k_i) F_{ii} \Lambda_i(k_i)^{-1} = k_i F_{ii}$$

$$\begin{aligned}
&\Lambda_i(k_i)H_{ii} = H_{ii} \\
&\Lambda_i(k_i)C_i^T G_1 C_i \Lambda_i(k_i)^{-1} = C_i^T G_1 C_i \\
&\Lambda_i(k_i)C_i^T = k_i^{N_i-2} C_i^T
\end{aligned}$$

then z^i fulfills

$$\begin{aligned}
\dot{z}^i = k_i F_{ii} z^i + H_{ii}\zeta_i + \Lambda_i(k_i) \sum_{j=1}^{i-1} \left(F_{ij}\Lambda_j(k_j)^{-1} z^j + H_{ij}\zeta_j\right) \\
+ C_i^T G_1 C_i z^i + k_i^{N_i-2} C_i^T (\eta_2 - G_1\eta_1)
\end{aligned}$$

Let $(\alpha_1^i, \dots, \alpha_{N_i-1}^i) \in \mathbb{HC}(N_i - 1)$ and, with $D^i := \begin{pmatrix} -\alpha_1^i I_{m_i} & \cdots - \alpha_{N_i-1}^i I_{m_i} \end{pmatrix}$ and $\bar{\zeta}^i \in \mathbb{R}^{m_i}$, let

$$\zeta^i = \bar{\zeta}^i + k_i D^i z^i \,. \tag{2.57}$$

Since $F_{ii} + H_{ii}D^i$ is Hurwitz, it can be shown that there exists $k^\star > 1$ such that, for any $k_i > k^\star$, the following holds

$$\begin{aligned}
|z^i(t)| \le b_1 e^{-b_2 k_i t}|z^i(0)| + b_3 k^{N_i-3}\Big(|\bar{\zeta}|_{[0,t)} + |\eta|_{[0,t)}\Big) \\
+ b_4 k_i^{N_i-3} \sum_{j=1}^{i-1} k_j \int_0^t e^{-b_2 k_i (t-\tau)} |z^j(\tau)| d\tau
\end{aligned} \tag{2.58}$$

for some $b_1, b_2, b_3, b_4 > 0$. We can partition $\bar{e}$ as $\bar{e} = \text{col}(\bar{e}^1, \dots, \bar{e}^{r_e})$, with $\bar{e}^i := k_i^{2-N_i} C_i z^i$. Pick $\epsilon > 0$ and pick $i \in \{1, \dots, r\}$. Suppose that, for each $j = 1, \dots, i-1$, $z^j(t)$ fulfills

$$|z^j(t)| \le h_1^i e^{-h_2^i t}|z(0)| + h_3^i \Big(|\bar{\zeta}|_{[0,t)} + |\eta|_{[0,t)}\Big) \tag{2.59}$$

for some $h_1^i, h_2^i, h_3^i \ge 0$. Then, by letting for convenience $\bar{k}_i := \max_{1\le j<i} k_j$, (2.58)

gives

$$
\begin{aligned}
|z^i(t)| \le{}& b_1 e^{-b_2 k_i t}|z(0)| + b_3 k^{N_i-3}\Big(|\bar{\zeta}|_{[0,t)} + |\eta|_{[0,t)}\Big) \\
&+ b_4 k_i^{N_i-3}(i-1)\bar{k}_i\bigg(\int_0^t e^{-b_2 k_i(t-\tau)} h_1^i |z(0)| d\tau \\
&+ \int_0^t e^{-b_2 k_i(t-\tau)} h_3^i \big(|\bar{\zeta}|_{[0,\tau)} + |\eta|_{[0,\tau)}\big) d\tau\bigg) \\
\le{}& \left(b_1 + \frac{2b_4 r \bar{k}_i h_1^i k^{N_i-4}}{b_2}\right) e^{-b_2 k_i t}|z(0)| \\
&+ \left(b_3 + \frac{2b_4 r \bar{k}_i h_3^i}{b_2 k_i}\right) k_i^{N_i-3}(|\tilde{\zeta}|_{[0,t)} + |\tilde{\eta}|_{[0,t)})
\end{aligned} \tag{2.60}
$$

Fix

$$
k_i = \max\left\{k^\star,\ \max_{1\le j<i} k_j,\ \frac{r}{\epsilon}\left(b_3 + \frac{2b_4 r h_3^i}{b_2}\right)\right\}. \tag{2.61}
$$

Then, in view of (2.60), the fact that (2.59) holds for $j = 1, \ldots, i-1$ implies that the same bound also holds for $j = 1, \ldots, i$, with

$$
\begin{aligned}
h_1^{i+1} &:= \max\left\{h_1^i, b_1 + \frac{2b_4 r h_1^i k^{N_i-3}}{b_2}\right\} \\
h_2^{i+1} &:= \min\left\{b_2 k_i, h_2^i\right\}, \\
h_3^{i+1} &:= \max\left\{h_3^i, \frac{r}{\epsilon} k^{N_i-2}\right\}
\end{aligned}
$$

Moreover, noting that $|\bar{e}^i(t)| \le k_i^{2-N_i}|z^i(t)|$, in view of (2.61), we have

$$
|\bar{e}^i(t)| \le q_1^i e^{-q_2^i t}|z(0)| + \frac{\epsilon}{r}\big(|\bar{\zeta}|_{[0,t)} + |\eta|_{[0,t)}\big) \tag{2.62}
$$

with $q_1^i := h_1^{i+1} k_i^{2-N_i}$ and $q_2^i := h_2^{i+1}$. Fix k_1 so that

$$
k_1 \ge \max\left\{k^\star, \frac{b_3 r}{\epsilon}\right\}.
$$

Then the sequence (2.61) is well-defined and (2.58) implies (2.59) for $j < 2$ and (2.62) for $i = 1$, with $h_1^2 := b_1$, $h_2^2 := -b_2 k_1$ and $q_1^1 := b_1 k^{2-N_1}$, $q_2^1 := b_2 k_1$. Hence, by induction we conclude that, for each $i = 1, \ldots, r$

$$
|z^i(t)| \le \bar{h}_1 e^{-\bar{h}_2 t}|z(0)| + \bar{h}_3\Big(|\bar{\zeta}|_{[0,t)} + |\eta|_{[0,t)}\Big)
$$

and for each $i = 1, \dots, r_e$,

$$|\bar{e}^i(t)| \le \bar{q}_1 e^{-\bar{q}_2 t}|z(0)| + \frac{\epsilon}{r}\big(|\bar{\zeta}|_{[0,t)} + |\eta|_{[0,t)}\big)$$

with $\bar{h}_1 := h_1^{r+1}$, $\bar{h}_2 := h_2^{r+1}$, $\bar{h}_3 := h_3^{r+1}$, $\bar{q}_1 := q_1^r$ and $\bar{q}_2 := q_2^r$. Noting that $|\chi(t)| \le \sum_{i=1}^r |\chi^i(t)| \le \sum_{i=1}^r |z^i(t)|$, $|z(0)| \le k^{N_r-1}|\chi(0)|$ and $|\bar{e}(t)| \le \sum_{i=1}^{r_e} |\bar{e}^i(t)|$, then we obtain (2.29)-(2.30), with

$$\begin{aligned} &\beta_\chi(s,t) := k^{N_r-1} r \bar{h}_1 \exp(-\bar{h}_2 t)s && a_5 := k^{N_r-1} r \bar{h}_3 \\ &\beta_{\bar{e}}(s,t) := r_e \bar{q}_1 \exp(-\bar{q}_2 t)s \end{aligned}$$

and the claim follows with

$$K := \Big(k_1 D^1 \Lambda_1(k_1) \quad \cdots \quad k_r D^r \Lambda_r(k_r)\Big) .$$

■

Proof of Lemma 2.2. Fix the compact sets $W \times X \subset \mathbb{R}^{n_w+n}$ and define the function

$$V(w,x) = \sqrt{\bar{\zeta}^T \mathcal{P}(w,x) \bar{\zeta}} \tag{2.63}$$

on a neighborhood of $W \times X$. Point a of Assumption 2.2 implies the existence of $\underline{\lambda}, \bar{\lambda} > 0$ such that

$$\underline{\lambda}|\bar{\zeta}| \le V(w,x) \le \bar{\lambda}|\bar{\zeta}|$$

for all $(w,x) \in W \times X$. Taking the Dini derivative of V along the solutions of the closed-loop system yields

$$\begin{aligned} D^+V(w,x) = \frac{1}{2V(w,x)}\Big(&-\ell\bar{\zeta}^T\Big(\mathcal{L}^T B(w,x)^T \mathcal{P}(w,x) + \mathcal{P}(w,x)B(w,x)\mathcal{L}\Big)\bar{\zeta} \\ &+ 2\bar{\zeta}^T\mathcal{P}(w,x)\big(\rho(\eta,\chi,\bar{\zeta}) + q(w,x)\big) \\ &+ \bar{\zeta}\big(L_s^{(w)}\mathcal{P}(w,x) + L_f^{(x)}\mathcal{P}(w,x) + L_{b(w,x)u}^{(x)}\mathcal{P}(w,x)\big)\bar{\zeta}\Big). \end{aligned}$$

Point c of Assumption 2.2 implies

$$\bar{\zeta}^T\Big(\mathcal{L}^T B(w,x)^T \mathcal{P}(w,x) + \mathcal{P}(w,x)B(w,x)\mathcal{L}\Big)\bar{\zeta} \ge |\bar{\zeta}|^2$$

so as

$$-\ell\bar{\zeta}^T\Big(\mathcal{L}^T B(w,x)^T\mathcal{P}(w,x) + \mathcal{P}(w,x)B(w,x)\mathcal{L}\Big)\bar{\zeta} \le -\ell|\bar{\zeta}|^2.$$

Let $\bar{q}(w,x)$ be any function that agrees with q on $\mathcal{A}$ and that satisfies $|q(w,x)| \le \sup_{(w,x)\in\mathcal{A}} |q(w,x)|$ for all $(w,x) \in \mathbb{R}^{n_w+n}$. Adding and subtracting $2\bar{\zeta}^T\mathcal{P}(w,x)\bar{q}(w,x)$ yields

$$2\bar{\zeta}^T\mathcal{P}(w,x)\big(\rho(\eta,\chi,\bar{\zeta}) + q(w,x)\big) = 2\bar{\zeta}^T\mathcal{P}(w,x)\big(\rho(\eta,\chi,\bar{\zeta}) + \tilde{q}(w,x) + \bar{q}(w,x)\big)$$

with $\tilde{q} := q - \bar{q}$. As $\mathcal{P}$ is continuous, ρ is Lipschitz and, as $\tilde{q}$ vanishes on $\mathcal{A}$ and it is locally Lipschitz, then for some $M_1(g) > 0$ we have

$$\begin{aligned}&2\bar{\zeta}^T\mathcal{P}(w,x)\big(\rho(\eta,\chi,\bar{\zeta}) + q(w,x)\big)\\ &\quad\le M_1(g)|\bar{\zeta}|\Big(|\eta| + |\chi| + |\bar{\zeta}| + |(w,x)|_{\mathcal{A}} + \sup_{(w,x)\in\mathcal{A}} |q(w,x)|\Big),\end{aligned}$$

as long as $(w,x) \in W\times X$. Point b of Assumption 2.2 implies that $L^{(x)}_{b(w,x)u}\mathcal{P}(w,x) = 0$, so as by continuity of $\mathcal{P}$, as long as $(w,x) \in W \times X$, we can write

$$\bar{\zeta}\big(L^{(w)}_s\mathcal{P}(w,x) + L^{(x)}_f\mathcal{P}(w,x) + L^{(x)}_{b(w,x)u}\mathcal{P}(w,x)\big)\bar{\zeta} \le M_2|\bar{\zeta}|^2$$

for some $M_2 > 0$. Since

$$\frac{1}{\bar{\lambda}} \le \frac{|\bar{\zeta}|}{V(w,x)} \le \frac{1}{\underline{\lambda}},$$

then there exist $\alpha_1, \alpha_2(g) > 0$ such that, as long as $(w,x) \in W \times X$, we have

$$D^+V(w,x) \le (\alpha_2(g) - \ell\alpha_1)V(w,x) + \alpha_2(g)\Big(|\eta| + |\chi| + |(w,x)|_{\mathcal{A}} + \sup_{(w,x)\in\mathcal{A}} |q(w,x)|\Big),$$

and the result thus follows by taking, for some arbitrary $\epsilon \in (0,\alpha_1)$, $\ell^\star_2(g) := \max\{\ell^\star_1, \alpha_2(g)/(\alpha_1 - \epsilon)\}$. ■

2.3.5 On the Controllability Assumption

Although Assumption 2.2 might seem to be constructed ad-hoc for the stabilization problem inside Proposition 2.1, it turns out that it is general enough to be actually implicated by many state-of-art assumptions routinely used in the context of high-gain stabilization and regulation of multivariable systems. In the following we prove this fact for some relevant papers appeared in the literature. In this respect we notice that the following results give constructive procedures to define the matrix $\mathcal{L}$ ($\mathcal{P}$ is not required to be known as it is not used by the control) using only quantities that are known in the respective frameworks. In the following we assume that $B(w,x)$ is C^1 and, for ease of notation, we let $\mathbf{x} := (w,x)$.

Strong Invertibility in the sense of (Wang et al., 2015a) implies Assumption 2.2

Here we prove that the assumption of invertibility used, for instance, in the recent papers (Wang et al., 2015a) and (Wang et al., 2016), implies Assumption 2.2.

Lemma 2.3. *Suppose that $m = n_y$ (i.e. $B(\mathbf{x})$ is square), $B(\mathbf{x})$ is bounded, $L^{(x)}_{g(\mathbf{x})u} B(\mathbf{x}) = 0$ for all $\mathbf{x} \in W \times \mathbb{R}^n$, and there exists $\epsilon > 0$ such that all its principal minors $\Delta_i(\mathbf{x})$, $i = 1, \dots, r$ satisfy*

$$|\Delta_i(\mathbf{x})| \geq \epsilon, \tag{2.64}$$

for all $\mathbf{x} \in W \times \mathbb{R}^n$. Then Assumption 2.2 holds.

Proof. According to (Wang et al., 2015a, Lemma 1), if (2.64) holds then $B(\mathbf{x})$ can be written as

$$B(\mathbf{x}) = EM(\mathbf{x})(I + U(\mathbf{x})),$$

with $M(\mathbf{x}) = M(\mathbf{x})^T$ positive definite for all $\mathbf{x} \in W \times \mathbb{R}^n$, $U(\mathbf{x})$ a strictly upper triangular matrix and with E a diagonal matrix satisfying $EE = I$. Lemma 2 of (Wang et al., 2015a) shows that, if $B(\mathbf{x})$ is bounded there exists $c > 1$ such that, with $C := \text{diag}(c^{m-1}, c^{m-2}, \dots, c, 1)$, we have

$$(I + U(\mathbf{x}))C + C(I + U(\mathbf{x}))^T \geq I.$$

Let

$$\mathcal{L} := EC, \qquad \mathcal{P}(\mathbf{x}) := EM(\mathbf{x})^{-1}E,$$

then, noting that

$$B(\mathbf{x})^T EM(\mathbf{x})^{-1} = (I + U(\mathbf{x}))^T M(\mathbf{x})^T E^T EM(\mathbf{x})^{-1} = (I + U(\mathbf{x}))^T$$

then we have (note that by construction $\mathcal{L} = \mathcal{L}^T = EC = CE$)

$$\begin{aligned}\mathcal{L}^T B(\mathbf{x})^T \mathcal{P}(\mathbf{x}) + \mathcal{P}(\mathbf{x}) B(\mathbf{x}) \mathcal{L} &= ECB(\mathbf{x})^T EM(\mathbf{x})^{-1} E + EM(\mathbf{x})^{-1} EB(\mathbf{x}) CE \\ &= E\Big[C(I + U(x))^T + (I + U(x))C\Big]E \\ &\geq I,\end{aligned}$$

for all $\mathbf{x} \in W \times \mathbb{R}^n$. Thus, since $EE = I$, point c of Assumption 2.2 holds. Furthermore, as $M(\mathbf{x}) > 0$, then $M(\mathbf{x})^{-1} > 0$ as well and, hence, $\mathcal{P}(\mathbf{x})$ is positive definite and point a holds. Finally, according to (Wang et al., 2015a, Lemma 3), $L^{(x)}_{g(\mathbf{x})u} B(\mathbf{x}) = 0$ implies $L^{(x)}_{g(\mathbf{x})u} M(\mathbf{x}) = 0$. Since

$$L^{(x)}_{g(\mathbf{x})u} M(\mathbf{x})^{-1} = -M(\mathbf{x})^{-1} L^{(x)}_{g(\mathbf{x})u} M(\mathbf{x}) M(\mathbf{x})^{-1} = 0,$$

then also point b holds, hence the result. ■

Strong invertibility in the sense of (Wang et al., 2015b, 2017) implies Assumption 2.2

Here we prove that the assumption of invertibility used, for instance, in the recent papers (Wang et al., 2015b) and (Wang et al., 2017), implies Assumption 2.2.

Lemma 2.4. *Suppose that $m = n_y$ (i.e. $B(\mathbf{x})$ is square) and that there exist a nonsingular matrix $M \in \mathbb{R}^{m \times m}$ and a constant $\delta_0 \in (0, 1)$ such that*

$$\max_{\substack{\Lambda \in \mathbb{R}^{m \times m} \\ |\Lambda| \leq 1}} \left| \Big(B(\mathbf{x}) - M \Big) \Lambda M^{-1} \right| \leq \delta_0 \tag{2.65}$$

holds for all $\mathbf{x} \in W \times \mathbb{R}^n$, then Assumption 2.2 holds.

Proof. As (2.65) holds for all $\Lambda \in \mathbb{R}^{m \times m}$ satisfying $|\Lambda| \leq 1$, it holds in particular for $\Lambda = I$, thus yielding

$$|B(\mathbf{x}) M^{-1} - I| \leq \delta_0.$$

Thus, for all $\mathbf{p} \in \mathbb{R}^m$ and $\mathbf{x} \in W \times \mathbb{R}^n$, it holds that

$$\begin{aligned} 2\mathbf{p}^T \Big(I - B(\mathbf{x}) M^{-1} \Big) \mathbf{p} \leq |2\mathbf{p}^T (B(\mathbf{x}) M^{-1} - I) \mathbf{p}| &\leq 2|\mathbf{p}|^2 \cdot |I - B(\mathbf{x}) M^{-1}| \\ &\leq 2\delta_0 |\mathbf{p}|^2 = \mathbf{p}^T (2\delta_0 I) \mathbf{p}. \end{aligned}$$

Therefore, we obtain

$$\begin{aligned}\mathbf{p}^T\Big(2I - M^{-T}B(\mathbf{x})^T - B(\mathbf{x})M^{-1}\Big)\mathbf{p} &= 2\mathbf{p}^T\mathbf{p} - 2\mathbf{p}^T B(\mathbf{x})M^{-1}\mathbf{p} \\ &= 2\mathbf{p}\Big(I - B(\mathbf{x})M^{-1}\Big)\mathbf{p} \le \mathbf{p}^T(2\delta_0 I)\mathbf{p}.\end{aligned}$$

As it holds for all $\mathbf{p} \in \mathbb{R}^m$ and $\mathbf{x} \in W \times \mathbb{R}^n$ then, necessarily

$$2I - M^{-T}B(\mathbf{x})^T - B(\mathbf{x})M^{-1} \le 2\delta_0 I$$

and, thus, letting $\mathcal{L} := M^{-1}$ and $\mathcal{P}(\mathbf{x}) := I/(2(1-\delta_0))$, yields the point c of Assumption 2.2. Point a of the assumption follows by noting that $\delta_0 \in (0,1)$ and point b is straightforward as $\mathcal{P}$ is constant in $\mathbf{x}$. ■

Positivity and negativity in the sense of (McGregor et al., 2006; Back, 2009; Astolfi et al., 2013) imply Assumption 2.2

Going back few years we find other three papers in which a "negativity" or a "positivity" assumption on B is made. The following lemma refers to (McGregor et al., 2006, Ass. 4.4).

Lemma 2.5. *Suppose that $m = n_y$ (i.e. $B(\mathbf{x})$ is square) and that there exists $M \in \mathbb{R}^{n_y \times n_y}$ such that the following negativity condition holds:*

$$B(\mathbf{x})M + M^T B(\mathbf{x})^T < 0 \tag{2.66}$$

for all $\mathbf{x} \in W \times \mathbb{R}^n$. Then Assumption 2.2 holds.

Proof. Let $\delta(\mathbf{x}) := \min \sigma(B(\mathbf{x})M + M^T B(\mathbf{x}))$. Equation (2.66) implies, for any $\mathbf{p} \in \mathbb{R}^m$ and $\mathbf{x} \in W \times \mathbb{R}^n$,

$$\mathbf{p}^T(\delta(\mathbf{x})I)\mathbf{p} = \delta(\mathbf{x})|\mathbf{p}|^2 \le \mathbf{p}^T(B(\mathbf{x})M + M^T B(\mathbf{x})^T)\mathbf{p},$$

i.e.

$$M^T B(\mathbf{x})^T + B(\mathbf{x})M \ge \delta(\mathbf{x})I.$$

Thus Assumption 2.2 holds with $\mathcal{L} := M$ and $\mathcal{P}(\mathbf{x}) := I/\delta(\mathbf{x})$. ■

The following results instead refers to the positivity assumption of (Astolfi et al., 2013, Assumption 1).

Lemma 2.6. *Suppose that $m = n_y$ (i.e. $B(\mathbf{x})$ is square) and that there exists $K \in \mathbb{R}^{n_y \times n_y}$ such that the following positivity condition holds:*

$$B(\mathbf{x})K + K^T B(\mathbf{x})^T \geq I \tag{2.67}$$

for all $\mathbf{x} \in W \times \mathbb{R}^n$. Then Assumption 2.2 holds.

Proof. The proof follows by noting that the hypotheses of Lemma 2.5 hold with $M = -K$. ■

Finally, the following lemma shows that also the quite intricate condition of (Back, 2009, Ass. 3) implies Assumption 2.2.

Lemma 2.7. *Suppose that $m = n_y$ (i.e. $B(\mathbf{x})$ is square), and assume that there exist a nonsingular matrix K, $G^- := \operatorname{diag}(g_1^-, \ldots, g_m^-)$ and $G^+ := \operatorname{diag}(g_1^+, \ldots, g_m^+)$ such that $0 < G^- < G^+$ and that*

$$\Big(B(\mathbf{x})K\mathbf{p} - G^-\mathbf{p}\Big)^T \Pi^2 \Big(B(\mathbf{x})K\mathbf{p} - G^+\mathbf{p}\Big) \leq 0, \tag{2.68}$$

for all $\mathbf{p} \in \mathbb{R}^m$ and all $\mathbf{x} \in W \times \mathbb{R}^n$ and where $\Pi := 2(G^+ + G^-)^{-1}$. Then Assumption 2.2 holds.

Proof. Equation (2.68) implies ($G^- = (G^-)^T$)

$$-K^T B(\mathbf{x})^T \Pi^2 G^+ - G^- \Pi^2 B(\mathbf{x})K + K^T B(\mathbf{x})^T \Pi^2 B(\mathbf{x})K + G^- \Pi^2 G^+ \leq 0,$$

that in turn implies

$$\mathcal{M}(\mathbf{x}) := K^T B(\mathbf{x})^T \Pi^2 G^+ + G^- \Pi^2 B(\mathbf{x})K > 0$$

for all $\mathbf{x} \in W \times \mathbb{R}^n$. Noting that

$$\Pi^2 G^+ = \Pi \cdot 2(G^+ + G^-)^{-1} G^+ = \Pi \cdot 2(G^+ + G^-)^{-1}(G^+ + G^- - G^-) = 2\Pi - \Pi^2 G^-$$
$$G^- \Pi^2 = 2 \cdot G^-(G^+ + G^-)^{-1}\Pi = 2\Pi - G^+\Pi^2$$

then

$$\mathcal{M}(\mathbf{x}) = 2\Big(K^T B(\mathbf{x})^T \Pi + \Pi B(\mathbf{x})K\Big) - \mathcal{M}(\mathbf{x})^T$$

Thus, $\mathcal{M} > 0$ implies

$$K^T B(\mathbf{x})^T \Pi + \Pi B(\mathbf{x}) K = \frac{1}{2}\Big(\mathcal{M}(\mathbf{x}) + \mathcal{M}(\mathbf{x})^T\Big) > 0$$

and the claim follows with the same arguments of Lemma 2.5. ■

3

Robustness in Output Regulation

The most celebrated property of the linear regulator (see property P1 in Section 1.1.3) is *robustness* to plant's uncertainties. Namely, if the internal model unit is appropriately chosen, asymptotic regulation is ensured despite plant's uncertainties that do not destroy *linearity* and *closed-loop stability*, with the stabilizer chosen on the basis of the plant's nominal value. The whole nonlinear regulation theory developed so far (the regulators mentioned in Section 1.2.3 included) failed to extend, in its full generality, this robustness property and robustness itself, quite surprisingly, has been almost left out by the majority of the output regulation literature of the last 20 years. This chapter deals with the robustness issue in output regulation schemes. We fist analyze, by means of a quite informal discussion, the reason why robustness is such a far concept for nonlinear systems (Section 3.1). Then we present a (pre-processing) regulator design based on *low-power high-gain observers* that extends to a class of nonlinear systems the "structural robustness" properties of the local approach of (Byrnes et al., 1997a) (Section 3.2). Then we present a framework in which the usual notions of steady state and zero dynamics, as originally introduced in

(Byrnes and Isidori, 2003), can be extended to the case in which the exosystem is a differential inclusion, thus shifting the robustness issue to the exosystem definition (Section 3.3). Finally, we present a new framework in which a generalized notion of robustness with respect to arbitrary topologies is defined relative to arbitrary steady-state properties of the closed-loop trajectories (Section 3.4). We review in this framework different existing regulators and we present new results about robustness of post-processing schemes.

3.1 The Robustness Issue

This section contains some original considerations extracted by the tutorial paper (Bin and Marconi, 2018b). For further details see also (Bin et al., 2018a) and (Bin and Marconi, 2018a). While for linear systems the concept of robustness has a clear and agreed meaning, for nonlinear system robustness is still quite a vague concept that usually is treated in ad hoc manners. The reason of this fact is perhaps that the type of plant's perturbations captured by the concept of *"structural robustness"* originally given by Francis and Wonham in (Francis and Wonham, 1975), that refers to perturbations obtained by changing the matrices entries, for linear plants are general enough to include more "exotic" types of perturbations, such as those framed in the context of differential topology (Hirsch, 1994). For nonlinear plants, however, parametric uncertainties are way far to be sufficient to describe the whole set of perturbations that may affect the plant's functions, and a general unifying concept of perturbation is probably not yet taken into consideration. The first frameworks developed for nonlinear systems (see e.g. Byrnes et al., 1997a,b), as well as the majority of the subsequent designs, just focused on the extension of the parametric notion of perturbation and on the corresponding generalization of the concept of "structural robustness", with few exceptions such as (Astolfi et al., 2015; Astolfi and Praly, 2017) in which the C^1 topology was considered.

This considerable gap in the characterization of robustness that is present between linear and nonlinear systems, led us to wonder two questions: *What would be the right way to extend the notion of "structural robustness" to nonlinear systems?* and *Why robustness seems so far in nonlinear output regulation?* While the first section is answered thoroughly in Section 3.4, we focus here on an informal simple answer to the second, more conceptual, question.

The key point that makes linear systems so special is that they obey the superposition principle and, as such, they do not *distort* linear input signals (i.e. signals generated by linear exosystems). With reference to Section 1.1.3, superposition principle means that if the linear stabilized plant (1.5), (1.13), (1.15) is asymptotically stable, and it is excited by the input w produced by a linear exosystem (1.7), at the steady state all the signals in the closed-loop will contain exclusively the *same* harmonics of $w(t)$. In other words, all the signals in the closed-loop systems can be generated by the same linear process that generated $w(t)$, i.e. (1.7). As made clear in the nonlinear frameworks (and that is true also in the linear case), *the internal model should not be a model of the exosystem. Rather it should be a model of any process that can generate the ideal error-zeroing control law $u^\star(t)$*. The superposition principle implies that for linear systems $u^\star(t)$ can be generated by any process that has the same modes of the exosystem, and hence, the fact that Φ replicates S turns out to be sufficient to ensure that η has the internal model property. This, in turn, is also the reason why the same linear regulator holds for *any* choice of the matrices P and Q in (1.5). For nonlinear systems this fortunate relation between internal model and exosystem simply does not hold, and seeking or trying to force it in nonlinear contexts results in reductive and astray approaches[1]. This can be easily seen by considering the following simple system[2]:

$$\begin{array}{rcl} \dot{x}_1 & = & x_2 + x_3 \\ \dot{x}_2 & = & -x_1 - \beta x_2 + \epsilon x_1^3 + Pw + x_3 \\ \dot{x}_3 & = & u - x_1 \\ e & = & x_3 \end{array} \tag{3.1}$$

with $\beta > 0$ and $\epsilon \in \mathbb{R}$ small numbers, $P := (0 \ \ 1)$, and where w is generated by the following linear oscillator:

$$\begin{array}{rcl} \dot{w}_1 & = & w_2 \\ \dot{w}_2 & = & -w_1 . \end{array}$$

[1]In this respect, we observe that this is directly implied by Assumption 1.8, as $u^\star$ is given as a function of only w.

[2]Notice that letting $z = (x_1, x_2)$ and $e = x_3$ yields a system of the form (1.20).

Every regulator that ensures $e(t) = 0$ for all $t \in \mathbb{R}_+$, also must ensure $\dot{e}(t) = 0$ almost everywhere in $\mathbb{R}_+$. This means that $u(t)$ must compensate, at the steady state, the effect of $x_1(t)$ produced by the zero dynamics

$$\begin{aligned} \dot{w}_1 &= w_2 \\ \dot{w}_2 &= -w_1 \\ \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -x_1 - \beta x_2 + \epsilon x_1^3 + w_1, \end{aligned} \tag{3.2}$$

i.e. $u^\star$ must equal the output x_1 of (3.2) a.e. in $\mathbb{R}_+$. Therefore, any regulator solving the problem must embed a suitable replica of (3.2) inside the control loop to be able to generate $u^\star(t)$ at the steady state. Nevertheless, no matter how small β and $|\epsilon|$ are, for sufficiently large initial conditions of (w, x) the system (3.2) admits *chaotic solutions* (see Sprott, 2010, Sec. 2.4). As a consequence, the information given by the exosystem (a simple linear oscillator) is arbitrarily far to be sufficient to individuate a model for the desired $u^\star$, that is potentially non-periodic and has a chaotic attractor, and the role of the exosystem in generating $u^\star(t)$ confuses and melts with the residual dynamics of the plant (in this case the dynamics of x_1 and x_2 restricted to the set in which $x_3 = 0$). This example also shows that Assumption 1.8 is very restrictive, since even a simple example like (3.1) does not satisfy it outside a neighborhood of the origin[3].

When (3.1) is linear (take $\epsilon = 0$) and asymptotically stable, then the plant's residual dynamics, represented by the equations of x_1 and x_2 in (3.2), just act as a linear filter on w_1, with the only effect to change its phase and amplitude. That means that no matter how we chose $\beta > 0$ or even add new linear terms to (3.1), if $\epsilon = 0$ and (x_1, x_2) is asymptotically stable, then $u^\star(t)$ can be always produced by the system

$$\begin{aligned} \dot{\eta}_1 &= \eta_2, \\ \dot{\eta}_2 &= -\eta_1 \\ u^\star &= \eta_1 \end{aligned}$$

which is completely determined by the knowledge of the exosystem.

This example shows how in nonlinear systems the plant itself plays a strong

[3]To see this, note that if Assumption 1.8 holds, then $u^\star$ is a function of w and, as such, must be periodic. This however contradicts the fact that (3.2) has chaotic solutions for large enough initial conditions.

active part in the definition of the law $u^\star(t)$, and hence, if a regulator is constructed to embed a copy of the process that generates $u^\star(t)$ then any arbitrarily slight perturbations in the plant, as well as in the exosystem, can in principle invalidate such an internal model, thus breaking the possibility of obtaining asymptotic regulation. This leads us to conclude the following important, even if straightforward, fact:

> *A regulator can in principle guarantee robust asymptotic regulation only under perturbations of the plant or the exosystem that do not affect the process that generates $u^\star$.*

Therefore, looking for a nonlinear regulator that is robust with respect to arbitrary, even if small, perturbations of the plant is not conceptually different than looking for a linear regulator that is robust also with respect to variations in the exosystem matrix S, which is a property that also the linear regulator does not have. This fact motivated the content of Section 3.4, where we formalize the fact that the very special robustness property of the linear regulator is just a fortunate consequence of linearity, and we conjecture that, in the general nonlinear context, *no regulator is robust*.

3.2 Robust Internal Models by Immersion and the Low-Power Construction

This section contains original results adapted by the author's paper (Bin et al., 2016). We propose a regulator that guarantees some form of robustness to a class of parametric uncertainties. We first provide a procedure to immerse the (uncertain) process that produces the error-zeroing control law $u^\star$ into a known system of higher order. Then we propose a design of the internal model unit based on the low-power high-gain observers of (Astolfi and Marconi, 2015) to make the implementation of internal model units of large dimension more convenient.

We follow here the same "structural robustness" concept of (Byrnes et al., 1997a). The idea is to approach the design of "robust" regulators by assuming that the whole uncertainty is concentrated in a fixed number of parameters of the ideal internal model unit (that is, we assume that (1.7) holds with ϕ that depends on some uncertain parameters); thus we define a simple procedure to immerse that uncertain ideal internal model in a larger system that does not depend on the uncertain parameters. As the dimension of the new model might increase considerably, we substitute the "high-gain" design (1.30) with an equivalent "low-power high-gain" version, so as to avoid the power explosion of terms of the form g^i that, as g has to be chosen large, might result in infeasible practical implementation.

3.2.1 Robustification by Immersion

Even though the considerations of Section 3.1 constitute a negative answer to the robustness quest, for certain kind of problems robustness can be achieved without adaptation by means of a design by *immersion*. Given two systems of the form

$$\begin{aligned} \dot{x} &= f(x) \\ y &= h(x) \end{aligned} \qquad\qquad \begin{aligned} \dot{x}' &= f'(x') \\ y' &= h'(x') \end{aligned}$$

defined on the subsets $X \subset \mathbb{R}^n$ and $X' \subset \mathbb{R}^{n'}$, $n, n' \in \mathbb{N}$, with state $x \in \mathbb{R}^n$ and $x' \in \mathbb{R}^{n'}$ and output $y, y' \in \mathbb{R}^p$ $p \in \mathbb{N}$, we define the concept of *immersion* of systems as follows:

Definition 3.1. *(Byrnes et al., 1997a) The system x is said to be immersed into x' if there exists a smooth mapping $\tau : X \to X'$ satisfying $\tau(0) = 0$ and*

$$h(x_1) \neq h(x_2) \implies h'(\tau(x_1)) \neq h'(\tau(x_2))$$

for all $x_1, x_2 \in X$ and such that

$$\begin{aligned} \frac{\partial \tau}{\partial x} f(x) &= f'(\tau(x)) \\ h(x) &= h'(\tau(x)) \end{aligned}$$

for all $x \in X$.

In other words, saying that x is immersed in x' means to say that the output y corresponding to each solution x with values in X can be obtained as an output y' of the system x' on X'. Immersion assumptions are at the base of many "robust" approaches. For instance, nonlinear systems with parametric uncertainties have been considered in (Byrnes et al., 1997a) (see also (Byrnes et al., 1997b)), where robustness is achieved by assuming that the (uncertain) process generating the ideal error-zeroing control input $u^\star$ is immersed into a linear system whose dynamics does not depend on the uncertain parameters. The same assumption appeared for instance in (Khalil, 1998; Serrani and Isidori, 2000; Serrani et al., 2001; Byrnes and Isidori, 2003; Huang and Chen, 2004), while immersion in a known nonlinear system is exactly Assumption 1.7.

Here we start from Assumption 1.7, where however, ϕ is not known. We assume thought that ϕ admits an affine parametrization in the unknown parameters, i.e. we can write

$$\phi(\cdot) = h(\cdot) + \theta^T \psi(\cdot), \tag{3.3}$$

for some *known* $h : \mathbb{R}^d \to \mathbb{R}$, $\psi : \mathbb{R}^d \to \mathbb{R}^p$ and $p \in \mathbb{N}$ and for some *unknown* $\theta \in \mathbb{R}^p$. The idea is to find a new $d' \in \mathbb{N}$, larger than d, and $\phi' : \mathbb{R}^{d'} \to \mathbb{R}$, *independent on* θ, such that Assumption 1.7 holds with (d, ϕ) substituted by (d', ϕ'). In other words, we want to immerse the unknown nonlinear system[4]

$$u^{\star(d)} = \phi(u^{\star(0,d-1)})$$

into the known system

$$u^{\star(d')} = \phi'(u^{\star(0,d'-1)}).$$

The idea is not new in its essence. In (Isidori et al., 2012) the same idea has been used to cope with an unknown linear ϕ with arbitrary dimension, while in (Forte et al., 2013) the same idea has been extended to some nonlinear oscillators. Here we provide a formal extension to (Forte et al., 2013).

The idea pursued here is illustrated in the following example

Example 3.1. Consider a linear oscillator of the form

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -\gamma^2 x_1 \end{aligned} \tag{3.4}$$

with output

$$y = x_1,$$

[4]Recall from the notation section that $u^{\star(i,j)} := (u^{\star(i)}, u^{\star(i+1)}, \ldots, u^{\star(j)})$.

where the frequency $\gamma \in \mathbb{R}$ is an unknown parameter. Differentiating $\dot{x}_2$ yields

$$\ddot{x}_2 = -\gamma^2 \dot{x}_1.$$

As $\dot{x}_2 = \ddot{x}_1$ and $\ddot{x}_2 = x_1^{(3)}$, then we have

$$\begin{pmatrix} \ddot{x}_1 \\ x_1^{(3)} \end{pmatrix} = -\gamma^2 \begin{pmatrix} x_1 \\ \dot{x}_1 \end{pmatrix}.$$

Hence, along each solution of (3.4) that does not originate in the origin (and thus guarantee that $x_1^2 + \dot{x}_1^2 > 0$) we can write

$$\gamma^2 = -\frac{\ddot{x}_1 x_1 + \dot{x}_1 x_1^{(3)}}{x_1^2 + \dot{x}_1^2}.$$

Differentiating further $x_1^{(3)}$ and substituting such expression of γ yields

$$x_1^{(4)} = \frac{\ddot{x}_1 x_1 + \dot{x}_1 x_1^{(3)}}{x_1^2 + \dot{x}_1^2} \ddot{x}_1,$$

namely, by letting $z := (x_1, \dot{x}_1, \ddot{x}_1, x_1^{(3)})$, we have that the system (3.4) is immersed into the following system:

$$\begin{aligned} \dot{z}_1 &= z_2 \\ \dot{z}_2 &= z_3 \\ \dot{z}_3 &= z_4 \\ \dot{z}_4 &= \frac{z_3 z_1 + z_2 z_4}{z_1^2 + z_2^2} z_3 \end{aligned}$$

that has twice the dimension of the original system (3.4), but that does not depend on the unknown parameter γ. The intuition is that, if we have a model of the z system, we also have a model of any system obtained from (3.4) by letting γ vary in $\mathbb{R}$, without the need of knowing γ explicitly. $\triangle$

We extend now the idea illustrated in the previous example to more general classes of systems. With reference to Section 1.2.3, suppose that Assumption 1.6 holds, and let

$$\mathcal{U}^\star := \left\{ u^\star \text{ given by (1.27)} \;:\; (w, z) \in \mathcal{S}_{(1.25)}(\mathcal{A}) \right\}$$

be the set of all the possible error-zeroing inputs. We start from (3.3), by assuming that all $u^\star \in \mathcal{U}^\star$ fulfill

$$u^{\star(d)} = h\big(u^{\star(0,d-1)}\big) + \psi\big(u^{\star(0,d-1)}\big)^T\theta, \tag{3.5}$$

with d, h, ψ, θ defined as above. Note that from the the smoothness of f, q and b in (1.24), from the sign-definiteness of b and since $\mathcal{A}$ is compact and forward invariant for (1.25), then for each $d \in \mathbb{N}$ there exists a compact set $U^d \subset \mathbb{R}$ such that, for all $u^\star \in \mathcal{U}^\star$,

$$u^{\star(d)}(t) \in U^d, \qquad \forall t \in \operatorname{dom} u^\star. \tag{3.6}$$

For any $d \in \mathbb{N}$, we then define $\mathbf{U}_d := U^0 \times \cdots \times U^d$.

For $i \in \mathbb{N}$ we define the functions $h^j : \mathbb{R}^{d+j} \to \mathbb{R}$ as:

$$\begin{aligned} h^0\big(u^{\star(0,d+j-1)}\big) &:= h\big(u^{\star(0,d+j-1)}\big) \\ h^j\big(u^{\star(0,d+j-1)}\big) &:= \frac{\partial h^{j-1}\big(u^{\star(0,d+j-2)}\big)}{\partial u^{\star(0,d+j-2)}} u^{\star(1,d+j-1)}, \qquad j = 1, \dots, i, \end{aligned}$$

and, for each $k = 1, \dots, p$, we let $\psi_j^k : \mathbb{R}^{d+j} \to \mathbb{R}$ be the functions

$$\begin{aligned} \psi_k^0\big(u^{\star(0,d+j-1)}\big) &:= \psi_k\big(u^{\star(0,d+j-1)}\big) \\ \psi_k^j\big(u^{\star(0,d+j-1)}\big) &:= \frac{\partial \psi_{j-1}^k\big(u^{\star(0,d+j-2)}\big)}{\partial u^{\star(0,d+j-2)}} u^{\star(1,d+j-1)}, \qquad j = 1, \dots, i, \end{aligned}$$

where we let ψ_k denote the k-th component of ψ. We then let $\psi^j := \operatorname{col}(\psi_1^j, \dots, \psi_p^j)$ and

$$\begin{aligned} H_i\big(u^{\star(0,d+i-1)}\big) &:= \operatorname{col}\Big(h^j\big(u^{\star(0,d+j-1)}\big) \,:\, j \le i\Big) \\ \Psi_i\big(u^{\star(0,d+i-1)}\big) &:= \operatorname{col}\Big(\phi^j\big(u^{\star(0,d+j-1)}\big)^T \,:\, j \le i\Big). \end{aligned}$$

Then we have the following result:

Proposition 3.1. *Assume that there exist $d, p \in \mathbb{N}$, smooth $h : \mathbb{R}^d \to \mathbb{R}$ and $\psi : \mathbb{R}^d \to \mathbb{R}^p$, and $\theta \in \mathbb{R}^p$ such that, for all $u^\star \in \mathcal{U}^\star$,* (3.5) *holds, and suppose that there exists $m \ge p-1$ such that*

$$\operatorname{rank} \Psi_m(\mathbf{u}) = p, \qquad \forall \mathbf{u} \in \mathbf{U}_{d+m-1}.$$

Then there exist $d' \in \mathbb{N}$ and a locally Lipschitz $\phi' : \mathbb{R}^{d'} \to \mathbb{R}$, independent on θ, such *that Assumption 1.7 holds.*

Proof. Differentiating (3.5) m times and collecting the obtained equations yields

$$u^{\star(d,d+m)} = H_m(u^{\star(0,d+m-1)}) + \Psi_m(u^{\star(0,d+m-1)})\theta.$$

Solving for θ leads to

$$\theta = \Psi_m\big(u^{\star(0,d+m-1)}\big)^\dagger\Big(u^{\star(d,d+m)} - H_m(u^{\star(0,d+m-1)})\Big). \tag{3.7}$$

As Ψ_m has constant rank on the whole $\mathbf{U}_{d+m-1}$, then the map

$$\mathbf{u} \in \mathbb{R}^{d+m} \mapsto \Psi_m(\mathbf{u}_{[1,d+m]})^\dagger(\mathbf{u}_{[d+1,d+m+1]} - H_m(\mathbf{u}_{[1,d+m]}))$$

is smooth in an open set containing $\mathbf{U}_{d+m}$, and, thus, (3.7) is well defined along each $u^\star \in \mathcal{U}^\star$.

Taking the $(m+1)$-th derivative of (3.5) yields

$$u^{\star(d+m+1)} = h_{m+1}\big(u^{\star(0,d+m)}\big) + \psi_{m+1}\big(u^{\star(0,d+m-1)}\big)^T\theta,$$

and substituting (3.7) yields

$$u^{\star(d')} = \bar{\phi}\big(u^{\star(0,d'-1)}\big),$$

where $d' := d+m+1$ and $\bar{\phi} : \mathbf{U}_{d+m} \to \mathbb{R}$ is the function

$$\begin{aligned}\mathbf{u} \mapsto h_{m+1}(\mathbf{u}_{[1,d+m+1]}) &+ \psi_{m+1}(\mathbf{u}_{[1,d+m+1]})^T\cdot\\ &\cdot \Psi_m(\mathbf{u}_{[1,d+m]})^\dagger(\mathbf{u}_{[d+1,d+m+1]} - H_m(\mathbf{u}_{[1,d+m]})).\end{aligned}$$

The existence of a locally Lipschitz map $\phi' : \mathbb{R}^{d'} \to \mathbb{R}$ that agrees with $\bar{\phi}$ on $\mathbf{U}_{d+m}$ is then provided by the Kirszbraun theorem (see e.g. Federer, 1969, Theorem 2.10.43), and this concludes the proof. ■

As a consequence of Proposition 3.1, we may end up with a pair (d, ϕ) such that Assumption 1.7 holds and no uncertainty is present. This results in a control design that is robust in the canonical sense. Nevertheless, it comes with a regression order d that might be very large. As the Byrnes-Isidori regulator is build to implement, at a given time scale and in given coordinates, a *high-gain* *observer* (Gauthier and Kupka, 2001; Khalil and Praly, 2013) of the quantity $u^\star$, large model dimensions d would lead to an explosion of the power of the high-gain parameter (denoted by g in (1.30)), with consequent problems in terms of peaking, noise amplification and implementation issues (Khalil and Praly, 2013), (Astolfi et al., 2016). In the next section we present a regulator design in which the high-gain internal model unit of the Byrnes-Isidori regulator is substituted by a "low-power" version based on the recent *low-power high-gain observers* introduced in (Astolfi and Marconi, 2015), resulting in the same asymptotic behavior but without having to implement terms of the form g^i with $i > 2$.

3.2.2 Low-Power High-Gain Internal Models

We consider here the same class of SISO normal forms (1.24) under Assumption 1.6 and with the following relaxation of Assumption 1.7.

Assumption 3.1. *There exists $d \in \mathbb{N}$, a locally Lipschitz function $\phi : \mathbb{R}^d \to \mathbb{R}$, a $\bar{\delta} \in \mathbb{R}_+$ and, for each $u^\star \in \mathcal{U}^\star$, a continuous $\delta : \mathbb{R} \to \mathbb{R}$, such that $|\delta|_\infty \le \bar{\delta}$ and*

$$u^{\star(d)}(t) = \phi\big(u^\star(t), \dot{u}^\star(t), \dots, u^{\star(d-1)}(t)\big) + \delta(t).$$

We observe that the same assumption has been considered in (Isidori et al., 2012) for the Byrnes-Isidori regulator, under the additional Assumption 1.8 and with $\mathcal{A} = \operatorname{graph} \pi$ (in this case δ can be taken as $\delta(t) = \nu(w(t))$, for some $\nu : W \to \mathbb{R}$). In (Isidori et al., 2012), the authors showed that if $\nu \neq 0$, then the following asymptotic bound on the regulation error holds:

$$\limsup_{t\to\infty} |e(t)| \le \frac{c}{g^{d+1}} |\nu(W)|, \tag{3.8}$$

with $c > 0$ a constant not depending on k or g. We will provide here an equivalent result adapted to Assumption 3.1.

With (A, B, C) a triplet in prime form of dimension 2 and $g > 0$ a high-gain parameter, let

$$\begin{array}{rcll} N & := & B^T B & \in \mathbb{R}^{2\times 2} \\ \Gamma & := & \begin{pmatrix} C & 0 & \dots & 0 \end{pmatrix} & \in \mathbb{R}^{d\times(2d-2)} \\ T & := & \operatorname{diag}(C, \dots, C, I_2) & \in \mathbb{R}^{d\times(2d-2)} \\ D_2(g) & := & \operatorname{diag}(g, g^2) & \in \mathbb{R}^{2\times 2}. \end{array}$$

Let $\mathbf{U}_{d-1} \subset \mathbb{R}^d$ be a compact set such that $u^{\star(0,d-1)}(t) \in \mathbf{U}_{d-1}$ for all $u^\star \in \mathcal{U}^\star$ and all $t \in \operatorname{dom} u^\star$. With $r > 0$ arbitrary, let $\phi_s : \mathbb{R}^d \to \mathbb{R}$ be any bounded Lipschitz function such that $\phi_s(\mathbf{u}) = \phi(\mathbf{u})$ for all $\mathbf{u} \in \mathbf{U}_{d-1} + r\overline{\mathbb{B}}$ and $|\phi_s(\mathbb{R}^d)| \leq C_\phi$, for some $C_\phi > 0$. For a $\eta \in \mathbb{R}^{2d-2}$ consider the partition $\eta = \operatorname{col}(\eta^1, \dots, \eta^{d-1})$ with $\eta^i \in \mathbb{R}^2$, and let $F : \mathbb{R}^{2d-2} \to \mathbb{R}^{2d-2}$ be the linear map

$$F(\eta) := \operatorname{col}\big(F_1(\eta), \dots, F_{d-1}(\eta)\big)$$

where the elements $F_i(\eta) \in \mathbb{R}^2$ are defined as

$$\begin{array}{rcll} F_1(\eta) & := & A\xi_1 + N\eta_2 & \\ F_i(\eta) & := & A\eta_i + N\eta_{i+1} + D_2(g)L_i(B^T\eta_{i-1} - C\eta_i), & \quad i = 2, \dots, d-2 \\ F_{d-1}(\eta) & := & A\eta_{d-1} + B\phi_s(T\eta) + D_2(g)L_{d-1}(B^T\eta_{d-2} - C\eta_{d-1}), & \end{array}$$

with $L_i := \begin{pmatrix} \ell_{i1} & \ell_{i2} \end{pmatrix} \in \mathbb{R}^{1\times 2}$ coefficients to be designed. Finally let

$$G := \operatorname{col}\big(D_2(g)L_1,\, 0_{2\times 1}, \cdots, 0_{2\times 1}\big) \quad \in \mathbb{R}^{2d-2}.$$

Then we define the low-power high-gain regulator as a system with state $\eta \in \mathbb{R}^{2d-2}$ and input v, satisfying the following equations

$$\begin{array}{rcl} \dot\eta & = & F(\eta) + Gv \\ u & = & \Gamma\eta + v \\ v & = & -ke, \end{array} \tag{3.9}$$

with $k > 0$ a further control parameter. Overall, the coefficients to be fixed are the high-gain parameters g and k and the coefficients $L_i = (\ell_{i1} \quad \ell_{i2})$ for $i = 1, \dots, d-1$. The existence of a choice for them, and the resulting asymptotic properties of (3.9), are expressed by the following proposition:

Proposition 3.2. *With $W \subset \mathbb{R}^{n_w}$ and $Z \subset \mathbb{R}^{n_z}$ arbitrary compact sets, let Assumptions 1.6 and 3.1 bu fulfilled, with $\mathcal{A}$ that is also locally exponentially stable for* (1.25)*. Let $E \subset \mathbb{R}$ and $H \subset \mathbb{R}^{2d-2}$ be arbitrary compact subsets, then there exist $(\ell_{i1}, \ell_{i2}) \in \mathbb{R}^2$, $i = 1, \ldots, d-1$, $g^\star > 0$, $c > 0$ and, for each $g > g^\star$, a $k^\star(g) > 0$, such that for all $g > g^\star$ and $k > k^\star(g)$ the trajectories of the closed loop system* (1.24), (3.9) *originating from $W \times Z \times E \times H$ are bounded and such that*

$$\limsup_{t\to\infty} |e(t)| \le \frac{c}{k\, g^d}\bar{\delta}. \tag{3.10}$$

Proof. Let $\mathcal{O}$ be an open set containing $W \times Z$. Let us define recursively the functions $\tau_i : \mathcal{O} \to \mathbb{R}$ as

$$\begin{aligned}
\tau_1(w,z) &= -\frac{q(w,z,0)}{b(w,z,0)} \\
\tau_i(w,z) &= \frac{\partial \tau_{i-1}(w,z)}{\partial w} s(w) + \frac{\partial \tau_{i-1}(w,z)}{\partial z} f(w,z,0), \quad i = 2, \ldots, d,
\end{aligned}$$

and let $\tau^e : \mathcal{O}_{\mathcal{A}} \to \mathbb{R}^{2d-2}$ be the function

$$\begin{aligned}
\tau^e(w,z) &:= \mathrm{col}\left(\tau_i^e(w,z) \;:\; i = 1, \ldots, d-1\right) \\
\tau_i^e(w,z) &:= \mathrm{col}\left(\tau_i(w,z), \tau_{i+1}(w,z)\right).
\end{aligned}$$

The closed loop system (1.24), (3.9) is a system with unitary relative degree between the input v and the output e and zero dynamics described by

$$\begin{aligned}
\dot{w} &= s(w) \\
\dot{z} &= f(w,z,0) \\
\dot{\eta}_1 &= A\eta_1 + N\eta_2 + D_2(g)L_1(\tau_1(w,z) - C\eta_1) \\
\dot{\eta}_i &= A\eta_i + N\eta_{i+1} + D_2(g)L_i(B^T\eta_{i-1} - C\eta_i) \qquad i = 2, \ldots, d-2 \\
\dot{\eta}_{d-1} &= A\eta_{d-1} + B\phi_s(T\eta) + D_2(g)L_{d-1}(B^T\eta_{d-2} - C\eta_{d-1})
\end{aligned} \tag{3.11}$$

System (3.11) is characterized by the following lemma.

Lemma 3.1. *There exists $g^\star > 0$ and a compact set $\mathcal{B} \subset \mathbb{R}^{n_w+n_z+2d-2}$ such that, for all $g > g^\star$, $\mathcal{B}$ is asymptotically stable for* (3.11) *with a domain of attraction including*

$W \times Z \times H$. Moreover, there exists $c_0 > 0$ such that the following bound holds:

$$|\tau_1(w,z) - C\eta_1| \le \frac{c_0}{g^d}\bar{\delta}, \qquad \forall (w,z,\eta) \in \mathcal{B}. \tag{3.12}$$

The lemma is proved below this proof. The rest of the proof follows from quite standard arguments. Consider the change of variables

$$\eta \mapsto \chi := \eta - G\int_0^e \frac{1}{b(w,z,s)}ds,$$

which is well defined as b is sign definite and which transforms η to the system

$$\dot{\chi} = F(\chi) + G(\tau_1(w,z) - C\chi_1) + \Delta(w,z,e),$$

with

$$\begin{aligned}\Delta(w,z,e) := F\left(G(I-\Gamma)\int_0^e \frac{1}{b(w,z,s)}ds\right) \\ - G\Bigg(\frac{q(w,z,e)}{b(w,z,e)} + \tau_1(w,z) \\ - \int_0^e \frac{1}{b(w,z,s)^2}\left(\frac{\partial b(w,z,s)}{\partial w}s(w) + \frac{\partial b(w,z,s)}{\partial f(w,z,e)}\right)ds\Bigg)\end{aligned}$$

that, in each compact subset of $\mathbb{R}^{n_w+n_z+1}$ is linearly bounded by $|e|$. The equation of e, instead, reads as

$$\dot{e} = q(w,z,e) + b(w,z,e)(\Gamma\chi + v) + \Lambda(w,z,e) \tag{3.13}$$

being

$$\Lambda(w,z,e) := b(w,z,e)\Gamma G\int_0^e \frac{1}{b(w,z,s)}ds$$

that on each compact subset of $\mathbb{R}^{n_w+n_z+1}$ is linearly bounded by $|e|$. Developing further (3.13) yields

$$\begin{aligned}\dot{e} &= b(w,z,e)\left(\frac{q(w,z,e)}{b(w,z,e)} \pm \tau_1(w,z) + \Gamma\chi + v\right) + \Lambda(w,z,e) \\ &= \rho_1(w,z,e) + \rho_2(w,z,e,\chi) + \Lambda(w,z,e) + b(w,z,e)v,\end{aligned}$$

with

$$\rho_1(w,z,e) := b(w,z,e)\left(\frac{q(w,z,e)}{b(w,z,e)} + \tau_1(w,z)\right)$$
$$\rho_1(w,z,e) := b(w,z,e)\Gamma\Big(\Gamma\chi - \tau_1(w,z)\Big)$$

that vanish with $(w,z,\eta) \in \mathcal{B}$, $e=0$ and $\bar{\delta}=0$. In particular, in view of Lemma 3.1, for each compact subset of $\mathbb{R}^{n_w+n_z+1}$ there exists $M>0$ such that

$$|\rho_1(w,z,e)+\rho_2(w,z,e)| \le M\Big(|e| + |(w,z,\eta)|_{\mathcal{B}} + \frac{\alpha_3}{g^d}\bar{\delta}\Big).$$

Thus, by noting that the zero dynamics between the input v and the output e coincide with (3.11) (with $\chi=\eta$), and since $\mathcal{A}$ is locally exponentially stable, standard high-gain arguments (see e.g. Byrnes et al., 2003; Isidori, 1995, 1999) can be used to show that there exists $k^\star(g)>0$ such that the claim holds. ■

Proof of Lemma 3.1. Let

$$\Delta_i(g) := g^{2-i}D_2(g)^{-1},$$

and with

$$\Delta(g) := \mathrm{diag}\left(\Delta_0(g),\dots,\Delta_{d-1}(g)\right)$$

consider the change of variables

$$\eta \mapsto \varepsilon := \Delta(g)(\eta - \tau^e(w)) \tag{3.14}$$

We start analyzing the dynamics of ε component-wise. Consider the partition $\varepsilon = \mathrm{col}(\varepsilon_1,\dots,\varepsilon_{d-1})$, where for each $i=1,\dots,d-1$ we let $\varepsilon_i = \mathrm{col}(\varepsilon_{i1},\varepsilon_{i2}) \in \mathbb{R}^2$. For $i=1$ we have

$$\begin{aligned}\dot{\varepsilon}_1 &= \Delta_0(g)(\dot{\eta}_1 - \dot{\tau}_1^e(w,z))\\ &= \Delta_0(g)\Big(A\eta_1 + N\eta_2 + D_2(g)L_1(\tau_1(w,z) - C\eta_1) - \dot{\tau}_1^e(w,z)\Big)\end{aligned}$$

Noting that:

$$\eta_i = \Delta_{i-1}(g)^{-1}\varepsilon_i + \tau_i^e(w,z)$$

$$\begin{aligned}
&\Delta_i(g) A \Delta_i(g)^{-1} = gA, && \forall i = 0, \dots, d-1 \\
&\Delta_i(g) N \Delta_{i+1}(g)^{-1} = gN, && \forall i = 0, \dots, d-2 \\
&g^{2-i} C \Delta_i(g)^{-1} = gC, && \forall i = 0, \dots, d-1 \\
&\dot{\tau}_i^e(w,z) = \tau_{i+1}^e(w,z) = A\tau_i^e(w,z) + N\tau_{i+1}^e(w,z), && \forall i = 1, \dots, d-2
\end{aligned}$$

then we obtain

$$\begin{aligned}
\dot{\varepsilon}_1 &= gA\varepsilon_1 + \Delta_0(g)\Big(A\tau_1^e(w,z) + N\Delta_1(g)^{-1}\varepsilon_2 + N\tau_2^e(w,z) - \dot{\tau}_1^e(w,z)\Big) \\
&\qquad + g^2 L_1 C \Delta_0(g)^{-1}\varepsilon_1 \\
&= g(A + L_1 C)\varepsilon_1 + gN\varepsilon_2.
\end{aligned}$$

For $i \in \{2, \dots, d-2\}$, instead:

$$\begin{aligned}
\dot{\varepsilon}_i &= \Delta_{i-1}(g)\Big(A\eta_i + N\eta_{i+1} + D_2(g) L_i (B^T \eta_{i-1} - C\eta_i) - \dot{\tau}_i^e(w,z)\Big) \\
&= gA\varepsilon_i + gN\varepsilon_{i+1} + g^{2-(i-1)} L_i (B^T\eta_{i-1} - C\eta_i)
\end{aligned}$$

Since

$$B^T\eta_{i-1} - C\eta_i = B^T\Delta_{i-2}(g)^{-1}\varepsilon_{i-1} - C\Delta_{i-1}(g)^{-1}\varepsilon_i + B^T\tau_{i-1}^e(w,z) - C\tau_i^e(w,z)$$

and, by construction,

$$\begin{aligned}
&B^T\tau_{i-1}^e(w,z) - C\tau_i^e(w,z) = 0 \\
&g^{2-(i-1)} B^T \Delta_{i-2}(g)^{-1} = gB^T
\end{aligned}$$

then we obtain

$$\dot{\varepsilon}_i = g(A - L_i C)\varepsilon_i + gN\varepsilon_{i+1} + L_i B^T \varepsilon_{i-1}.$$

Finally, for $i = d-1$, we have

$$\begin{aligned}
\dot{\varepsilon}_{d-1} &= \Delta_{d-2}(g)\Big(A\eta_{d-i} + B\phi_s(T\eta) + D_2(g) L_{d-1}(B^T\eta_{d-2} - C\eta_{d-1}) - \dot{\tau}_{d-1}^e(w,z)\Big) \\
&= g(A - L_{d-1}C)\varepsilon_{d-1} + gL_{d-1}B^T\varepsilon_{d-2} \\
&\qquad + \Delta_{d-2}(g)\Big(A\tau_{d-1}^e(w,z) + B\phi_s\big(T\Delta(g)^{-1}\varepsilon + T\tau^e(w,z)\big) - \dot{\tau}_{d-1}^e(w,z)\Big).
\end{aligned}$$

Noting that

$$\dot{\tau}_e^{d-1}(w,z) = A\tau_{d-1}^e(w,z) + B\dot{\tau}_d(w,z), \tag{3.15}$$

and since $T\tau^e(w,z) = \tau(w,z)$, then, by letting

$$\tau_{d+1}(w,z) := \frac{\partial \tau_d(w,z)}{\partial w} s(w) + \frac{\partial \tau(w,z)}{\partial z} f(w,z,0),$$

we obtain

$$\begin{aligned}\dot{\varepsilon}_{d-1} &= g(A - L_{d-1}C)\varepsilon_{d-1} + gL_{d-1}B^T\varepsilon_{d-2} \\ &\quad + \frac{1}{g^{d-1}} B\Big(\phi_s\big(T\Delta(g)^{-1}\varepsilon + \tau(w,z)\big) - \tau_{d+1}(w,z)\Big).\end{aligned}$$

Hence, by letting for $i = 1, \dots, d-1$, $E_i \in \mathbb{R}^{2\times 2}$, $Q_i \in \mathbb{R}^{2\times 2}$

$$\Sigma := \begin{pmatrix} 0_{(2d-1)\times 1} \\ 1 \end{pmatrix} \qquad E_i := \begin{pmatrix} -\ell_{i1} & 1 \\ -\ell_{i2} & 0 \end{pmatrix}, \qquad Q_i := \begin{pmatrix} 0 & \ell_{i1} \\ 0 & \ell_{i2} \end{pmatrix}$$

and

$$M := \begin{pmatrix} E_1 & N & 0 & & \dots & \dots & 0 \\ Q_2 & E_2 & N & 0 & & & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & \ddots & Q_i & E_i & N & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & Q_{d-2} & E_{d-2} & N \\ 0 & \dots & \dots & \dots & 0 & Q_{d-1} & E_{d-1} \end{pmatrix},$$

the system ε can be compactly rewritten as

$$\dot{\varepsilon} = gM\varepsilon + g^{1-d}\Sigma\Big(\phi_s\big(T\Delta(g)^{-1}\varepsilon + \tau(w,z)\big) - \tau_{d+1}(w,z)\Big). \tag{3.16}$$

By using Lemma 1 in (Astolfi and Marconi, 2015), it is possible to show that we can always choose the matrices L_i such that M is Hurwitz. Let Ξ be a compact set such that $\eta \in H$ and $(w,z) \in W \times Z$ imply $\varepsilon \in \Xi$. As M is Hurwitz and, by construction, ϕ_s is bounded by C_ϕ, then the reachable sets[5] $\mathcal{R}^{\tau}_{(3.16)}(\Xi)$ are uniformly bounded for each $\tau > 0$, and we can assume without loss of generality

[5]See the notation section.

that $\mathcal{R}^{\tau}_{(3.16)}(\Xi) \subset \operatorname{int}\Xi$ for sufficiently large $\tau > 0$, as C_ϕ is independent on Ξ. This in turn implies that the Ω-limit set $\mathcal{B}' := \Omega_{(3.11)}(W \times Z \times \Xi)$ of the zero dynamics (3.11) is compact invariant and (Assumption 1.6) included in $\operatorname{int}(W \times Z \times \Xi)$. Point 1 of Assumption 1.6 also implies that for each $(w, z, \varepsilon) \in \mathcal{B}'$, $(w, z) \in \mathcal{A}$.

Add and subtract to (3.16) the term $g^{1-d}\Sigma\phi(\tau(w, z))$ to obtain

$$\dot{\varepsilon} = gM\varepsilon + g^{1-d}\Sigma\Big(\phi_s\big(T\Delta(g)^{-1}\varepsilon + \tau(w, z)\big) - \phi\big(\tau(w, z)\big) + \nu\Big).$$

with

$$\nu := \phi\big(\tau(w, z)\big) - \tau_{d+1}(w, z).$$

As $\mathcal{A}$ is asymptotically stable for the (w, z) subsystem of (3.11), then there exists $\bar{t} > 0$ such that, for all $t \geq \bar{t}$, $\tau(w, z) \in \mathbf{U}_{d-1} + r\overline{\mathbb{B}}$ and from the properties of ϕ_s we claim the existence of a $L_\phi > 0$ such that, for all $t > \bar{t}$,

$$|\phi_s\big(T\Delta(g)^{-1}\varepsilon + \tau(w, z)\big) - \phi\big(\tau(w, z)\big)| \leq L_\phi|T\Delta(g)\varepsilon| \leq g^{d-1}L_\phi|\varepsilon|.$$

As M is Hurwitz, standard high-gain arguments (Gauthier and Kupka, 2001; Khalil and Praly, 2013; Isidori, 2017) show the existence of a $g^\star > 0$ such that, for all $g > g^\star$, the following estimate holds:

$$\limsup_{t\to\infty} |\varepsilon(t)| \leq \frac{\alpha_3}{g^d} \limsup_{t\to\infty} |\nu(t)|, \tag{3.17}$$

uniformly in the initial conditions and for some $\alpha_1, \alpha_2, \alpha_3 > 0$ independent on g.

Pick now $(w, z, \varepsilon) \in \mathcal{B}'$. By definition of $\mathcal{B}'$, there exists a sequence $((w^n, z^n, \varepsilon^n))_n$ in $\mathcal{S}_{(1.25)}(W \times Z \times \Xi)$ and a strictly increasing sequence $(t_n)_n$ in $\mathbb{R}_+$ with $t_n \to \infty$, such that

$$(w^n(t_n), z^n(t_n), \varepsilon^n(t_n)) \to (w, z, \varepsilon). \tag{3.18}$$

We thus have:

$$\begin{aligned} |\varepsilon| &\leq |\varepsilon - \varepsilon^n(t_n)| + |\varepsilon^n(t_n)| \\ &\leq |\varepsilon - \varepsilon^n(t_n)| + \frac{\alpha_3}{g^d} \limsup_{t\to\infty} |\phi(\tau(w^n(t), z^n(t))) - \tau_{d+1}(w^n(t), z^n(t))| \\ &\leq |\varepsilon - \varepsilon^n(t_n)| \\ &\qquad + \frac{\alpha_3}{g^d} \limsup_{t\to\infty} \Big(|\phi(w, z) - \tau_{d+1}(w, z)| + |\phi(\tau(w^n(t), z^n(t))) - \phi(w, z)| + \end{aligned}$$

$$|\tau_{d+1}(w,z) - \tau_{d+1}(w^n(t), z^n(t))|\Big)$$

Due to (3.18), and since $(w,z) \in \mathcal{A}$ implies $\phi(w,z) - \tau_{d+1}(w,z) = \delta$, then, for each $\mu > 0$ and each large enough n the previous inequality yields

$$|\varepsilon| \leq \mu + \frac{\alpha_3}{g^d}\bar{\delta},$$

and from the arbitrariness of μ we claim that

$$|\varepsilon| \leq \frac{\alpha_3}{g^d}\bar{\delta}.$$

Let $\mathcal{B}$ be the compact set such that $(w,z,\eta) \in \mathcal{B}$ if and only if $(w,z,\varepsilon) \in \mathcal{B}'$. Since $|\tau_1(w,z) - C\eta_1| \leq |\varepsilon|$ a then the claim follows from the arbitrariness of (w,z,ε). ■

3.2.3 An Example

In this example the low-power high-gain observer regulator (3.9) and the design approach by nonlinear regression presented above are applied together to address a robust output regulation problem. The control goal is to asymptotically reject, by means of the same regulator, a disturbance which can be indistinguishably generated by *uncertain* linear, Duffing or Van der Pool oscillators. To this end, the immersion argument introduced above is used to find a system (with an overall order of $d = 7$) in which all the three uncertain oscillators can be immerse, and the low-power high-gain regulator (3.9) is used to implement an internal model unit of such system.

In this example we consider the following controlled plant:

$$\begin{aligned}\dot{x}_1 &= -2x_1 + x_2^3 \\ \dot{x}_2 &= 2x_2 - 2x_1 + u - w_1\end{aligned}$$

where u is the control input,

$$e = x_2$$

is the (measured) regulation error, and w_1 is the exogenous disturbance, which is assumed to be generated from a linear, a Duffing or a Van der Pool oscillator with unknown parameters. In particular, we can model w_1 as the first component of a system of the form

$$\ddot{w} = \alpha w + \beta \dot{w} + \upsilon w^3 + \gamma w^2 \dot{w} \tag{3.19}$$

which for different configurations of the parameters includes, among all the others, also the dynamics of interest. To put the plant in the normal form (1.24) we simply let $z = x_1$ and we rewrite the plant as

$$\begin{aligned} \dot{z} &= -2z + e^3, \\ \dot{e} &= 2e - 2z + u - w_1. \end{aligned}$$

Note that the steady state control law able to maintain the error to zero is exactly $u^\star = w_1$, which satisfies Assumption 1.7, with $d = 2$ and ϕ given by (3.19). By following the same notation as in the first part of the section, we can express ϕ as (3.3), with

$$\psi(u^\star, \dot{u}^\star) := \begin{pmatrix} u^\star \\ \dot{u}^\star \\ u^{\star 3} \\ u^{\star 2}\dot{u}^\star \end{pmatrix} \quad \theta := \begin{pmatrix} \alpha \\ \beta \\ \upsilon \\ \gamma \end{pmatrix}.$$

The procedure detailed in Proposition 3.1 can be applied to construct a system of dimension $d = 7$ into which (3.19) can be immersed. Hence, a regulator of the form (3.9) is used to control the system, with an overall dimension of $2(d-1) = 12$. The controller design is completed by the choice $v = -ke$, where $k > 0$ is chosen large enough. Figure 3.1 shows the simulation results of the overall closed-loop systems subject to a disturbance w_1 which in the first 10 seconds is produced by a sinusoid at frequency 3 rad/s (obtained from (3.19) with $\alpha = -9$ and $\beta = \upsilon = \gamma = 0$). At time $t = 10s$ it switches to the output of a Duffing oscillator obtained by letting $\alpha = 2$, $\upsilon = -1$ and $\beta = \gamma = 0$. Finally at time $t = 20s$ it switches to the output of a Van der Pool oscillator obtained with $\alpha = -4$, $\beta = 1$, $\upsilon = 0$ and $\gamma = -1$. In order to dominate the dynamics of the 7^{th} derivative of the considered exosystem, we used a gain $g = 200$.A standard high gain design we would have had a term of $g^d = 200^7$, which is a 17-digit number.

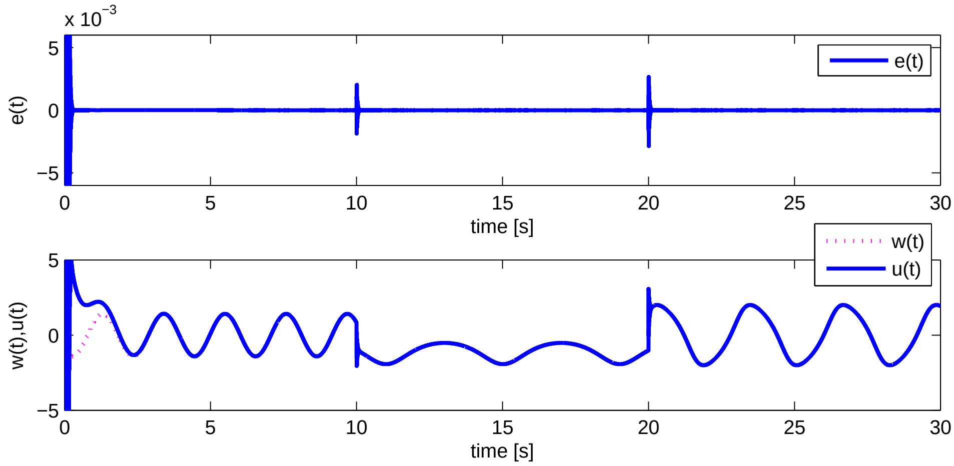

Figure 3.1: Simulation results.

3.3 Output Regulation with Set-Valued Exosystems

This section contains original results published in (Petit et al., 2018). Here we approach the problem of robustness from a different perspective: instead of modeling the exosystem as an ordinary differential equation (ODE), we model it as a *differential inclusion* (Aubin and Cellina, 1984; Aubin, 1991), i.e. in place of $\dot{w} = s(w)$, we write

$$\dot{w} \in S(w), \tag{3.20}$$

with $w \in \mathbb{R}^{n_w}$ and being $S : \mathbb{R}^{n_w} \rightrightarrows \mathbb{R}^{n_w}$ a set-valued map. The solutions to (3.20) are *absolute continuous* functions (Aubin and Cellina, 1984) that, thus, need not to be differentiable but, rather, admit a *distributional derivative* that in general may differ among any two solutions. In particular $w : [0, t] \to \mathbb{R}^{n_w}$ is absolute continuous if there exists a Lebesgue integrable function $w' : [0, t] \to \mathbb{R}^{n_w}$ such that we can write

$$w(t) = w(0) + \int_0^t w'(s)ds.$$

Then we say that $w(t)$ solves (3.20) if $w'(s) \in S(w(s))$ a.e. in $[0, t]$. Clearly, system (3.20) can generate a consistently larger multitude of signals than an ordinary differential equation (which is obtained whenever $S(w) = \{s(w)\}$ and when we restrict to C^1 functions). A relevant case for regulation, for instance, is the ability to model differential equations subject to *uncertain* time-varying parameters (in this case S is a *parametrized map*), or exosystems with variable structure. For instance, let $\mu \in \mathbb{R}^{n_\mu}$ be an unmodeled time-varying vector of unknown parameters and suppose we have an uncertain exosystem of the form

$$\dot{w} = s(w, \mu). \tag{3.21}$$

Suppose moreover that we know that μ ranges in a set $M \subset \mathbb{R}^{n_\mu}$. Then, in place of (3.21), we can consider (3.20) where

$$S(w) := \Big\{ w' \in \mathbb{R}^{n_w} \,:\, w' = s(w, \mu),\, \mu \in M \Big\}.$$

Modeling the exosystem with a differential inclusion means translating the uncertainties in the right internal model to use into uncertainties in the model of the exogenous signals. At the *analysis level*, the exosystem is usually needed to define fundamental notions such as the *steady state* and the *zero dynamics* of a nonlinear system, moving the attention from "signals" to "systems". Under a *synthesis* point of view, the exosystem is generally exploited to identify an ideal steady state in which the regulated variables vanish, and thus to chose the degrees of freedom of the regulator. In most of the designs (see Section 1.2.3), the structure of the exosystem enters explicitly in the definition of the regulator, and generally only a perfect knowledge of the exosystem dynamics can guarantee asymptotic tracking. Nevertheless, as in this thesis we eventually look towards regulator designs that can adapt at run time, it is worth wondering if adaptive regulators will still be as tied to the model of the exogenous signals as non adaptive ones or if the hypothesis of the exosystem being an ODE could be weakened in future.

As a first preliminary work, in this section we follow the line of (Byrnes and Isidori, 2003), by extending the concepts of steady state and zero dynamics to the case in which the exosystem has the form (3.20). We also give necessary conditions for the solvability of the output regulation problem and we extend the characterization in terms of zero dynamics and the notion of efficient controllers as given in (Byrnes and Isidori, 2003).

In this section we will adopt the following additional notations: L_1^{loc} denotes the space of functions that are locally in L_1. With $\dot{x} = f(x, u)$ a differential equation with input, where $x \in \mathcal{X}$ and $u \in \mathcal{U}$, being $\mathcal{X}$ and $\mathcal{U}$ vector spaces, we denote by $(t, x, u) \mapsto \phi_x(t, x, u)$ the value of the solution originating in $x \in \mathcal{X}$

at time $t = 0$ with input u. Moreover, for all fixed u and all $X \subset \mathcal{X}$, we let $\mathcal{S}_x(X, u) := \{\phi_x(\cdot, x, u) \, : \, x \in X\}$ be the set of all the solutions starting in X driven by the input u. With $\mathcal{F}$ a set of functions from $\mathbb{R}_+$ to $\mathbb{R}^d$, $d \in \mathbb{N}^*$, and with $t \geq 0$, we denote by $\mathcal{F}|_t$ the set of all functions obtained by restricting an element of $\mathcal{F}$ to the interval $[0, t]$. We denote by $AC(\mathcal{X}, \mathcal{Y})$ the set of all the absolute continuous functions from $\mathcal{X}$ to $\mathcal{Y}$. We denote by $(t_n)_n \nearrow$ a sequence of $t_n \in \mathbb{R}_+$ that are strictly increasing and $\lim_{n\to\infty} t_n = \infty$. Let $x = (x_1, x_2, \ldots, x_m) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \ldots \times \mathbb{R}^{n_m}$, for $i = 1, \ldots, m$ we denote by $\mathrm{Pr}_{x_i}(x) := x_i$ the projection of x on $\mathbb{R}^{n_i}$. For a set $X \subset \mathbb{R}^d$ and $x \in \mathbb{R}^d$, we denote by $T_X(x)$ the tangent space to X at point x. In the following we will also often call φ_x a solution to a system with state x to avoid confusion between solutions and point while keeping the notation simple.

3.3.1 Preliminaries

In this section we consider the following interconnection

$$\dot{w} \in S(w) \tag{3.22}$$

$$\dot{\xi} = \psi(w, \xi) \tag{3.23}$$

in which an autonomous differential inclusion with state $w \in \mathbb{R}^{n_w}$, $n_w \in \mathbb{N}^*$, drives a nonlinear system with state $\xi \in \mathbb{R}^{n_\xi}$, $n_\xi \in \mathbb{N}^*$. We suppose that the initial conditions of (3.22), (3.23) range in a compact subset $W \times \Xi \subset \mathbb{R}^{n_w} \times \mathbb{R}^{n_\xi}$. We assume that $\psi : W \times \mathbb{R}^{n_\xi} \to \mathbb{R}^{n_\xi}$ is locally Lipschitz and $S : W \rightrightarrows \mathbb{R}^{n_w}$ is Lipschitz on W and has non-empty and compact values at each $w \in W$.

We shall introduce now the essential preliminary concepts instrumental for the forthcoming analysis.

Preliminary definitions:

With $N, M > 0$, we define the set of *admissible solutions* of (3.22) as

$$\mathcal{L}_w(w_0) = \left\{ \varphi_w \in \mathcal{S}_w(w_0) \, : \, |\varphi_w|_\infty \leq M \text{ and } \forall \varphi_\xi \in \mathcal{S}_\xi(\Xi, \varphi_w), |\varphi_\xi|_\infty \leq N \right\}$$

which is the set of all the bounded solutions to (3.22) which produce bounded solutions to (3.23). For ease of notation, with $Z \subset \mathbb{R}^{n_w} \times \mathbb{R}^{n_\xi}$, we define the set

$$\mathcal{A}(Z) = \Big\{(\varphi_\xi, \varphi_w) \in AC(\mathbb{R}_+, \mathbb{R}^{n_w} \times \mathbb{R}^{n_\xi}) \; : \\ \varphi_w \in \mathcal{L}_w(w_0), \; \varphi_\xi \in \mathcal{S}_\xi(\xi_0, \varphi_w), \; (\xi_0, w_0) \in Z\Big\}.$$

which is the set of the *admissible solutions* to (3.22)-(3.23) from Z. With $\mathcal{B}$ a set of functions from $\mathbb{R}_+$ into $\mathbb{R}^d$, $d \in \mathbb{N}^*$, we define the *flow* of $\mathcal{B}$ as the set-valued map $\Phi_\mathcal{B} : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$

$$\Phi_\mathcal{B}(t, b_0) = \Big\{b \in \mathbb{R}^d \; : \; \exists \varphi \in \mathcal{B}, \varphi(t) = b, \; \varphi(0) = b_0\Big\}.$$

With $\varphi \in \mathcal{B}$ we define the *ω-limit set* of φ as the set

$$\omega(\varphi) = \Big\{b \in \mathbb{R}^d \; : \; \exists (t_n)_n \nearrow, \varphi(t_n) \to_{n \to +\infty} b\Big\}$$

or equivalently $\omega(\varphi) = \cap_{t \geq 0} \Phi_\varphi(t, \varphi(0))$. Furthermore, we define the *Ω-limit set* of $\mathcal{B}$ as

$$\Omega(\mathcal{B}) = \Big\{b \in \mathbb{R}^d \; : \; \exists (t_n)_n \nearrow, \; \exists (\varphi^n)_n, \; \varphi^n \in \mathcal{B}, \varphi^n(t_n) \to b\Big\}.$$

Let A be a set such that $A = \cup_{\varphi \in \mathcal{B}} \{\varphi(0)\}$. We shall say that the set A is *Poisson Stable* if $A = \omega(\mathcal{B}) = \cup_{\varphi \in \mathcal{B}} \omega(\varphi)$. With $A \subset \mathbb{R}^d$, we say A *uniformly attracts* $\mathcal{B}$ if

$$\forall \varepsilon > 0, \; \exists T > 0, \; \forall \varphi \in \mathcal{B}, \; \forall t > T, \; |\varphi(t)|_A \leq \varepsilon .$$

We say A is *invariant for* $\mathcal{B}$ if

$$\forall t \in \mathbb{R}_+, \; \forall \varphi \in \mathcal{B}, \; \varphi(t) \in A.$$

If the set $\mathcal{B}$ is clear from the context, we omit to mention it. When invariance or attractiveness refer to the solutions to a differential equation (or inclusion), we always refer to the set of complete solutions if not other set is mentioned.

Properties of Limit Sets:

We study now the asymptotic behavior of the admissible solutions to the interconnection (3.22), (3.23). Under mild existence and regularity assumptions, we show that the Ω-limit set of $\mathcal{A}(W \times \Xi)$ is a well-defined compact set that uniformly attracts $\mathcal{A}(W \times \Xi)$. This results are instrumental for the forthcoming analysis in the context of output regulation. From now on, we assume W to be *invariant* for (3.22) and we fix $N > 0$ in the definition of $\mathcal{L}_w(W)$. We denote with Φ the flow of the set $\mathcal{A}(W \times \Xi)$. Finally, with slight abuse of notation, we let $\Omega := \Omega(\mathcal{A}(W \times \Xi))$ and we make the following existence and admissibility assumption:

Assumption 3.2. *The following hold:*

A1) *For all $w_0 \in W$, $\mathcal{L}_w(w_0)$ is non empty.*

A2) $\mathcal{S}_w(W) = \mathcal{L}_w(W)$

A relevant case in which Assumptions A1-A2 hold is when the system (3.23) is input-to-state-stable relatively to the origin and with respect to the input w. In this case, indeed, there exist $\gamma, \rho \in \mathcal{K}_\infty$ such that (Sontag, 1995)

$$\forall \varphi_w \in \mathcal{S}_w(W), \quad |\varphi_\xi(t)| \leq \gamma(|\varphi_\xi(0)|) + \rho(|\varphi_w|_\infty)$$

for all $\varphi_\xi \in \mathcal{S}_\xi(\Xi)$ and for all $t \in \mathbb{R}_+$. Therefore, every pair (φ_w, φ_ξ) is in $\mathcal{A}(W \times \Xi)$ with $M := \max_{w \in W} |w|$ and $N := \rho(M) + \max_{\xi \in \Xi} \gamma(|\xi|)$.

With the next theorem we show that under Assumption 3.2 the Ω-limit set of $\mathcal{A}(W \times \Xi)$ is a well-defined compact attractor for the admissible solutions to (3.22), (3.23).

Theorem 3.1. *Assume A1. Then Ω is non-empty, compact, and uniformly attracts $\mathcal{A}(W \times \Xi)$. Moreover, Ω is the graph of an upper semicontinuous set-valued map and, if in addition A2 holds, then Ω is invariant for $\mathcal{A}(\Omega)$.*

Before proving Theorem 3.1, we prove the following technical lemma.

Lemma 3.2. *For all $t \in \mathbb{R}_+$ and any two solutions $\varphi_w^0 \in \mathcal{S}_w(W)$ and $\varphi_w^1 \in \mathcal{S}_w(\varphi_w^0(t))$, let $\oplus_t$ be the* concatenation operator

$$\varphi_w^0 \oplus_t \varphi_w^1(s) := \mathbf{1}_{[0,t]}(s)\varphi_w^0(s) + \mathbf{1}_{]t,+\infty[}(s)\varphi_w^1(s-t)$$

Then $\varphi_w^0 \oplus_t \varphi_w^1 \in \mathcal{S}_w(W)$.

Proof. Since, by definition, φ_w^i, $i = 0, 1$, is absolutely continuous then it admits derivative in L_1^{loc} such that

$$\varphi_w^i(s) = \varphi_w^i(0) + \int_0^s (\varphi_w^i)'(u)du$$

and $(\varphi_w^i)'(s) \in S(\varphi_w^i(s))$ a.e., $i = 0, 1$. Then

$$\varphi_w^0 \oplus_t \varphi_w^1(s) = \varphi_w^0(0) + \int_0^s \big(\mathbf{1}_{[0,t]}(u)(\varphi_w^0)'(u) + \mathbf{1}_{]t,+\infty[}(u)(\varphi_w^1)'(u-t)\big)\, du$$

and then

$$(\varphi_w^0 \oplus_t \varphi_w^1)'(s) = \mathbf{1}_{[0,t]}(s)(\varphi_w^0)'(s) + \mathbf{1}_{]t,+\infty[}(s)(\varphi_w^1)'(s-t)$$

which proves that $\varphi_w^0 \oplus_t \varphi_w^1$ is absolutely continuous and satisfies $(\varphi_w^0 \oplus_t \varphi_w^1)'(s) \in S(\varphi_w^0 \oplus_t \varphi_w^1)$ a.e. on $\mathbb{R}_+$. ∎

Proof of Theorem 3.1.

First we prove that Ω is compact. Boundedness follows from the definition of $\mathcal{L}_w(W)$, hence it suffices to prove it is closed. Let $(w_n, \xi_n)_{n\in\mathbb{N}}$ be a sequence in Ω converging to (w, ξ). By definition of Ω, for all $n \in \mathbb{N}$,

$$\exists (t_k^n)_k \nearrow, \exists (\varphi_\xi^{n,k}, \varphi_w^{n,k})_{k\in\mathbb{N}} \in \mathcal{A}(W \times \Xi)^{\mathbb{N}}, (\varphi_\xi^{n,k}(t_k^n), \varphi_w^{n,k}(t_k^n)) \to_{k\to+\infty} (\xi_n, w_n).$$

We can index k on n to obtain for all $n \in \mathbb{N}$

$$|(\varphi_\xi^{n,k}(t_k^n), \varphi_w^{n,k}(t_k^n)) - (\xi_n, w_n)| \le 2^{-n}$$

that in turn implies

$$|(\varphi_\xi^{n,k}(t_k^n), \varphi_w^{n,k}(t_k^n)) - (\xi, w)| \le 2^{-n} + |(\xi_n, w_n) - (\xi, w)|.$$

This shows that $(\xi, w) \in \Omega$ and thus Ω is closed, hence compact.

We now show uniform attractiveness of Ω for $\mathcal{A}(W \times \Xi)$. By contradiction, assume

$$\exists \varepsilon > 0, \forall T > 0, \exists t > T, \exists (\varphi_\xi, \varphi_w) \in \mathcal{A}(W \times \Xi), |(\varphi_\xi(t), \varphi_x(t))|_\Omega > \varepsilon.$$

Then there exist a sequence $(t_n)_n \nearrow$ and a sequence $(\varphi_\xi^n, \varphi_w^n)_{n\in N}$ in $\mathcal{A}(W \times \Xi)$ such that

$$|(\varphi_\xi^n(t_n), \varphi_w^n(t_n))|_\Omega > \varepsilon.$$

By definition of $\mathcal{A}(W \times \Xi)$, $(\varphi_\xi^n(t_n), \varphi_w^n(t_n))_n$ lives in a compact set and thus there exists a subsequence which converges to (ξ, w), which is in Ω by definition. As this is a contradiction and we claim the uniformly attractiveness of Ω.

We now show invariance of Ω for $\mathcal{A}(\Omega)$. Pick arbitrarily $(w_0, \xi_0) \in \Omega$ and let $\varphi_w \in \mathcal{S}_w(w_0)$. We now show a element $\varphi_\xi \in \mathcal{S}_\xi(\xi_0, \varphi_w)$ is defined at least on $[0, T]$, where $T > 0$ does not depend on (w_0, ξ_0) picked in Ω. Consider $\varphi_\xi^{ref} \in \mathcal{S}_\xi(\Xi, \varphi_w)$, so as, by definition, $|\varphi_\xi^{ref}|_\infty \leq N$. Let $\eta > 0$ be such that $\Omega \subset \eta\mathbb{B}$ (with $\mathbb{B}$ the unit open ball in $\mathbb{R}^{n_w+n_\xi}$). By the fact ψ is locally Lipschitz there exists $T > 0$ such that φ_ξ is defined on $[0, T]$ and $\varphi_\xi(t) \leq \eta$ for all $t \in [0, T]$. In fact Gronwall lemma gives us :

$$\forall t \in [0, T], |\varphi_\xi^{ref}(t) - \varphi_\xi(t)| \leq |\varphi_\xi^{ref}(0) - \varphi_\xi(0)|e^{Lt} \tag{3.24}$$

where L is the Lipschitz constant of ψ on $\eta\overline{\mathbb{B}}$. With

$$\mu > \max\left\{|\varphi_\xi^{ref}(0) - \xi| \;:\; \xi \in \mathrm{Pr}_\xi(\Omega)\right\},$$

let

$$T^* = \min_{\xi\in\mathrm{Pr}_w(\Omega)}\left\{\frac{1}{L}\ln\left(\frac{\mu}{|\varphi_\xi^{ref}(0) - \xi|}\right)\right\},$$

which is non-negative by the choice of μ. Let us take η big enough to have

$$\forall \xi \in \mathbb{R}^{n_\xi}, \forall t \in \mathbb{R}_+, |\varphi_\xi^{ref}(t) - \xi| < \mu \Rightarrow |\xi| < \eta,$$

which is possible as φ_ξ^{ref} has a compact positive orbit. In view of (3.24), if $T < T^*$, then $|\varphi_\xi^{ref}(T) - \varphi_\xi(T)| < \mu$ and $|\varphi_\xi(T)| < \eta$. T^* is independent of (w_0, ξ_0) picked in Ω. From now on we considered only maximal solutions that, in view of the previous analysis are defined for $T > T^*$. By definition of Ω, there exist $(t_n)_n \nearrow$ and $(\varphi_w^n, \varphi_\xi^n)$ such that

$$(\varphi_w^n(t_n), \varphi_\xi^n(t_n)) \to_{n\to+\infty} (w_0, \xi_0).$$

By the hypotheses on S (see Aubin and Cellina, 1984, Thm. 1, ch. 2.4), for all

$n \in \mathbb{N}$, there exists $\bar{\varphi_w}^n \in \mathcal{S}_w(\varphi_w^n(t_n))$ such that

$$\sup_{t\in[0,T^*]} |\bar{\varphi_w}^n(t) - \varphi_w(t)| \le |\bar{\varphi_w^n}(t_n) - \varphi_w(0)| e^{\alpha T^*} .$$

By lemma 3.2, $\varphi_w^n \oplus_{t_n} \bar{\varphi_w}^n \in \mathcal{S}_w(W)$ and, from the last estimate, we obtain

$$\forall t \in [0, T^*], \ \varphi_w^n \oplus_{t_n} \bar{\varphi_w}^n(t_n + t) \to_{n\to+\infty} \varphi_w(t) .$$

We now have to prove the same kind of result for the variable ξ. For all $n \in \mathbb{N}$, consider the solution $\bar{\varphi_\xi^n}$ of (3.23) with initial condition $\varphi_\xi^n(0)$ and input $\bar{\varphi_w^n}$. Since $\varphi_w^n \oplus_{t_n} \bar{\varphi_w^n} \in \mathcal{S}_w(W)$ then, by using the fact that $\mathcal{S}_w(W) = \mathcal{L}_w(W)$ and from the definition of $\bar{\varphi_\xi^n}$, we deduce that $\bar{\varphi_\xi^n}$ is defined on $\mathbb{R}_+$ and bounded by N. From the Gronwall lemma we get

$$|\phi_\xi(t, \varphi_\xi^n(0), \bar{\varphi_w^n}) - \phi_\xi(t, \xi_0, \varphi_w)| \le \left(|\varphi_\xi^n(0) - \xi_0| + LT^* \sup_{s\in[0,T^*]} |\bar{\varphi_w^n}(s) - \varphi_w(s)| \right) e^{Lt}$$

for all $t \in [0, T^*]$. Hence, for all $t \in [0, T^*]$,

$$(\varphi_w^n \oplus_{t_n} \bar{\varphi_w^n}(t_n + t), \varphi_\xi^n \oplus_{t_n} \bar{\varphi_\xi^n}(t_n + t)) \to_{n\to+\infty} (\varphi_w(t), \varphi_\xi(t))$$

and then $(\varphi_w(t), \varphi_\xi(t)) \in \Omega$. Since $(\varphi_w(t), \varphi_\xi(t)) \in \Omega$ and T^* does not depends of the element chosen in it, invariance is obtained by induction.

Finally, to prove Ω is the graph of an upper semi continuous map. Define

$$\pi(w) = \left\{ \xi \in \mathbb{R}^{n_\xi} \ : \ (w, \xi) \in \Omega \right\}$$

then π is well-defined and upper-semicontinuity follows from (see Aubin and Cellina, 1984, Thm. 1, ch. 1). ■

The following proposition also shows that Ω is the smallest set that has the properties of Theorem 3.1.

Proposition 3.3. *Assume Ω is not empty. Then Ω is the smallest closed set (in the sense of inclusion) which uniformly attracts $\mathcal{A}(W \times \Xi)$.*

Proof. Assume a closed set K uniformly attracts $\mathcal{A}(W \times \Xi)$. Pick a point $\omega \in \Omega$, it suffices to prove that $\omega \in K$. By definition there exist sequences $(t_n)_n \nearrow$ and

$(\varphi_\xi^n, \varphi_w^n)_{n\in\mathbb{N}}$ in $\mathcal{A}(W \times \Xi)$ such that

$$(\varphi_w^n(t_n), \varphi_\xi^n(t_n)) \to \omega$$

and, from uniform attractiveness of K,

$$\forall \varepsilon > 0, \exists T > 0, \forall t \geq T, \forall (\varphi_\xi, \varphi_w) \in \mathcal{A}(W \times \Xi), \quad |(\varphi_w(t), \varphi_\xi(t)|_K \leq \frac{\varepsilon}{2}.$$

Fix $\varepsilon > 0$. Then, for n big enough,

$$d(\omega, K) \leq |(\varphi_w^n(t_n), \varphi_\xi^n(t_n)) - \omega| + |(\varphi_w^n(t_n), \varphi_\xi^n(t_n)|_K \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

For arbitrariness of $\varepsilon > 0$, we thus claim that $|\omega|_K = 0$, which proves $\omega \in K$, as K is closed. ■

3.3.2 Necessary Conditions for Output Regulation

In this section, we show how the asymptotic characterization of the interconnections of the kind (3.22), (3.23) presented so far can be used to deduce necessary conditions for the output regulation problem. In doing this we follows the line of development of (Byrnes and Isidori, 2003).

We consider here systems of the kind

$$\begin{aligned} \dot{x} &= f(w, x, u) \\ y &= \begin{pmatrix} e \\ y_a \end{pmatrix} = \begin{pmatrix} h_e(w, x) \\ h_a(w, x) \end{pmatrix} =: h(w, x) \end{aligned} \tag{3.25}$$

with state $x \in \mathbb{R}^n$, control input $u \in \mathbb{R}^m$, output $y = (y_a, e) \in \mathbb{R}^{n_a} \times \mathbb{R}^{n_e}$ and with $w \in \mathbb{R}^{n_s}$ that is generated by the exosystem

$$\dot{w} \in S(w) \tag{3.26}$$

with initial conditions that range in a compact invariant set $W \subset \mathbb{R}^{n_w}$. As before, we assume $S : \mathbb{R}^{n_w} \rightrightarrows \mathbb{R}^{n_w}$ to be Lipschitz on W with non-empty compact values at each $w \in W$, and we assume that f and h are locally Lipschitz. As in the rest of the text, the output e represents the system outputs that need to asymptotically vanish, while y_a is the set of measured outputs that might be needed for stabilization purposes but that are not required to vanish at the steady state. In this framework the problem of semiglobal *output regulation* reads as follows: given a compact set $X \subset \mathbb{R}^n$ of initial conditions for (3.25), find a controller of the form

$$\begin{array}{rcl} \dot{\eta} &=& g(\eta, y) \\ u &=& \gamma(\eta, y) \end{array} \tag{3.27}$$

with state $\eta \in \mathbb{R}^{n_\eta}$ and with $g : \mathbb{R}^{n_\eta} \times \mathbb{R}^{n_y} \to \mathbb{R}^{n_\eta}$ ($n_y := n_e + n_a$), $\gamma : \mathbb{R}^{n_\eta} \times \mathbb{R}^{n_y} \to \mathbb{R}^m$ locally Lipschitz, and a compact set $H \subset \mathbb{R}^{n_\eta}$, such that, with $K := h_e^{-1}(0)$, the closed-loop system

$$\begin{aligned} &\dot{w} \in S(w) \\ &\dot{x} = f(w, x, \gamma(\eta, h_y(w, x))) \\ &\dot{\eta} = \varphi(\eta, h_y(w, x)) \end{aligned} \tag{3.28}$$

satisfies the following:

1. All the solutions to (3.28) originating in $W \times X \times H$ are admissible.

2. K uniformly attracts $\mathcal{A}(W \times X \times H)$.

With $\xi := \mathrm{col}(x, \eta)$ and $\Xi := X \times H$, the first requirement can be equivalently expressed as $\mathcal{S}_w(W) = \mathcal{L}_w(W)$. The second one, instead, requires the regulation error e to vanish at the steady state.

As a consequence of Theorem 3.1 and Proposition 3.3 we obtain the following necessary conditions for the solvability of the output regulation problem.

Proposition 3.4. *Suppose the problem of output regulation is solvable on $W \times X$. Then there exists an upper semicontinuous set valued map*

$$\pi : \operatorname{dom} \pi \subset W \rightrightarrows \mathbb{R}^n$$

with compact graph, such that

a) $\operatorname{graph} \pi \subset K$.

b) For each $(w, x) \in \operatorname{graph} \pi$, the set of all input functions $u \in \mathbb{R}^m$ such that

$$S(w) \times \{f(x, w, u)\} \subset T_{\operatorname{graph} \pi}(w, x),$$

is non empty.

Proof. Assume that the problem of output regulation is solved, i.e. points 1 and 2 of the definition hold. Point 1 implies Assumption 3.2 and, hence, Ω is well defined and Theorem 3.1 holds. Let $\operatorname{dom}\pi := \{w \in \mathbb{R}^{n_w} : (w, x, \eta) \in \Omega\}$. From the invariance of W it follows that $\operatorname{dom}\pi \subset W$. For each $w \in \operatorname{dom}\pi$, let

$$\pi(w) = \Big\{x \in \mathbb{R}^n : (w, x, \eta) \in \Omega\Big\}.$$

Then, $\pi : \operatorname{dom}\pi \to \mathbb{R}^n$ is well-defined and has compact graph. Upper semi-continuity follows from Theorem 3.1. Thus, point a) of the claim follows from Proposition 3.3.

By considering the restrictions of S, f and g to any open neighborhood of Ω and applying (Aubin, 1991, Thm. 5.3.4) we obtain that

$$S(w) \times \{f(w, x, \gamma(\eta, h_y(w, x)))\} \times \{g(\eta, h_y(w, x))\} \subset T_\Omega(w, x, \eta).$$

Hence, with $E(w, x) := \{\eta \in \mathbb{R}^\ell : (w, x, \eta) \in \Omega\}$, the set $\mathcal{U}(w, x) := \{u \in \mathbb{R}^m : u = \gamma(\eta, h_y(w, x)),\ \eta \in E(w, x),\ (w, x) \in \operatorname{graph}\pi\}$ is precisely the (non-empty) set of inputs for which point b) holds. ∎

Remark 3.1. If in addition we assume W to be *Poisson Stable* for $\mathcal{A}(W \times X \times H)$ then $\operatorname{dom}\pi = W$, as in the case in (Byrnes and Isidori, 2003). △

3.3.3 Output Regulation and Zero Dynamics

In this section we extend the concept of zero dynamics as given in (Byrnes and Isidori, 2003) to the case of exosystems given by a differential inclusion. Consider the system

$$\begin{aligned} \dot{z} &\in F(z, u) \\ y &= H(z) \end{aligned} \tag{3.29}$$

with $z \in \mathbb{R}^{n_z}$ and $y \in \mathbb{R}^{n_y}$. We say that $Z \subset \mathbb{R}^{n_z}$ is a *viability domain* for (3.29) if, for all $z \in Z$, there exists u such that $F(z, u) \subset T_Z(z)$. We define the *regulation map* $r : Z \rightrightarrows \mathbb{R}^{n_z}$ as

$$r(z) = \{u \in \mathbb{R}^m : F(z, u) \subset T_Z(z)\}.$$

If the regulation map $r(\cdot)$ admits a selection $\alpha : \mathbb{R}^{n_z} \to \mathbb{R}^m$ such that the map $z \mapsto F(z, \alpha(z))$ is sufficiently regular, then from from (Aubin, 1991, Thm, 5.3.4) we obtain that Z is invariant for (3.29).

We say (3.29) possesses a well-defined *zero dynamics* if there exists a non-empty closed subset Z in $\mathbb{R}^n$ such that

1. $Z \subset H^{-1}(0)$
2. Z is a viability domain and the regulation map possesses a continuous selection α such that $z \mapsto F(z, \alpha(z))$ is Lipschitz and has compact values.
3. If $z_0 \in \mathbb{R}^{n_z}$, $u_0 : \mathbb{R}_+ \to \mathbb{R}^m$ and $z \in \mathcal{S}_z(z_0, u_0)$ are defined on an interval I and such that u_0 is C^0 and

$$\forall t \in I, H(z(t)) = 0$$

then

$$z_0 \in Z \text{ and } u_0(t) = \alpha(z(t)) \; a.e. \text{ on } I$$

The above definition generalizes zero dynamics notion appearing in (Byrnes and Isidori, 2003) and, in particular, the third condition ensures the uniqueness of the selection α.

Consider now the system

$$\begin{array}{rcl} \dot{w} & \in & S(w) \\ \dot{x} & = & f(w, x, u) \\ e & = & h_e(w, x) \end{array} \tag{3.30}$$

and assume that (3.30) possesses well-defined zero dynamics. Let Z_e denote the zero dynamics kernel of (3.30) and let $\alpha : \mathbb{R}^{n_w} \times \mathbb{R}^n \to \mathbb{R}^m$ be the unique continuous selection of the regulation map associated to Z_e. Let $A \subset W \times \mathbb{R}^n$ be the (possibly empty) set for which all the solutions to the system

$$\begin{gathered} \dot{w} \in S(w) \\ \dot{x} = f(w, x, \alpha(w, x)) \end{gathered} \tag{3.31}$$

originating in A are admissible. Then the following result holds.

Proposition 3.5. *Assume* (3.27) *solves the problem of output regulation for* (3.25), (3.26) *on* $W \times X$ *and suppose that* (3.30) *possesses a well defined zero dynamics. Then,*

1. $\mathrm{Pr}_{(w,x)}(\Omega) \subset A$
2. *Any trajectory in* $\mathcal{A}(\Omega)$ *is* Pr_ξ*-related to a trajectory of the zero dynamics system.*
3. *For any* $(x_0, w_0) \in \mathrm{Pr}_\xi(\Omega)$ *there exists* ξ_0 *such that the response* u_δ *of the system*

$$\begin{aligned} \dot{w} &\in S(w) \\ \dot{x} &= f(x, w, \alpha(x, w)) \\ \dot{\xi} &= \varphi(\xi, k(x, w)) \\ u_\delta &= \alpha(x, w) - \gamma(\xi, k(x, w)) \end{aligned} \tag{3.32}$$

is define for all $t \geq 0$ *and identically zero.*

Proof. The result directly follows from the properties of Ω with the same arguments used in (Byrnes and Isidori, 2003, Prop. 6.1). ■

3.3.4 Efficient controllers

Finally, in this section we restrict the focus on the case in which the map π introduced in Proposition 3.4 is single valued. By borrowing the terminology of (Byrnes and Isidori, 2003), we say that (3.27) is an *efficient controller* if it solves the problem of output regulation and there exist two single-valued maps $\varpi : \mathrm{Pr}_w(\Omega(W \times X \times H)) \to \mathbb{R}^n$ and $\rho : \mathrm{Pr}_w(\Omega(W \times X \times H)) \to \mathbb{R}^{n_\eta}$ such that

$$\begin{aligned} \Omega(W \times X \times H) = \{(w, x, \eta) \in W \times \mathbb{R}^n \times \mathbb{R}^{n_\eta} \ : \\ x = \varpi(w) \text{ and } \eta = \rho(w)\}. \end{aligned}$$

The following results hold :

Proposition 3.6. *If* (3.27) *is an efficient controller, then* ϖ *and* ρ *are continuous.*

Proposition 3.7. *Suppose a controller of the form* (3.27) *is efficient and* (3.25) *possesses a well-defined zero dynamics. Then there exist* $\varpi : \mathrm{Pr}_w(\Omega(W \times X \times H)) \to \mathbb{R}^n$ *and* $\rho : \mathrm{Pr}_w(\Omega(W \times X \times H)) \to \mathbb{R}^{n_\eta}$ *such that*

1. $\operatorname{graph} \varpi$ *is invariant for the set of solutions of* (3.31) *starting in* $W \times X$*. Moreover* $h_e(\varpi(w), w) = 0$*.*

2. *For any* $(w_0, \eta_0) \in \{(w, \eta) \in \mathbb{R}^{n_w} \times \mathbb{R}^{n_\eta} \,:\, \eta = \rho(w)\}$ *the response of the system*

$$\begin{aligned} \dot{w} &\in S(w) \\ \dot{x} &= \varphi(\xi, k(\nu(w), w)) \\ u_\delta &= \alpha(\rho(w), w) - \gamma(\rho(w), k(\nu(w), w)) \end{aligned} \tag{3.33}$$

is defined for all $t \geq 0$ *and is identically equal to zero.*

3.4 A Framework for Robustness in Output Regulation

This section contains original contributions published by the author in (Bin et al., 2018b). We do a step back to Section 3.1, and in particular to the question "*what would be the right way to extend the notion of "structural robustness" to nonlinear systems?*" that remained open. In this section we give an answer to that question, by seeking a more abstract definition of robustness able to deal with more general concepts of "perturbation" and with properties milder than asymptotic regulation. The motivation behind the work presented in this section lies in the fact that, while the concept of robustness defined by Francis and Wonham for linear systems has a clear and well-defined meaning, the notion of robustness in a nonlinear setting is still vague, and quite often "robustness results" are claimed in ad-hoc contexts using custom definitions. As mentioned in Section 3.1, the "structural robustness" of (Byrnes et al., 1997a) extends the "parametric" interpretation of the notion of robustness given by (Francis and Wonham, 1976), while the perhaps more interesting notion of perturbation considered in (Astolfi and Praly, 2017) refers to the "functional" nature of the same definition. Although these two notions are equivalent for linear functions between finite-dimensional vector spaces, they do not lead to equivalent notions if nonlinear functions are considered. As a matter of fact, if $f : \mathbb{R}^n \to \mathbb{R}^n$ is linear, then fixing a basis of $\mathbb{R}^n$ defines a function m that sends f to its matrix representation $F = m(f) \in \mathbb{R}^{n\times n}$. We may consider parametric variations of f by "moving" $F = m(f)$ in $\mathbb{R}^{n\times n}$ (thus obtaining the parametric interpretation of the perturbations of the linear plants, with a total of n^2 parameters). If $|\cdot|$ is a norm on $\mathbb{R}^{n\times n}$, these variations are quantified by looking at the norm $|F - F^\circ|$ of the deviation of F relatively to its nominal value F°, so as variations in the same neighborhood of F° are quantified equally. If we induce a topology on the space $\mathcal{F}$ of linear maps $\mathbb{R}^n \to \mathbb{R}^n$ by completing the collection of subsets of the form $m^{-1}(V)$, where V is a neighborhood in the topology induced by $|\cdot|$ on $\mathbb{R}^{n\times n}$, then we obtain a topology on $\mathcal{F}$ which is equivalent to the one considered in (Astolfi and Praly, 2017) (i.e. the C^1 topology (Hirsch, 1994)), and that is in direct relation with the parametric interpretation of the matrix variations. If f is not linear, however, this correspondence is not true anymore. For instance, consider the family $\mathcal{F}$ of functions of the kind $f_\mu(x) = \mu x^2$, with μ a parameter ranging in an interval $I \subset \mathbb{R}$. While for any $\mu_1, \mu_2 \in I$, f_{μ_1} and f_{μ_2} belong to the same class $\mathcal{F}$, the function $g(x) = f_\mu(x) + \epsilon x^3$, $\epsilon > 0$, does not belong to $\mathcal{F}$ for any choice of $\mu \in I$ and $\epsilon > 0$. While f_{μ_1} and f_{μ_2} are obtained in the spirit of the "structural" notion of (Byrnes et al., 1997a), g is obtained in the "differential topology" spirit of (Astolfi and Praly, 2017), and they lead to totally different concept of variation.

In this section we present a unifying concept of robustness relying of a general topological notion of "variation" that includes all the previous cases as particular examples. We also extend the notion of robustness to properties more general than "asymptotic regulation", by capturing in this way a wider variety of "robust behaviors" exhibited by practical and approximate regulation designs. We then review some of the main control approaches for linear and nonlinear systems, by characterizing their robustness properties in terms of the proposed framework. We show that the robustness property of the linear regulator can be framed in this language and it naturally extends to a milder robustness condition relative to the Fourier expansion of the regulation error when applied to nonlinear systems (in this respect, we thus re-frame the result of (Astolfi et al., 2015) in more abstract terms). We also review the general design of (Marconi et al., 2007) for nonlinear systems and we show that robustness of asymptotic regulation does not hold for smooth plant's variations, while a *practical regulation* property does. The section then concludes with a conjecture stating that, in a general nonlinear context, asymptotic regulation is on its nature a fragile property that cannot be achieved in a robust way with a finite dimensional regulator.

3.4.1 The Framework

In this work we deal with a controlled system which is described by a *nominal model* of the form

$$\begin{aligned} \dot{x} &= f^o(w,x,u) \\ y &= h^o(w,x) \end{aligned} \tag{3.34}$$

where $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^m$ is the control input,$y \in \mathbb{R}^{n_y}$, is the measured output, and the functions (f^o, h^o) are supposed to be smooth enough. The measured output y is partitioned as $y := \mathrm{col}(e, y_m)$ in which $e \in \mathbb{R}^{n_e}$, with $p_e \leq n_y$, denotes a regulation error on which there are "asymptotic performance expectations", as detailed later, and $y_m \in \mathbb{R}^{n_y - n_e}$ are possible extra measurements.

The input $w \in \mathbb{R}^{n_w}$ is supposed to be a bounded external signal that may represent a reference to be tracked or a disturbance to be rejected. As in the rest of the text, we suppose that w is generated by a nominal *exosystem* of the form

$$\dot{w} = s^o(w), \tag{3.35}$$

although, most of the considerations reported below do not necessarily rely upon this description.

System (3.34) represents a nominal model on which we argue that an output feedback controller of the form

$$\begin{aligned} \dot{\eta} &= \phi(\eta, y) \\ u &= \theta(\eta, y) \end{aligned} \tag{3.36}$$

with state $\eta \in \mathbb{R}^{n_\eta}$ has been designed so that the resulting closed-loop system satisfies certain properties. For instance, in the problem of *asymptotic output regulation*, it is required that the trajectories of the resulting closed-loop system (3.34)-(3.36) are bounded and the associated regulation error $e(t)$ is asymptotically vanishing, while *approximate regulation* asks for a possibly non-zero asymptotic bound on the regulation errors that, though, enjoys a given meaning or satisfies some optimality conditions (see e.g. Astolfi et al., 2015; Forte et al., 2017). Here, we generically denote by P a *property* expected on the *asymptotic* regulation error and we say that the "regulation objective P" is achieved by a regulator (3.36) if the error trajectories associated to the closed-loop system *asymptotically* satisfy the property P. Examples of properties are clearly "error identically zero" as in the asymptotic output regulation problem or "error with bounded amplitude" in the approximate version.

The problem of designing a regulator so that a regulation objective P is fulfilled assumes a conceptual and practical relevance as soon as robustness aspects are taken into account, namely as soon as the regulation objective must be guaranteed not only in nominal conditions, namely when the regulated system behaves as *nominal* dynamics (3.34), but also when parametric or structural uncertainties are present. In this respect, in the following, we consider the case in which the *real* (unknown) model of the controlled system has the form

$$\begin{array}{rcl} \dot{w} & = & s(w) \\ \dot{x} & = & f(w,x,u) \\ y & = & h(w,x) \end{array} \tag{3.37}$$

in which s, f and h are obtained by "perturbing" the nominal models (3.34), (3.35). In the following, we denote $F := (s, f, h)$ and we assume that F belongs to a given function space $\mathcal{F}$ that includes the nominal value $F^\circ := (s^\circ, f^\circ, h^\circ)$. We thus consider the problem in which the regulation objective P is required to hold, not only for the *nominal* system (3.34), (3.35), but also for all possible actual dynamics (3.37) that are "close enough" to (3.34), (3.35), in some sense that will be made precise below. Namely, the property P is preserved under plant perturbations.

We end the section by stating a fundamental (yet well known) fact. Consider a system of the form

$$\mathcal{H}: \ \dot{z} = g(z), \tag{3.38}$$

defined over a normed vector space $\mathbb{Z}$, with g a continuously differentiable function. Given any subset $Z \subset \mathbb{Z}$, We say that $\mathcal{H}$ is *uniformly eventually bounded* from Z if exists $\tau \geq 0$ such that $\mathcal{R}^{\tau}_{\mathcal{H}}(Z)$ is bounded.

Proposition 3.8. *$\Omega_{\mathcal{H}}(Z)$ exists and is closed. If $\mathcal{H}$ is uniformly eventually bounded from Z, $\Omega_{\mathcal{H}}(Z)$ is compact, non empty, invariant, uniformly attractive from Z and is the smallest (in the sense of inclusion) closed set with this latter property.*

Proposition 3.8 is a direct consequence of the definition (see the Notation section) and of the group property the flow of (3.38). Proposition 3.8, and in particular the fact that $\Omega_{\mathcal{H}}(Z)$ is the smallest set with such properties, motivates referring to $\Omega_{\mathcal{H}}(Z)$ as the *steady state* locus of the trajectories of (3.8) originating in Z.

3.4.2 A Definition of Robustness

The extended controlled plant, given by (3.37) is thus defined by the functions $F := (s, f, h)$ and the regulator (3.36) is designed to enforce a given property P on the steady-state trajectories under the assumption that F equals a nominal value $F^\circ := (s^\circ, f^\circ, h^\circ)$ and the initial conditions range in a nominal set $W_0^\circ \times X_0^\circ \subset \mathbb{R}^{n_w} \times \mathbb{R}^n$. This yields a definition of the maps ϕ and θ in (3.36) and of a set $H_0 \subset \mathbb{R}^{n_\eta}$ of initial conditions for η. In order to study the "robustness" features of the regulator relatively to the property P and under variation of the plant's function F and of the initial set $W_0 \times X_0$, in the following we consider a closed-loop system of the form

$$\mathcal{H}_F: \quad \begin{array}{rcl} \dot{w} &=& s(w) \\ \dot{x} &=& f(w, x, \theta(\eta, h(w,x))) \\ \dot{\eta} &=& \phi(\eta, h(w,x)). \end{array} \tag{3.39}$$

in the case in which the actual plant's function F and the actual initial set $W_0 \times X_0$ possibly differ from the nominal values.

For simplicity, in the following we let $\mathbf{x} := (w, x)$ and $\mathbf{n} = \dim(\mathbf{x})$ and we denote by $\mathcal{S}_F(\mathbf{X}_0 \times H_0)$ and $\Omega_F(\mathbf{X}_0 \times H_0)$ the quantities $\mathcal{S}_{\mathcal{H}_F}(\mathbf{X}_0 \times H_0)$ and $\Omega_{\mathcal{H}_F}(\mathbf{X}_0 \times H_0)$ obtained with a given $F \in \mathcal{F}$ and a given $\mathbf{X}_0 \subset \mathbb{R}^{\mathbf{n}}$. We let $\mathcal{X} := \mathcal{K}(\mathbb{R}^{\mathbf{n}})$, with $\mathcal{K}(\bigstar)$ denoting the set of all the compact subsets of $\bigstar$, and we endow $\mathcal{X}$ with the Hausdorff topology[6] $\tau_{\mathcal{X}}$. We consider functions F that belong to the generic functional space $\mathcal{F}$ introduced above, that we equip with a topology $\tau_{\mathcal{F}}$, and to compact sets of initial condition in $\mathcal{X}$. The particular value of $\mathcal{F}$ and $\tau_{\mathcal{F}}$ will be specified later depending on the context. We equip the product space $\mathcal{F} \times \mathcal{X}$ with the product topology $\tau_{\mathcal{F}\times\mathcal{X}}$ and we make the notion of "perturbation" $(F, \mathbf{X}_0)$ of $(F^\circ, \mathbf{X}_0^\circ)$ introduced in the previous section precise by the

[6] Namely, the topology induced by the distance function:

$$d(\mathbf{X}_1, \mathbf{X}_2) = \max\left\{ \sup_{\mathbf{x}_1 \in \mathbf{X}_1} \inf_{\mathbf{x}_2 \in \mathbf{X}_2} |\mathbf{x}_1 - \mathbf{x}_2|, \ \sup_{\mathbf{x}_2 \in \mathbf{X}_2} \inf_{\mathbf{x}_1 \in \mathbf{X}_1} |\mathbf{x}_2 - \mathbf{x}_1| \right\}$$

for $\mathbf{X}_1, \mathbf{X}_2 \in \mathcal{K}(\mathbb{R}^{\mathbf{n}})$.

following definition.

Definition 3.2. $(F, \mathbf{X}_0) \in \mathcal{F} \times \mathcal{X}$ *is called a perturbation of* $(F^\circ, \mathbf{X}_0^\circ)$ *if it belongs to a* $\tau_{\mathcal{F}\times\mathcal{X}}$*-neighborhood of* $(F^\circ, \mathbf{X}_0^\circ)$.

The regulation objective has been previously introduced throughout the informal definition of a property P that we wish the trajectories of the closed-loop system to asymptotically have. With $\mathbb{S}$ the set of all the functions $\mathbb{R}_+ \to \mathbb{R}^{\mathbf{n}} \times \mathbb{R}^{n_\eta}$, we can formally define the property P by associating to its informal statement the subset $\mathcal{P} \subset \mathbb{S}$ given by $\mathcal{P} := \{(\mathbf{x}, \eta) \in \mathbb{S} \,:\, \text{P holds}\}$. We then say that $(\mathbf{x}, \eta) \in \mathbb{S}$ has the property P if $(\mathbf{x}, \eta) \in \mathcal{P}$. To formally express what we mean by saying that the property P should hold "asymptotically", we restrict our attention to the trajectories of (3.39) that originates in the attractor $\Omega_F(\mathbf{X}_0 \times H_0)$. In order to work with a well-defined (in the sense of Proposition 3.8) set $\Omega_F(\mathbf{X}_0 \times H_0)$, we also need as a basic robustness ingredient that the regulator (3.36) guarantees that $\mathcal{H}_F$ has the desired boundedness properties for the considered perturbations. We put all together within the following definition:

Definition 3.3. *With* $(F, \mathbf{X}_0) \in \mathcal{F} \times \mathcal{X}$*, we say that the regulator* (3.36) *achieves the regulation objective* P *asymptotically at* $(F, \mathbf{X}_0)$ *if* $\mathcal{H}_F$ *is uniformly eventually bounded from* $\mathbf{X}_0 \times H_0$ *and* $(\mathbf{x}, \eta) \in \mathcal{S}_F(\Omega_F(\mathbf{X}_0 \times H_0))$ *implies* $(\mathbf{x}, \eta) \in \mathcal{P}$.

Let $V \subset \mathcal{F} \times \mathcal{X}$, then we say that V *generates an equibounded family of systems*[7] if for every $(F, \mathbf{X}_0) \in V$ the system $\mathcal{H}_F$ defined as in (3.39) is uniformly eventually bounded from $\mathbf{X}_0 \times H_0$ and there exists a compact set $\mathcal{O} \subset \mathbb{R}^{\mathbf{n}}$ such that $\Omega_F(\mathbf{X}_0 \times H_0) \subset \mathcal{O}$ for all $(F, \mathbf{X}_0) \in V$. We define now a formal notion of *robustness* for the regulator (3.36) associated to the property P.

Definition 3.4. *We say that the regulator* (3.36) *is* P*-robust at* $(F^\circ, \mathbf{X}_0^\circ)$ *with respect to* $\tau_{\mathcal{F}}$ *if there exists a* $\tau_{\mathcal{F}\times\mathcal{X}}$*-neighborhood* V *of* $(F^\circ, \mathbf{X}_0^\circ)$ *that generates an equibounded family of systems such that, for all* $(F, \mathbf{X}_0) \in V$*, the regulator achieves the regulation objective* P *asymptotically at* $(F, \mathbf{X}_0)$.

3.4.3 Robustness in Regulators with Linear Internal Model

In this section we consider the class of regulators (3.36) obtained by partitioning the state η as $\eta = (\eta_{\text{im}}, \eta_{\text{s}})$, with $\eta_{\text{im}} \in \mathbb{R}^{n_{\text{im}}}$ and $\eta_{\text{s}} \in \mathbb{R}^{n_{\text{s}}}$ that satisfy the dynamic

[7] Equiboundedness is needed to avoid unfortunate limit cases in which there exists a sequence $((F^n, \mathbf{X}_0^n))_n$ in V such that the corresponding sequence $(\Omega_{F^n}(\mathbf{X}_0^n \times H_0))_n$ escapes to the horizon.

equations

$$\begin{aligned} \dot{\eta}_{\rm im} &= \Phi\eta_{\rm im} + Ge \\ \dot{\eta}_{\rm s} &= \phi_{\rm s}(\eta_{\rm s}, \eta_{\rm im}, y) \qquad \eta(0) \in H_0 \\ u &= \theta(\eta, y). \end{aligned} \tag{3.40}$$

We will focus here on "smooth variations" of F. More precisely, with $\mathbf{X} \subset \mathbb{R}^{\mathbf{n}}$ and $\mathbf{U} \subset \mathbb{R}^m$ arbitrarily large compact sets, throughout this section we suppose that F ranges in the set $\mathcal{C}^1(\mathbf{X}\times\mathbf{U})$ of all the continuously differentiable functions defined on $\mathbf{X} \times \mathbf{U}$ and with values in $\mathbb{R}^{\mathbf{n}} \times \mathbb{R}^p$. We endow $\mathcal{C}^1(\mathbf{X} \times \mathbf{U})$ with the *weak topology* (Hirsch, 1994) $\tau_{\mathcal{C}^1}$ defined as follows: with $\ell := \max\{\max_{\mathbf{x}\in\mathbf{X}} |\mathbf{x}|, \max_{\mathbf{u}\in\mathbf{U}} |\mathbf{u}|\}$ and for any $F \in \mathcal{C}^1(\mathbf{X} \times \mathbf{U})$ and $\epsilon > 0$, an ϵ-neighborhood of F is given as

$$\mathcal{N}_\epsilon(F) = \Big\{G \in \mathcal{C}^1(\mathbf{X} \times \mathbf{U}) \,:\, \max_{\mathbf{p}\in\mathbf{X}\times\mathbf{U}} |F(\mathbf{p}) - G(\mathbf{p})| < \ell\epsilon, \\ \max_{\mathbf{p}\in\mathbf{X}\times\mathbf{U}} |F'(\mathbf{p}) - G'(\mathbf{p})| < \epsilon\Big\},$$

where F' and G' denote the derivatives of F and G. We stress that $\mathbf{X}$ and $\mathbf{U}$ are arbitrary and can be chosen large enough to encompass all the solutions of interest. Restricting the functions F to $\mathbf{X}\times\mathbf{U}$ though allows us to consider a nicer topology $\tau_{\mathcal{C}^1}$ which is first-countable and metrizable rather than the alternative *strong topology* (Hirsch, 1994).

Robustness of the Linear Regulator

We start considering the case in which $F^\circ \in \mathcal{C}^1(\mathbf{X} \times \mathbf{U})$ is linear and the linear regulator (see Section 1.1.3) is used. The regulator is obtained from (3.40) by **a)** letting $H_0 \in \mathcal{K}(\mathbb{R}^{n_\eta})$ be arbitrary, **b)** letting $(\phi_{\rm s}(\cdot), \theta_{\rm s}(\cdot))$ be linear functions whose matrix representation with respect to a fixed basis of $\mathbb{R}^{\mathbf{n}} \times \mathbb{R}^{n_\eta}$ is of the kind

$$\begin{aligned} \dot{\eta}_{\rm s} &= A_s\eta_{\rm s} + B_{s1}\eta_{\rm im} + B_{s2}y \\ u &= K_1\eta + K_2 y, \end{aligned} \tag{3.41}$$

c) choosing (Φ, G) as any controllable pair with Φ that has a characteristic polynomial which coincides with the minimal polynomial of the corresponding matrix representation of $s^\circ(\cdot)$ and, finally, **d)** by fixing $A_s, B_{s1}, B_{s2}, K_1, K_2$ so as to stabilize the nominal system $\mathcal{H}_{F^\circ}$ originating from $\mathbf{X}_0^\circ \times H_0$.

We let $\mathcal{F}_L \subset \mathcal{C}^1(\mathbf{X}\times\mathbf{U})$ be the set of all the linear functions in $\mathcal{C}^1(\mathbf{X}\times\mathbf{U})$ and we let

$$\mathcal{F} := \left\{F \in \mathcal{F}_L \,:\, s = s^\circ\right\}. \tag{3.42}$$

We endow $\mathcal{F}$ with the subset topology $\tau_{\mathcal{F}}$ derived by $(\mathcal{C}^1(\mathbf{X}\times\mathbf{U}), \tau_{\mathcal{C}^1})$ and we let P_0 the property

$$\mathrm{P}_0 \quad = \quad \text{“ } e = h_e(w,x) = 0 \text{ ”}. \tag{3.43}$$

The next result, that follows from (Davison, 1976), captures the main robustness property of the linear regulator.

Proposition 3.9. *The linear regulator* (3.41) *is* P_0*-robust at* $(F^\circ, \mathbf{X}_0^\circ)$ *with respect to* $\tau_{\mathcal{F}}$.

Proof. The existence of a neighborhood $\mathcal{N}^\circ$ of F° for which $\mathcal{H}_F$ is uniformly eventually bounded for each $F\in\mathcal{N}^\circ$ is a direct consequence of the definition of $\tau_{\mathcal{F}}$ and the continuity of the eigenvalues of the function F seen as maps $\mathcal{F}\to\mathbb{R}$. The fact that P_0 holds as long as the trajectories are bounded follows from the more general result of Proposition 3.10. ■

Actually, a stronger result can be given: the P_0-robustness property of Proposition 3.9 is *universal* in $\mathbf{X}_0$, i.e. the same regulator achieves P_0 for any initial condition of the plant. Nevertheless, although the linear regulator is P_0-robust for “$\mathcal{C}^1$ variations” of linear functions that let s° unchanged, this property is a mere consequence of linearity and it is broken by any slight nonlinear perturbation of F or by any (even linear) perturbation of s. More precisely, if instead of (3.42) we consider a larger sets of the kind $\mathcal{F}=\mathcal{F}_L$ or $\mathcal{F}=\left\{F\in\mathcal{C}^1(\mathbf{X}\times\mathbf{U}) \,:\, s=s^\circ\right\}$, endowed with the correspondent subset topologies, then Proposition 3.9 does not apply anymore and, as a simple counter-example can easily show, the property of P_0-robustness is lost.

Robustness in the P_T Sense

The limits of the P_0-robustness of the linear regulator motivate seeking for a regulation objective P for which a regulator that is P-robust relatively to more general topological spaces $(\mathcal{F},\tau_{\mathcal{F}})$ more likely could be constructed. As in (Astolfi et al., 2015; Astolfi and Praly, 2017), in this section we let $\mathcal{F}=\mathcal{C}^1(\mathbf{X}\times\mathbf{U})$, $\tau_{\mathcal{F}}=\tau_{\mathcal{C}^1}$ and we consider a regulator of the kind (3.40) with (ϕ,θ) possibly nonlinear and with $n_{\mathrm{im}} = (2d+1)n_e$, for some arbitrary $d \in \mathbb{N}$. We then choose a basis for $\mathbb{R}^{\mathbf{n}} \times \mathbb{R}^{n_\eta}$ in which (Φ, G) reads as

$$\Phi = \begin{pmatrix} 0 & I & 0 & \cdots & 0 \\ 0 & 0 & I & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & & & & I \\ -a_1 I & -a_2 I & \dots & & -a_{2d+1} I \end{pmatrix} \qquad G = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ I \end{pmatrix} \tag{3.44}$$

with $a_1, \dots, a_{2d+1}$ chosen so that the characteristic polynomial of Φ is

$$p_\Phi(\lambda) = \lambda \cdot \prod_{k=1}^{d} \left(\lambda^2 + \omega_k^2\right), \tag{3.45}$$

where we let

$$\omega_k := 2\pi k/T \tag{3.46}$$

for some $T \geq 0$. For a given continuous function $\alpha : \mathbb{R}_+ \to \mathbb{R}$, let

$$\mathrm{c}_k(\alpha) := \int_0^T \alpha(\nu) e^{-i2\pi k\nu/T} d\nu \tag{3.47}$$

be the Fourier coefficient corresponding to the k-th harmonic $2\pi k/T$ and let

$$\mathcal{Q}_d := \left\{ \alpha : \mathbb{R} \to \mathbb{R} \, : \, c_k(\alpha) = 0, \; k = 0, \dots, d \right\}$$

be the subspace of the functions $\mathbb{R} \to \mathbb{R}$ that have null harmonics at $\omega_k = 2\pi k/T$. We define the regulation objective:

$$\mathrm{P}_T^{weak} \quad = \quad \text{“ } \eta \text{ is not T-periodic or } e \in (\mathcal{Q}_d)^{n_e} \text{ ”}$$

and, with $F^\circ \in \mathcal{C}^1(\mathbf{X} \times \mathbf{U})$ and $\mathbf{X}_0^\circ \in \mathcal{X}$, we make the following assumption

Assumption 3.3. *There exists a $\tau_{\mathcal{C}^1 \times \mathcal{X}}$-neighborhood V of $(F^\circ, \mathbf{X}_0^\circ)$ that generates an equibounded family of systems.*

Then the following result holds:

Proposition 3.10. *Suppose that Assumption 3.3 holds. Then the regulator is P_T^{weak}-robust at $(F^\circ, \mathbf{X}_0^\circ)$ with respect to $\tau_{\mathcal{C}^1}$.*

Proof. Pick $(F, \mathbf{X}_0) \in V$ and let (w, x, η) be any solution in $\mathcal{S}_F(\Omega_F(\mathbf{X}_0 \times H_0))$. If η is not T-periodic then P_T^{weak} trivially holds, so we assume η periodic. In view of (3.44), if we partition η as $\eta = \mathrm{col}(\eta_1, \ldots, \eta_{2d+1})$ with $\eta_\ell = \mathrm{col}(\eta_{\ell 1}, \ldots, \eta_{\ell n_e}) \in \mathbb{R}^{n_e}$ for $\ell = 1, \ldots, 2d+1$, then we have $\dot{\eta}_\ell = \eta_{\ell+1}$ for all $\ell = 1, \ldots, 2d$ and hence, for each $j = 1, \ldots, n_e$ we obtain

$$\eta_{1j}^{(2d+1)} + a_{2d+1}\eta_{1j}^{(2d)} + \cdots + a_2\dot{\eta}_{1j} + a_1\eta_{1j} = e_j. \tag{3.48}$$

Integrating by parts, for $k \in \mathbb{N}$ and $n = 1, \ldots, 2d+1$ we have

$$\begin{aligned} c_k\left(\eta_{1j}^{(n)}\right) &= \int_0^T \eta_{1j}^{(n)}(\nu)e^{-i2\pi k\nu/T}d\nu \\ &= \left[\eta_{1j}^{(n-1)}(t)e^{-i2\pi kt/T}\right]_0^T + \frac{i2\pi k}{T}c_k\left(\eta_{1j}^{(n-1)}\right). \end{aligned}$$

As the first term is zero ($\eta_{1j}^{(n-1)}$ is T-periodic), then by induction on n we obtain

$$c_k\left(\eta_{1j}^{(n)}\right) = \lambda_k^n c_k\left(\eta_{1j}\right),$$

with $\lambda_k := i2\pi k/T$. From (3.48) we thus obtain

$$c_k(e_j) = \left(\lambda_k^{(2d+1)} + a_{2d+1}\lambda_k^{(2d)} + \cdots + a_2\lambda_k + a_1\right)c_k\left(\eta_{1j}\right).$$

Therefore, if $k \leq d$, by definition of $(a_\ell)_\ell$, λ_k solves (3.45) and hence $c_k(e_j) = 0$ for all $j = 1, \ldots, n_e$, and the claim follows. ∎

Proposition 3.10 states that as long as a steady-state is defined, either the closed-loop solutions converge to a solution where η is not periodic or asymptotically e has null mean value and null harmonics at ω_k, $k = 1, \ldots, d$. We can refine the result under additional assumption concluding P_T-robustness where

$$\mathrm{P}_T \quad = \quad \text{“ } e \in (\mathcal{Q}_d)^{n_e} \text{ ”}.$$

Let $z := \mathrm{col}(x, \eta)$ and let $g : \mathbb{R}^{n_w} \times \mathbb{R}^n \times \mathbb{R}^{n_\eta} \to \mathbb{R}^{n+n_\eta}$ be such that (3.39) with $F = F^\circ$ can be rewritten as

$$\dot{w} = s(w), \qquad \dot{z} = g(w, z).$$

Then the following holds:

Proposition 3.11. *Assume that the z subsystem is 0-locally asymptotically stable and the origin of the w subsystem is stable. Suppose, moreover, that there exists a $\tau_{\mathcal{C}^1\times\mathcal{X}}$-neighborhood V of $(F^\circ, \{0\})$ such that for all $(F, \mathbf{X}_0) \in V$ and all $(\mathbf{x}, \eta) \in \mathcal{S}_F(\mathbf{X}_0 \times H_0)$, w is T-periodic. Then the regulator is P_T-robust at $(F^\circ, \{0\})$ with respect to $\tau_{\mathcal{C}^1}$.*

Proof. It suffices to show that if z is 0-LES and w is periodic and sufficiently small then the closed loop trajectories are periodic, since in this case the result would follow from Proposition 3.10, as P_T^{weak} implies P_T.

Since the z subsystem is 0-LES, standard Lyapunov arguments can be used to show that, for each $\mu > 0$ there exist $\bar{z}(\mu), \bar{w}(\mu) > 0$ such that each solution to $\mathcal{H}_{F^\circ}$ satisfying $|w|_\infty \leq \bar{w}(\mu)$ and $|z(0)| \leq \bar{z}(\mu)$ also satisfies $|z|_\infty \leq \mu$. In particular, the fact that the origin of the w subsystem is stable guarantees that this set of solutions is not empty.

Let $A := (\partial g(0,0))/(\partial z)$ and $B := (\partial g(0,0))/(\partial w)$ and let $\tilde{g}(w,z) := g(w,z) - (Az + Bw)$. Let

$$\tilde{g}'(w,z) := \frac{\partial \tilde{g}(w,z)}{\partial(w,z)}.$$

Since $\tilde{g}'(0,0) = 0$ and $\tilde{g}'(w,z)$ is continuous then $\lim_{(w,z)\to 0} \tilde{g}'(w,z) = 0$. Hence, for all $\epsilon > 0$, there exists $\delta(\epsilon) > 0$ such that $|(w,z)| \leq \delta(\epsilon)$ implies $|\tilde{g}'(w,z)| \leq \epsilon$.

Since, moreover, $\tilde{g}'(w,z)$ is the (unique) linear map satisfying

$$\lim_{|h|\to 0} \frac{|\tilde{g}((w,z)+h) - \tilde{g}(w,z) - \tilde{g}'(w,z)|}{|h|} = 0$$

with $h \in \mathbb{R}^{n_w+n_z}$, then there exists a function $\rho : \mathbb{R}^{n_w+n_z} \to \mathbb{R}^{n_w+n_z}$ fulfilling $\lim_{|h|\to 0} |\rho(h)|/|h| = 0$ such that

$$\tilde{g}((w,z)+h) - \tilde{g}(w,z) = \tilde{g}'(w,z)h + \rho(h). \tag{3.49}$$

This, in particular, implies that for each $\epsilon > 0$, there exists $\sigma(\epsilon) > 0$ such that, for all $|h| \leq \sigma(\epsilon)$, $|\rho(h)| \leq \epsilon|h|$. Let

$$\gamma(\epsilon) := \max\left\{\delta(\epsilon/2), \sigma(\epsilon/2)/2\right\}.$$

Then for all $z_1, z_2 \in \mathbb{R}^{n_z}$ such that $|z_1| \leq \gamma(\epsilon)$ and $|z_2| \leq \gamma(\epsilon)$, we have $|z_2| \leq \delta(\epsilon/2)$

and $|z_1 - z_2| \le |z_1| + |z_2| \le \sigma(\epsilon/2)$, and hence, in view of (3.49) computed at (w, z_2) and with $h = z_1 - z_2$, we obtain

$$\begin{aligned} |\tilde{g}(w, z_1) - \tilde{g}(w, z_2)| &\le |\tilde{g}'(w, z_2)| \cdot |z_1 - z_2| + |\rho(z_1 - z_2)| \\ &\le (\epsilon/2)|z_1 - z_2| + (\epsilon/2)|z_1 - z_2| \le \epsilon|z_1 - z_2|. \end{aligned}$$

As $\mathcal{H}_{F^\circ}$ is 0-LES, A is Hurwitz. Let $P = P^T > 0$ be such that $A^T P + PA = -2I$ and let

$$\epsilon := 1/(2|P|\sqrt{2}), \qquad \mu := \gamma(\epsilon). \tag{3.50}$$

With $\mathcal{D} \subset \mathbb{R}^{n_z}$ the domain of attraction of the exponential stability of the origin of the z subsystem, let $Z_0 \subset \mathbb{R}^{n_w}$ be the maximal compact set such that $Z_0 \subset \{z \in \mathcal{D} : |z| \le \bar{z}(\mu)\}$, then Z_0 has non-empty interior. Let W_0 be such that any solution to the w subsystem originating in W_0 is T-periodic and fulfills $|w|_\infty \le \bar{w}(\mu)$ (which is not empty due to the assumptions). Then the projection of set $W_0 \times Z_0$ onto $\mathbb{R}^{n_w} \times \mathbb{R}^n$ contains a $\tau_{\mathcal{X}}$-neighborhood of $\{0\}$.

Let $V(z) := z^T P z$ and pick a solution (w, z) to $\mathcal{H}_{F^\circ}$ originating in $W_0 \times Z_0$. Let $U(t) := V(\tilde{z}(t))$, $\tilde{z}(t) := z(t+T) - z(t)$, then, as $w(t) - w(t+T) = 0$, we have

$$\begin{aligned} \dot{U}(t) &= 2\tilde{z}(t)^T P(A\tilde{z}(t) + \tilde{g}(w(t), z(t)) - \tilde{g}(w(t), z(t+T))) \\ &\le -|\tilde{z}(t)|^2 + |2P|^2 \cdot |\tilde{g}(w(t), z(t)) - \tilde{g}(w(t), z(t+T))|^2. \end{aligned}$$

Since $|w(t)| \le \bar{w}(\mu)$ for all $t \in \mathbb{R}_+$ and $|z(0)| \le \bar{z}(\mu)$, with μ given by (3.50), then $|z(t)| \le \gamma(\epsilon)$ and $|z(t+T)| \le \gamma(\epsilon)$ with ϵ given by (3.50). Thus (3.49) implies $\dot{U}(t) \le -|\tilde{z}|^2 + |\tilde{z}|^2/2 \le -|\tilde{z}|^2/2$ and this suffices to show that $|z(t) - z(t+T)| \to 0$.

The analysis above suffices to conclude that, for each compact subsets of $W_0 \times Z_0$, any solution originating in the set $\Omega_{F^\circ}(W_0 \times Z_0)$ is periodic. Robustness with respect to changes of F in the topology τ_{C^1} follow from usual Lyapunov arguments. ■

Proposition 3.11 somewhat generalizes the result of (Astolfi et al., 2015), where they give though a design of the stabilizer able to achieve 0-local exponential stability based on forwarding arguments. Another refinement of Proposition 3.10 is given as follows.

Proposition 3.12. *Suppose that Assumption 3.3 holds for some $\tau_{\mathcal{F}\times\mathcal{X}}$-neighborhood V of $(F^\circ, \mathbf{X}_0^\circ)$ and suppose that, for each $(F, \mathbf{X}_0) \in V$, the corresponding system* (3.39) *satisfies Assumption 1.9. Suppose in addition that, for each solution* $(w^\star, x^\star, \eta^\star)$ *in* $\mathcal{S}_F(\Omega_F(\mathbf{X}_0 \times H_0))$*,* w *is* T*-periodic. Then the regulator is* P_T*-robust at* $(F^\circ, \mathbf{X}_0^\circ)$ *with respect to* $\tau_{\mathcal{F}}$*.*

Proof. The result follows directly from Theorem 1.2. As a matter of fact, under the assumptions of the proposition, Theorem 1.2 implies that for each $(F, \mathbf{X}_0) \in V$, $\Omega_F(\mathbf{X}_0 \times H_0)$ is the graph of a continuous function $\pi : \operatorname{dom}\pi \subset \mathbb{R}^{n_w} \to \mathbb{R}^{n+n_\eta}$, i.e. each $(w^\star, x^\star, \eta^\star) \in \mathcal{S}_F(\Omega_F(\mathbf{X}_0 \times H_0))$ is of the form $(x^\star(t), \eta^\star(t)) = \pi(w^\star(t))$. As $w^\star$ is T-periodic, then so is $\eta^\star$, and the claim follows from Proposition 3.10. ■

Quasi-Periodic Robustness

We can deal with quasi-periodic responses in the same way as in the previous section. In the following we use the same linear regulator as before, with (Φ, G) given by (3.44) and with (3.45) that still holds, with the ω_k's that are though arbitrarily chosen in $\mathbb{R}_+$ and where we also let for convenience $\omega_0 := 0$. With $\alpha : \mathbb{R} \to \mathbb{R}_+$ and for $k = 0, \dots, d$, let define the (generalized) Fourier coefficients as

$$c'_k(\alpha) := \lim_{T\to\infty} \frac{1}{T} \int_0^T \alpha(\nu) e^{-i\omega_k \nu} d\nu$$

and let

$$\mathcal{Q}'_d := \Big\{ \alpha : \mathbb{R} \to \mathbb{R} \ : \ c'_k(\alpha) = 0,\ k = 0, \dots, d \Big\}.$$

Proceeding as before, we let

$$\mathrm{P}_{\mathrm{qp}}^{weak} \ ; = \quad \text{“ } \eta \text{ is not quasi-periodic or } e \in (\mathcal{Q}'_d)^{n_e} \text{ ”}.$$

Then the following result holds:

Proposition 3.13. *Suppose that Assumption 3.3 holds. Then the regulator is* $\mathrm{P}_{\mathrm{qp}}^{weak}$*-robust at* $(F^\circ, \mathbf{X}_0^\circ)$ *with respect to* $\tau_{\mathcal{C}^1}$*.*

Proof. The proof follows from the same argument of those of Proposition 3.10 once noted that, for each $n = 1, \dots, 2d+1$, each $j = 1, \dots, n_e$ and each $k = 0, \dots, d$, we have

$$\begin{aligned} c'_k\left(\eta_{1j}^{(n)}\right) &= \lim_{T\to\infty} \frac{1}{T} \int_0^T \eta_{1j}^{(n)}(\nu) e^{-i\omega_k \nu} d\nu \\ &= \lim_{T\to\infty} \frac{1}{T} \left[\eta_{1j}^{(n-1)}(t) e^{-i\omega_k t} \right]_0^T + i\omega_k c'_k\left(\eta_{1j}^{(n-1)}\right) \end{aligned}$$

where the integrals converge whenever η is quasi-periodic and with the first term that vanishes as $\eta_{1j}^{(n-1)}(t)e^{-i\omega_k t}$ is bounded and with $i\omega_k$ that solves (3.45) ■

3.4.4 Robustness in Nonlinear Regulators

In this section we focus on output regulation for nonlinear systems. For simplicity, we refer to the Marconi-Praly-Isidori regulator (Marconi et al., 2007), however the same conclusions apply to the Byrnes-Isidori one (Byrnes and Isidori, 2004) and to the related extensions. We also limit to the SISO case, in which $y = e$ and $m = n_y = 1$, and where the state of the plant can be decomposed as $x = \text{col}(z, e)$, with $z \in \mathbb{R}^{n-1}$, and with $f(\cdot)$ that is defined so as the plant has the following structure

$$\begin{aligned} \dot{z} &= g(w, z, e) \\ \dot{e} &= q(w, z, e) + b(w, z, e)u \end{aligned} \tag{3.51}$$

for some $g : \mathbb{R}^{n_w} \times \mathbb{R}^n \to \mathbb{R}^{n-1}$, $q, b : \mathbb{R}^{n_w} \times \mathbb{R}^n \to \mathbb{R}$ and with $h(w, x) = e$. The regulator design is made assuming that the initial conditions of (3.51) range in an arbitrary compact set $W \times Z \times E \subset \mathbb{R}^{n_w} \times \mathbb{R}^n$, which we will assume fixed from now on. In line with the previous sections, we let $\mathbf{X} \in \mathcal{K}(\mathbb{R}^{n_w} \times \mathbb{R}^n)$ and $\mathbf{U} \in \mathcal{K}(\mathbb{R})$ be arbitrarily large compact sets such that $W \times Z \times E \subset \mathbf{X}$ and with $\mathbf{U}$ taken sufficiently large to encompass all the solutions of interest. We also adapt the definitions of $\mathcal{C}^1(\mathbf{X} \times \mathbf{U})$ and $\tau_{\mathcal{C}^1}$ to this case accordingly. As basic assumptions on the plant's data, we assume the following

Assumption 3.4. *W is invariant for the exosystem w and Assumption 1.2 holds*

As the structure of (3.51), the invariance of W and the minimum phase assumption are a properties of $F = (s, f, h)$, we let $\mathcal{F}$ be the set of functions in $\mathcal{C}^1(\mathbf{X} \times \mathbf{U})$ for which these properties holds. The functions $\phi(\cdot)$ and $\theta(\cdot)$ of (3.36) are chosen so as

$$\begin{aligned} \dot{\eta} &= F\eta + G\big(\gamma(\eta) + \kappa(e)\big) \\ u &= \gamma(\eta) + \kappa(e), \end{aligned} \tag{3.52}$$

with H_0 arbitrary, (F, G) a controllable pair and with $n_\eta \in \mathbb{N}$, $\gamma : \mathbb{R}^{n_\eta} \to \mathbb{R}$ and $\kappa : \mathbb{R} \to \mathbb{R}$ that are continuous functions chosen on the basis of W, Z, E and $F^\circ \in \mathcal{F}$ by following the procedure of (Marconi et al., 2007). The set $H_0 \subset \mathbb{R}^{n_\eta}$ is instead chosen arbitrarily. The main result of (Marconi et al., 2007) can be framed in the language of this framework as follows:

Proposition 3.14. *Let $\tau_{\mathcal{F}}$ be any topology containing $\{F^\circ\}$. With P_0 defined as in* (3.43)*, $F^\circ \in \mathcal{F}$ and $\mathbf{X}_0^\circ := W_0^\circ \times Z_0^\circ \times E_0^\circ \subset \mathrm{int}(W \times Z \times E)$, the regulator* (3.52) *is P_0-robust at $(F^\circ, \mathbf{X}_0^\circ)$ with respect to $\tau_{\mathcal{F}}$.*

In other words, Proposition 3.14, which is indeed the result of (Marconi et al., 2007), states that the P_0-robustness is proved only for perturbations of $(F^\circ, \mathbf{X}_0^\circ)$ that keep F° constant and $\mathbf{X}_0^\circ$ inside the set $W \times Z \times E$. As a matter of fact, as $\mathbf{X}_0^\circ$ is in the interior of $W \times Z \times E$, we can always find a $\tau_{\mathcal{X}}$-neighborhood of $\mathbf{X}_0^\circ$ that stays inside $W \times Z \times E$. Hence to construct the $\tau_{\mathcal{F}\times\mathcal{X}}$-neighborhood V of Definition 3.4 we just need to find a $\tau_{\mathcal{F}}$-neighborhood (call it N) of F° for which P_0 holds. The result of (Marconi et al., 2007) states indeed that the choice $N = \{F^\circ\}$ works. The fact that also slight $\mathcal{C}^1$ perturbations of F° might destroy the property P_0 can be verified by means of simple counter examples. Even though a general result on the fragility of P_0 with respect to general $\mathcal{C}^1$ variations is not known at present, we believe that it holds true. In particular, we are led to believe that the property P_0 is in its nature "nominal" and cannot be preserved, in general, under arbitrary (even if small) variations. This belief is included as a particular case in the forthcoming conjecture stated below.

Robustness to more general variations is instead possible when an "approximate" regulation goal is considered. More precisely, for any $\varepsilon > 0$, let

$$\mathrm{P}_\varepsilon \;=\; \text{“}\, |e|_\infty < \varepsilon \,\text{”}.$$

Let $\tau_{\mathcal{F}}$ be now the subset topology induced by $(\mathcal{C}^1(\mathbf{X} \times \mathbf{U}), \tau_{\mathcal{C}^1})$, and let the regulator (3.36) be chosen so that P_ε holds at $(F^\circ, \mathbf{X}_0^\circ)$, with $F^\circ \in \mathcal{F}$ and $\mathbf{X}_0^\circ \subset \mathrm{int}(W \times Z \times E)$ compact. Then the following result comes from (Marconi et al., 2007, Thm. 2) as a consequence of the continuity of F, γ and κ:

Proposition 3.15. *The regulator* (3.51) *is P_ε-robust at $(F^\circ, \mathbf{X}_0^\circ)$ with respect to $\tau_{\mathcal{F}}$.*

3.4.5 In general, no regulator is P_0-robust

In this chapter we dealt with the robustness issue in output regulation schemes, by proposing a new milder definition of robustness relative to a steady-state property that generalizes the classical regulation goal of reaching a steady state in which the regulation errors are identically zero. We reviewed some of the main regulation schemes in linear and nonlinear frameworks, by showing, in relevant cases, what kind of robustness properties are meet. Starting from the special P_0-robustness property of the linear regulator, throughout the milder P_T-robustness for nonlinear systems up to the weak robustness property of nonlinear design of Proposition 3.14, we arrived to the claim that P_0-robustness might not be the most appropriate goal in general nonlinear regulation. We conclude this discussion with a conjecture saying that, even if the exosystem is known, no finite-dimensional regulator can guarantee asymptotic regulation under general "$\mathcal{C}^1$ variations" of the plant: with s° the nominal (known) exosystem function and with $\mathbf{X} \in \mathcal{K}(\mathbb{R}^{n_w} \times \mathbb{R}^n)$ and $\mathbf{U} \in \mathcal{K}(\mathbb{R}^m)$ arbitrary compact sets, let $\mathcal{F} := \{F \in \mathcal{C}^1(\mathbf{X} \times \mathbf{U}) \ : \ s = s^\circ\}$ and let $\tau_{\mathcal{F}}$ be the subset topology derived by $(\mathcal{C}^1(\mathbf{X} \times \mathbf{U}), \tau_{\mathcal{C}^1})$. Then we state the following

Conjecture 3.1. *Let $F^\circ \in \mathcal{F}$ and $\mathbf{X}_0^\circ \in \mathcal{K}(\mathbb{R}^{n_w} \times \mathbb{R}^n)$. Then no regulator of the kind* (3.36) *is P_0-robust at $(F^\circ, \mathbf{X}_0^\circ)$ with respect to $\tau_{\mathcal{F}}$.*

Conclusion

In this chapter we reviewed the state of the art of output regulation for nonlinear systems and we presented new contributions in terms of structural and robustness aspects. On the structural side, we highlighted how the whole nonlinear regulation literature has avoided in these years dealing with the chicken-egg dilemma, incurring in the structural obstructions of pre-processing schemes or in approaches that sacrifice asymptotic regulation. We gave sufficient conditions for the existence of a regulator of the post-processing type that overcomes some of the main limitations of pre-processing schemes. The proposed regulator, indeed, deals with more inputs than regulation errors and allows for additional non-vanishing measured outputs to be used in the closed-loop stabilization. On the robustness side, we proposed a framework in which robustness relatively to a general notion of perturbation and to general steady-state properties can be analysed. We re-stated in that framework some of the main existing results in the filed of linear and nonlinear regulation, and we arrived to conjecture that, at least for what concerns the "canonical" C^1 perturbations, asymptotic regulation is a fragile property that cannot be achieved robustly with a finite-dimensional regulator. We also presented a regulator based on low-power high-gain observers and a systematic procedure to deal with structural robustness for some classes of problems, and we recover the main elements of the "non-equilibrium theory" of output regulation of (Byrnes and Isidori, 2003) in the case of exosystems modeled by differential inclusions.

Output regulation is still quite an open problem, and the results presented here are still far to be a definite answer. For instance, the post-processing design proposed in Section 2.3 guarantees asymptotic regulation only in the case in which there actually exist a dimension d and a function ϕ such that the equation (2.22) holds. This condition, though, might not hold in general for any finite d and, even if such d and ϕ existed, they are not easy to find at all. Moreover, unless some particular immersion properties hold, their use is not robust, as the functions v_i, $i = 1, \dots, d + 1$, strongly depend on the plant's data. The next chapters are dedicated to construct a framework in which *adaptation* can be systematically included in the design of regulators. In this way we aim at building on the theory developed so far to create *constructive* regulator designs, that use adaptation to deal with the chicken-egg dilemma and to confer on the control system additional robustness properties.

Part II

Identification and Control

4

A Framework for Identifiers

System Identification is a classical branch of control theory, devoted to the creation of mathematical models from observations (Ljung, 1999; Tangirala, 2015). The interactions between system identification and control have been present since the dawn of times, especially in adaptive control problems, and they have generated interesting contributions generally grouped under the common name of *"Identification for Control"* (see e.g. (Gevers, 1993, 1996) for an historical perspective and (Gevers, 2005; Hjalmarsson, 2005) for excellent excursions on the results and open problems). The main aim of the literature of identification for control is the synergistic *co-design* of control systems and system identification schemes to create control solutions that can operate autonomously. Inside this large field we find for instance the theory of *dual control* (Feldbaum, 1960), in which the control laws have to be designed to pursue the *dual* goal of achieving the desired closed-loop behavior and inducing in the system the "right" movements allowing a meaningful identification of the dynamics of interest; we also find the theory of *iterative identification* (Gevers, 1996), in which phases of free evolution of the closed-loop system alternate with off-line identification phases; last but not least, we find the application of identification to *robust control* (Hjalmarsson, 2005), where identification has the role of reducing the "uncertainty ellipsoid" where the *true* system lies inside the perturbation radius in which the robust control system can guarantee the desired steady-state performances.

The interaction among identification and robust control is perhaps the most fortunate and developed field of identification for control, and, in general, the community has mainly focused on stabilization problems leaving fields like observation and regulation barely unexplored from the system identification perspective and at the mercy of the *adaptive control* community. Nevertheless, the ideas at the core of the synergistic design of identification and control reveal a much deeper significance that makes it worth investigating in broader contexts. *In the adaptive designs presented in this thesis we chose to approach adaptation as a system identification problem.* On a theoretical side, this allows us to consider more general concepts of "uncertainty" and of "model" than those typically considered in the adaptive control field, where the uncertainty is usually assumed to be concentrated into a single parameter of known dimension. On the practical side, in this way we have access to well-known and well-developed identification techniques (parametric and not), that we can use (and inherit their robustness properties) instead of relying on ad hoc adaptation laws, as it is instead typical of adaptive control approaches.

As a preliminary work towards this goal, in this chapter we define a system theoretical framework in which general "online" identification problems can be cast, with the identification algorithms that are seen as dynamical systems and treated in the context of nonlinear system theory (from now on we refer to these systems as *identifiers*). Identification problems are cast as optimization problems and the identifiers are asked to satisfy strong stability requirements with respect to an ideal steady state defined by the optimal solution of the underlying identification problem. For simplicity, we develop the framework in a deterministic setting, as the results we give here do not require any further characterization of the errors.

The framework developed in this chapter will be used Chapter 5, where we show how the identifiers that fit in the proposed framework can be used in conjunction with a high-gain observer to design adaptive observers for nonlinear systems, and in Chapters 6 and 7, where adaptive internal models will be developed. In both the applications it is shown that the asymptotic control performance (state estimation error in the case of observation and regulation error in case of regulation) is directly related to the *prediction* performances of the identified model, evaluated along the ideal steady state of the overall system.

The identification framework proposed in this chapter is a *hybrid systems framework* (see Goebel et al., 2012). In this way we can deal with both continuous-time and discrete-time identification algorithms in a common formal playground. In Sections 4.2 and 4.3 we show that ordinary *least-squares* problems fit into the framework, in Section 4.4 we consider algorithms for nonlinear parametrizations and in Section 4.5 we show how the proposed identifiers can be used to perform an online multiresolution non-parametric *wavelet identification*.

This chapter contains original contribution submitted for publication in (Bin and Marconi, 2018c). For what concerns the notation on hybrid system, we refer to (Goebel et al., 2012) and to Appendix A. For what concerns the terminology related to system identification we refer to (Ljung, 1999). We denote by $\langle\cdot,\cdot\rangle$ the canonical scalar product in $L_2(\mathcal{X})$ and we equip $L_2(\mathcal{X})$ with the norm $|\cdot| := \sqrt{\langle\cdot,\cdot\rangle}$.

4.1 The Framework

In this section we give a formal definition of the class of identifiers that we will treat in the rest of the chapter. We describe *identifiers* as hybrid dynamical systems and we express the main desired properties in terms of a requirement, the *identifier requirement*, asking for strong stability properties with respect to an ideal steady state defined by the underlying optimization problem.

4.1.1 Hybrid Identifiers

Let $\mathcal{A}$ and $\mathcal{B}$ be normed vector spaces and let $\alpha^\star$ and $\beta^\star$ be hybrid inputs with the same hybrid time domain and with values respectively in $\mathcal{A}$ and $\mathcal{B}$. The general problem that we consider in this section is to find a "good" *prediction model* relating the values of $\alpha^\star(t,j)$ and $\beta^\star(t,j)$ on the basis of their available measurements. More precisely, we aim at finding a function $\mathcal{A}\to\mathcal{B}$ that, at each $(t,j)\in\operatorname{dom}(\alpha^\star,\beta^\star)$, produces a "good" guess $\hat{\beta}^\star(t,j)$ of the value of $\beta^\star(t,j)$ given the knowledge of $\alpha^\star(t,j)$. As usually done in system identification (Ljung and

Söderström, 1985), we look for candidate functions inside a parametrized set of functions $\mathcal{M}$ whose elements are indexed by a vector θ ranging in $\mathbb{R}^d$, with $d \in \mathbb{N}^*$ that is a degree of freedom called the *model order*. The set $\mathcal{M}$ is usually called the *model set* and, by definition, it naturally induces a map $\Phi : \mathbb{R}^d \times \mathcal{A} \to \mathcal{B}$ such that, as θ ranges in $\mathbb{R}^d$, $\Phi(\theta, \cdot)$ generates all the models in $\mathcal{M}$. Thus, every possible model ϕ that we consider is given by $\phi(\cdot) = \Phi(\theta, \cdot)$ for some $\theta \in \mathbb{R}^d$. We refer to Φ as the *prediction model* and, in the following, we make reference to $\Phi(\theta, \cdot)$ to denote a candidate model.

The choice of model order d, of the model set $\mathcal{M}$ and of the procedure used to select θ is typically based on some a-priori knowledge of the signals $\alpha^\star$ and $\beta^\star$. In control applications the signals $\alpha^\star$ and $\beta^\star$ are usually defined by some combinations of the plant's state variables (this is for example the case of the observer developed later in Chapter 5), and the aforementioned degree of freedom are decided by using the a priori knowledge of the plant. Here we "encode" the available information on $\alpha^\star$ and $\beta^\star$ by assuming that they are generated by a hybrid inclusion of the form

$$\mathcal{H}_w : \begin{cases} \dot{w} & \in & S(w) & \quad w \in \mathcal{C}_w \\ w^+ & \in & R(w) & \quad w \in \mathcal{D}_w \end{cases} \tag{4.1}$$

with state w taking values in a normed vector space $\mathcal{W}$ and with output[1]

$$\alpha^\star = \alpha^\star(w), \quad \beta^\star = \beta^\star(w)\,, \tag{4.2}$$

being $\mathcal{C}_w$ and $\mathcal{D}_w$ closed subsets of $\mathcal{W}$ and $\alpha^\star : \mathcal{W} \to \mathcal{A}$ and $\beta^\star : \mathcal{W} \to \mathcal{B}$ locally bounded functions. We refer to the process (4.1) as the *exosystem*. Throughout the rest of the chapter we make the following standing assumption (Goebel et al., 2012)

Assumption 4.1. *$\mathcal{H}_w$ satisfies the hybrid basic conditions of Definition A.1.2, is forward complete and $W := \mathcal{C}_w \cup \mathcal{D}_w$ is compact.*

Once the model order d and the model set $\mathcal{M}$ are fixed, the problem boils down to choose the particular value of the parameter θ for which the corre-

[1]With the aim of avoid introducing too many symbols, despite the slight abuse of notation, we will denote with the same letters $\alpha^\star$ and $\beta^\star$ both the hybrid arcs introduced above and the functions in (4.2).

sponding model $\Phi(\theta, \cdot) \in \mathcal{M}$ has the desired representation capabilities. Here we restrict the attention to *online* algorithms (Ljung and Söderström, 1985), i.e. processes that produce and adapt a guess of θ based on some measures that are made available at run time on the signals $\alpha^\star$ and $\beta^\star$. We call processes of this kind *identifiers* and we describe them as hybrid systems driven by the signals $\alpha^\star$ and $\beta^\star$. What information is actually available on $(\alpha^\star, \beta^\star)$ depends on the definition of the particular underlying problem and the particular algorithm that is implemented. For instance, a discrete-time identifier will typically work on samples of $(\alpha^\star, \beta^\star)$, that corresponds to the values $(\alpha^\star(t_j, j), \beta^\star(t_j, j))$, $j \in \mathbb{N}$, while a continuous-time identifier will more likely work on the values of $(\alpha^\star, \beta^\star)$ during the flow intervals (i.e. those assumed at (t, j) with $t \in (t_j, t^j)$).

To support the forthcoming application to adaptive observers design we further complicate the problem by letting the measures of $\alpha^\star$ and $\beta^\star$ available to the identifier be corrupted by an additive unmodeled disturbance. Namely, we suppose that the identifier has access to the hybrid signal (α, β), where

$$\begin{aligned} \alpha(t, j) &= \alpha^\star(t, j) + \delta_\alpha(t, j) \\ \beta(t, j) &= \beta^\star(t, j) + \delta_\beta(t, j), \end{aligned} \tag{4.3}$$

being $\delta := (\delta_\alpha, \delta_\beta)$ a bounded hybrid input on $\mathcal{A} \times \mathcal{B}$. In this respect we better frame in the field of *errors-in-variables* identification (Söderström, 2007).

We define an identifier as a hybrid system of the form

$$\mathcal{H}_z : \begin{cases} \dot{z} \in F(z, \alpha, \beta) & (z, \alpha, \beta) \in \mathcal{Z} \times \mathcal{A} \times \mathcal{B} \\ z^+ \in G(z, \alpha, \beta) & (z, \alpha, \beta) \in \mathcal{Z} \times \mathcal{A} \times \mathcal{B} \end{cases} \tag{4.4}$$

with state z taking values in a normed vector space $\mathcal{Z}$, input (α, β) given by (4.3), output

$$\theta = h(z), \tag{4.5}$$

and where $F, G : \mathcal{Z} \times \mathcal{A} \times \mathcal{B} \rightrightarrows \mathcal{Z}$ and $h : \mathcal{Z} \to \mathbb{R}^d$. To the identifier (4.4) we associate the prediction model

$$\Phi(\theta, \cdot) \in \mathcal{M},$$

which is such that, at each (t, j), $\hat{\beta}^\star(t, j) = \Phi(\theta(t, j), \alpha^\star(t, j))$ represents the identifier's best guess of $\beta^\star(t,j)$ given $\alpha^\star(t,j)$. The *overall* system, obtained by interconnecting (4.1) and (4.4), and by letting (α,β) be defined by (4.2)-(4.3), reads as follows

$$\mathcal{H}_{cl}: \begin{array}{l} \left\{ \begin{array}{rcl} \dot{w} & \in & S(w) \\ \dot{z} & \in & F(z, \alpha^\star(w)+\delta_\alpha, \beta^\star(w)+\delta_\beta) \end{array} \right. \qquad (w,z,\delta)\in\mathcal{C}\times\mathcal{A}\times\mathcal{B} \\ \left\{ \begin{array}{rcl} w^+ & \in & R(w) \\ z^+ & \in & G(z, \alpha^\star(w)+\delta_\alpha, \beta^\star(w)+\delta_\beta) \end{array} \right. \qquad (w,z,\delta)\in\mathcal{D}\times\mathcal{A}\times\mathcal{B} \end{array} \tag{4.6}$$

where $\mathcal{C} := \mathcal{C}_w \times \mathcal{Z}$ and $\mathcal{D} := \mathcal{D}_w \times \mathcal{Z}$.

Definition 4.1. *The identifier* (4.4) *is said to be* well-defined *if it satisfies the hybrid basic conditions of Definition A.1.2 and the overall system* (4.6) *is forward complete.*

Even if Definition 4.1 involves the entire system (4.6), we remark that, in view of Assumption 4.1, being well-defined is only a property of the identifier. The flow and jump sets of (4.4) coincide with the whole state space $\mathcal{Z}$ and, hence, the time domain of the solutions of the overall system is only decided by the exosystem. This allows us to conclude that, for a well-defined identifier, (w,z,δ) is a solution pair of $\mathcal{H}_{cl}$, only if w is a solution of $\mathcal{H}_w$. If the solutions of $\mathcal{H}_{cl}$ are needed to satisfy certain constraints in terms of time domain, a *clock* can be included in the exosystem description. More in general, in the rest of the chapter, whenever we want to restrict the attention to the solutions of $\mathcal{H}_{cl}$ that fulfills some given properties, we will refer to a suitably defined subset $\mathcal{E}$ of $\mathcal{S}_{\mathcal{H}_{cl}}$ that we call a *restriction* on $\mathcal{H}_{cl}$.

4.1.2 The Identification Problem

We associate to system (4.6) the function $\varepsilon : W \times \mathbb{R}^d \to \mathcal{B}$ given by

$$\varepsilon(w,\theta) := \beta^\star(w) - \Phi(\theta, \alpha^\star(w)) \,. \tag{4.7}$$

For a given solution w of $\mathcal{H}_w$ and a given $(t,j)\in \operatorname{dom} w$, $\varepsilon(w(t,j),\theta)$ represents the *prediction error* obtained by the model $\Phi(\theta,\cdot)\in\mathcal{M}$ corresponding to the *uncorrupted* data pair $(\alpha^\star(w(t,j)), \beta^\star(w(t,j)))$. While the model set $\mathcal{M}$ can be defined on the basis of the a priori knowledge of (4.1), the design of the identifier (4.4) is usually cast as an optimization problem involving the minimization of some function of the history of the prediction errors (4.7) along the data set. By following this line, with $\mathbb{F}(\mathbb{R}^d)$ denoting the set of functionals $\mathbb{R}^d \to \mathbb{R}_+$, we associate to each solution w to $\mathcal{H}_w$ a *cost functional* $J_w : \operatorname{dom} w \to \mathbb{F}(\mathbb{R}^d)$ of the form

$$J_w(\theta)(t,j) := \sum_{i=0}^{j-1} \mathrm{d}_{i,j}\big(\varepsilon(w(t^i,i),\theta)\big) + \int_0^t \mathrm{c}_{s,t}\big(\varepsilon(w(s,j_s),\theta)\big)ds + \omega(\theta) \tag{4.8}$$

with $\mathrm{d}_{i,j}, \mathrm{c}_{s,t}, \mathcal{B} \to \mathbb{R}_+$, $i,j \in \mathbb{N}$, $s,t \in \mathbb{R}$, some user-defined positive functions satisfying $d_{i,j}(0) = 0$, $c_{s,t}(0) = 0$ and $\omega(0) = 0$ that characterize the particular optimization problem. The summation and the integral terms penalize the prediction error obtained by the model $\Phi(\theta,\cdot)$ during the jump and flow times respectively, while $\omega(\theta)$ represents a regularization term. We associate to J_w the *optimum map* (which is possibly set-valued) $\vartheta_w^\circ : \operatorname{dom} w \rightrightarrows \mathbb{R}^d$ defined as

$$\vartheta_w^\circ(t,j) := \arg\inf_{\theta\in\mathbb{R}^d} J_w(\theta)(t,j)\,. \tag{4.9}$$

We will refer to the pair $(\mathcal{H}_w, J_w)$ formed by an exosystem $\mathcal{H}_w$ of the form (4.1)-(4.2) and a cost functional of the form (4.8) as an *identification problem*.

Once J_w is fixed, the design of the identifier is done so that its output θ minimizes (4.8) along the solutions of the overall system. Since the correct initialization of the identifier (4.4) is not in general known, a point-wise minimization is not in general feasible and a requirement in terms of an *asymptotic* optimal behavior is more suitable. Moreover, as the identifier works on a perturbed data set given by (α,β) rather than the "ideal" one corresponding to $(\alpha^\star,\beta^\star)$, it makes sense to require to the identifier an additional *robustness* property with respect to the disturbance δ. Under a system theoretical perspective, we look at the overall system (4.6) as a system with input δ, and the aforementioned requirements are seen as strong stability properties (also known as input-to-state stability (Sontag, 1989)) with respect to the input δ and relatively to an ideal optimal *steady state*. This is formalized in the definition below, to which we refer as the *identifier requirement*.

Requirement 4.1. *The identifier* (4.4)-(4.5) *is said to fulfil the* identifier requirement *relatively to the identification problem* $(\mathcal{H}_w, J_w)$ *and under the restriction* $\mathcal{E} \subset \mathcal{S}_{\mathcal{H}^{cl}}$ *if it is well-defined and there exist* $V_z : \mathcal{Z} \to \mathbb{R}_+$, $\underline{\sigma}, \bar{\sigma} \in \mathcal{K}_\infty$, $\rho, \kappa_1, \kappa_2 \in \mathcal{K}$, $\nu > 0$ *and, for each solution* w *of* $\mathcal{H}_w$ *an unique* $z^\star : \operatorname{dom} w \to \mathcal{Z}$, *such that for all*

$(w, z, \delta) \in \mathcal{E}$ *the following hold:*

1) Optimality: *The function* $\theta^\star := h(z^\star)$ *fulfills*

$$\theta^\star(t, j) \in \vartheta^\circ_w(t, j), \qquad \forall (t, j) \in \operatorname{dom} z^\star.$$

2) Stability: *with* $\tilde{z} := z - z^\star$ *it holds that*

   (a) For each $(t, j) \in \operatorname{dom} \tilde{z}$

$$\underline{\sigma}(|\tilde{z}(t, j)|) \le V_z(\tilde{z}(t, j)) \le \bar{\sigma}(|\tilde{z}(t, j)|)$$

   (b) For all $(t, j) \in \mathcal{I}(\tilde{z})$

$$V_z(\tilde{z}(t, j)) \ge \rho(|\delta(t, j)|) \implies D^+ V_z(\tilde{z}(t, j)) \le -\nu V(\tilde{z}(t, j))$$

   (c) For all $(t, j) \in \Gamma(\tilde{z})$

$$V_z(\tilde{z}(t, j+1)) \le \max\{e^{-\nu} V_z(\tilde{z}(t, j)),\ \rho(|\delta(t, j)|)\}$$

3) Regularity: *There exists* $T \ge 0$ *such that*

   (a) For all $(t, j) \in \mathcal{I}(w, z)|_{\ge T}$*,* $\dot{h}(z(t, j))$ *exists and*

$$|\dot{h}(z(t, j)) - \dot{h}(z^\star(t, j))| \le \kappa_1(|z(t, j) - z^\star(t, j)|)$$

   (b) For all $(t, j) \in \Gamma(w, z)|_{\ge T}$

$$|h(z(t, j)) - h(z^\star(t, j))| \le \kappa_2(|z(t, j) - z^\star(t, j)|).$$

Point 1 of the identifier requirement asks that the steady-state trajectory $z^\star$ is optimal in the sense that the corresponding output $\theta^\star$ minimizes (4.8). Point 2 is instead an *input-to-state stability* (see Appendix A.2) requirement relatively to the ideal steady state $z^\star$. The third point is a regularity requirement of the output map along the solutions (also interpreted as a detectability property). The third requirement typically is fulfilled under some *persistence of excitation* (see the Proposition 4.1), and we observe that in a purely discrete-time identifier the condition 3.a) is always satisfied, while a purely continuous-time identifier always satisfies 3.b). The restriction $\mathcal{E}$ has been introduced in the definition of the requirement to take into account the cases in which the requirement can only be achieved "locally" (both in the state and the disturbance). For instance, $\mathcal{E}$ can be the set of the solution pairs of $\mathcal{H}_{cl}$ originating in a given set, or having a time domain fulfilling some *dwell-time* conditions or satisfying some persistence of excitation condition.

In the following sections we present examples of identifiers that fulfill the identifier requirement: in sections 4.2 and 4.3 we address the case of linear parametrizations with (4.8) that is a least-squares functional by means of a discrete and continuous-time identifier respectively; in Section 4.4 we propose a systematic procedure to fit mini-batch algorithms for nonlinear parametrizations in the proposed framework; finally, in Section 4.5 we derive an *universal approximator* using cascades of least-squares that leverage on the wavelets theory and on the (bi)orthogonal multiresolution decomposition of the prediction errors.

4.2 Discrete-Time Least-Squares Identifiers

In this section we consider a class of discrete-time *regularized weighted least-squares identifiers*. For simplicity, we consider here a scalar case (i.e. we let in (4.2) $\beta^\star(w) \in \mathbb{R}$), however we observe that a multivariable identifier can be obtained straightforwardly either as the composition of p single-variable identifiers or, as in Sections 6.4 and 7.3, by properly augmenting the dimension of the regressor. We consider the class of identification problems consisting of an exosystem of the generic form (4.1), (4.2), a prediction model $\Phi(\theta, \cdot)$ that is *linearly parametrized* in θ and a cost functional of the form (4.8) that weights the squares of the historical prediction errors obtained during jumps. More precisely, for some $d \in \mathbb{N}$, we let

$$\Phi(\theta, \cdot) = \theta^T \sigma(\cdot), \tag{4.10}$$

with $\sigma : \mathcal{A} \to \mathbb{R}^d$ a user-defined locally Lipschitz function that plays the role of a *regressor vector*. The cost functional is obtained by letting in (4.8): $\mathrm{c}_{s,t} = 0$ for all

$s, t \in \mathbb{R}$, $\omega(\theta) = \theta^T \Omega \theta$, with $\Omega \in \mathbb{R}^{d\times d}$ a positive semidefinite matrix, and

$$\mathrm{d}_{i,j}(x) := \mu^{j-1-i}|x|^2,$$

with $\mu \in (0,1)$ that plays the role of a *forgetting factor*. With such choices the cost functional (4.8) reads as

$$J_w(\theta)(t,j) = \sum_{i=0}^{j-1} \mu^{j-1-i} \left|\beta^\star(w(t^i,i)) - \theta^T \sigma(\alpha^\star(w(t^i,i)))\right|^2 + \theta^T \Omega \theta. \tag{4.11}$$

The degrees of freedom left to the user are: the regression order d, the regressor vector σ, the forgetting factor μ and the regularization matrix Ω. These parameters can be chosen by the user to characterize the particular desired instance of the least square problem. We define a least-squares identifier as a system of the form (4.4), (4.5), with state space $\mathcal{Z} := \mathbb{R}^{d\times d} \times \mathbb{R}^d$ and with state z partitioned as $z := (R, \zeta)$ with $R \in \mathbb{R}^{d\times d}$ and fulfilling the equations

$$\mathcal{H}_{ls} : \left\{ \begin{array}{rcl} \dot{R} &=& 0 \\ \dot{\zeta} &=& 0 \end{array} \right. \quad \left\{ \begin{array}{rcl} R^+ &=& \mu R + \sigma(\alpha)\sigma(\alpha)^T \\ \zeta^+ &=& \mu\zeta + \sigma(\alpha)\beta \end{array} \right. \tag{4.12}$$

with flow and jump set given by $\mathcal{Z} \times \mathcal{A} \times \mathcal{B}$, and with output

$$\theta = (R + \Omega)^\dagger \zeta, \tag{4.13}$$

where $\cdot^\dagger$ denotes the Moore-Penrose pseudoinverse on $\mathbb{R}^{d\times d}$. We endow $\mathcal{Z}$ with the norm $|z| := |R| + |\zeta|$. For ease of notation, with $\delta = (\delta_\alpha, \delta_\beta) \in \mathcal{A} \times \mathcal{B}$, we let

$$\begin{aligned} \Sigma(w,\delta) &:= \sigma(\alpha^\star(w) + \delta_\alpha)\sigma(\alpha^\star(w) + \delta_\alpha)^T \\ \gamma(w,\delta) &:= \sigma(\alpha^\star(w) + \delta_\alpha)(\beta^\star(w) + \delta_\beta). \end{aligned} \tag{4.14}$$

Then the overall system (4.1), (4.2), (4.12), (4.13) reads as

$$\mathcal{H}_{cl}^{ls} : \begin{array}{cc} \left\{ \begin{array}{rcl} \dot{w} &\in& S(w) \\ \dot{R} &=& 0 \\ \dot{\zeta} &=& 0 \end{array} \right. & \left\{ \begin{array}{rcl} w^+ &\in& R(w) \\ R^+ &=& \mu R + \Sigma(w,\delta) \\ \zeta^+ &=& \mu\zeta + \gamma(w,\delta) \end{array} \right. \\ (w,z,\delta) \in \mathcal{C} & (w,z,\delta) \in \mathcal{D} \end{array} \tag{4.15}$$

where $\mathcal{C} := \mathcal{C}_w \times \mathcal{Z} \times \mathcal{A} \times \mathcal{B}$ and $\mathcal{D} := \mathcal{D}_w \times \mathcal{Z} \times \mathcal{A} \times \mathcal{B}$.

As the pseudoinverse map is not in general continuous everywhere, the point 3) of the identifier requirement might not hold. As a matter of fact, if the state (R, ζ) converges to a point, say $(R^\star, \zeta^\star)$, there is no guarantee that θ converges to $\theta^\star := (R^\star + \Omega)^\dagger \zeta^\star$. In order to overcome this issue, we define the following *persistence of excitation* condition, where for a square matrix M, we let $\operatorname{msv}(M)$ be the minimum non-zero singular value of M.

Definition 4.2. *With $J \in \mathbb{N}$ and $\epsilon > 0$, a complete hybrid input $\alpha : \operatorname{dom}\alpha \to \mathcal{A}$ is said to be (J, ϵ)*-persistently exciting *if*

$$\operatorname{msv}\left(\sum_{i=0}^{j-1} \mu^{j-1-i} \Sigma(\alpha(t^i, i), 0) + \Omega\right) \geq \epsilon, \qquad \forall j \geq J.$$

In the following we will often abbreviate "(J, ϵ)-persistently exciting" by "(J, ϵ)-PE" and for any $(J, \epsilon) \in \mathbb{N} \times \mathbb{R}_+^*$ we let $\mathcal{E}_{PE}^{J,\epsilon} \subset \mathcal{S}_{\mathcal{H}_{cl}^{ls}}$ be the set of solution pairs (w, z, δ) of (4.15) for which $\alpha^\star(w)$ and $\alpha^\star(w) + \delta$ are (J, ϵ)-PE. We observe that, whenever $J' \leq J$ and $\epsilon' \geq \epsilon$, then $\mathcal{E}_{PE}^{J',\epsilon'} \subset \mathcal{E}_{PE}^{J,\epsilon}$.

Remark 4.1. It is worth noting that, even if the PE property is a property of a hybrid arc α, it is strongly influenced by the parameters μ and Ω. As a matter of fact, if Ω is chosen positive definite, then every hybrid input α is $(0, \nu)$-PE, with $\nu > 0$ the real part of the minimum eigenvalue of Ω. In fact, by construction, the sum

$$S(j) := \sum_{i=0}^{j-1} \mu^{j-1-i} \Sigma(\alpha(t^i, i), 0)$$

is positive semi-definite. Hence, if λ is an eigenvalue of $S(j) + \Omega$ and v is a corresponding eigenvector, then $(S(j) + \Omega)v = \lambda v$, that implies $S(j)v \leq (\lambda - \nu)v$. As $S(j)$ is positive semi-definite, we get $0 \leq v^T S(j) v \leq (\lambda - \nu) v^T v$, which implies $\lambda \geq \nu$, and this suffices to conclude $\operatorname{msv}(S(j) + \Omega) \geq \nu$. △

Remark 4.2. The set $\mathcal{E}_{PE}^{J,\epsilon}$ is defined as the set of all the solution pairs of (4.15) for which both the ideal input $\alpha^\star(w)$ and the perturbed input $\alpha(w) = \alpha^\star(w) + \delta_\alpha$ are (J, ϵ)-PE. Nevertheless, by continuity of the function msv in Definition 4.2, it is possible to conclude that if $\alpha^\star(w)$ is (J, ϵ)-PE and the disturbance δ_α is small enough at the instants preceding a jump, then also the perturbed input α is (J', ϵ')-PE, for some $J' \geq J$ and $\epsilon' \leq \epsilon$. This in turn makes the PE condition a property of the exosystem. We also observe that, when applied to the adaptive high-gain observer in Proposition 5.1, the disturbance δ_α equals $x - \hat{x}_{[1,n]}$ and, hence, standard high-gain arguments can be used to show that for g and λ sufficiently large, at jump-times δ_α can be made small enough to make sure that if x is (J,ϵ)-PE, then so is also $\hat{x}_{[1,n]}$, thus reducing the PE condition to a property of the plant. $\triangle$

Remark 4.3. We also observe that the PE condition of the input $\alpha^\star + \delta$ (and thus of $\alpha^\star$ if δ is sufficiently small at jumps) can be checked online by looking at the matrix $R(t,j) + \Omega$. As a matter of fact, it can be shown by the same arguments of the forthcoming Proposition 4.1 that $R + \Omega$ converges exponentially to the argument of $\operatorname{msv}(\cdot)$ in Definition 4.2. Therefore, by continuity of msv, it follows that $\alpha^\star + \delta_\alpha$ is (J,ϵ)-PE if and only if there exists $\epsilon' \in \mathbb{R}_+^*$ such that $\operatorname{msv}(R(t_j,j) + \Omega) \geq \epsilon'$ for all $(t,j) \in \Gamma(R)|_{\geq J}$. $\triangle$

With $(r,N) \in (\mathbb{R}_+)^2$ we define the restriction $\mathcal{E}_{radt}^{r,N}$ on $\mathcal{H}_{cl}^{ls}$ as the set of solution pairs in $\mathcal{S}_{\mathcal{H}^{ls}}$ fulfilling the reverse average dwell-time condition (A.3) with parameters (r,N). Then the following result holds.

Proposition 4.1. *With $\mathcal{H}_w$ given by* (4.1), (4.2) *and J_w by* (4.11)*, suppose that Assumption 4.1 holds and pick arbitrary $(r,N) \in (\mathbb{R}_+^*)^2$ and $(J,\epsilon) \in \mathbb{N} \times \mathbb{R}_+^*$. Then the identifier* (4.12), (4.13) *fulfils the identifier requirement relatively to the identification problem $(\mathcal{H}_w, J_w)$ with restriction $\mathcal{E}_{PE}^{J,\epsilon} \cap \mathcal{E}_{radt}^{r,N}$ and with any κ_1, with $\underline{\sigma}, \bar{\sigma}$ and κ_2 linear and ρ locally Lipschitz.*

We underline that the existence of the steady state $z^\star$ such that point 1 of the identifier requirement holds is guaranteed by construction. Point 2 requires instead the time domain of the solutions to satisfy (A.3) and the persistence of excitation condition is only needed to obtain the regularity requirement.

Proof of Proposition 4.1. Well-definiteness in the sense of Definition 4.1 is clear. Pick a solution $(w,z) \in \mathcal{S}_{\mathcal{H}_{cl}^{ls}}$. Then, at each $(t,j) \in \operatorname{dom}(w,z)$, the set (4.9) of minimizers of (4.11) is given by the set of θ that annihilate its gradient, i.e.

$$\vartheta_w^\circ(t,j) = \left\{\theta \in \mathbb{R}^d \,:\, (R^\star(t,j) + \Omega)\theta = \zeta^\star(t,j)\right\} \tag{4.16}$$

where

$$R^\star(t,j) := \sum_{i=0}^{j-1} \mu^{j-1-i}\Sigma(w(t^i,i),0)$$
$$\zeta^\star(t,j) := \sum_{i=0}^{j-1} \mu^{j-1-i}\gamma(w(t^i,i),0).$$

Let $z^\star : \operatorname{dom} z^\star = \operatorname{dom}(w,z) \to \mathcal{Z}$ be defined as $\zeta^\star(t,j) := (R^\star(t,j), \zeta^\star(t,j))$. By construction $R^\star$ and $\zeta^\star$ satisfy $\dot{R}^\star(t,j) = 0$ and $\dot{\zeta}^\star(t,j) = 0$ for all $(t,j) \in \mathcal{I}(w,z)$ and

$$\begin{array}{rcl} R^\star(t,j+1) & = & \mu R^\star(t,j) + \Sigma(w(t,j),0) \\ \zeta^\star(t,j+1) & = & \mu\zeta^\star(t,j) + \gamma(w(t,j),0) \end{array} \tag{4.17}$$

for all $(t,j) \in \Gamma(w,z)$. Let $\tilde{z} = (\tilde{R}, \tilde{z}) := z - z^\star$, then, in view of (4.15), (4.17), for all $(t,j) \in \mathcal{I}(\tilde{z})$, $\tilde{z}$ satisfies

$$\begin{aligned} |\tilde{z}(t,j+1)| &= |\tilde{R}(t,j+1)| + |\tilde{z}(t,j+1)| \\ &\le \mu|\tilde{z}(t,j)| + |\Sigma(w(t,j),\delta(t,j)) - \Sigma(w(t,j),0)| \\ &\quad + |\gamma(w(t,j),\delta) - \gamma(w(t,j),0)| \end{aligned}$$

As Σ and σ are locally Lipschitz and for all $w \in \mathcal{W}$ the quantities $\Sigma(w,\delta) - \Sigma(w,0)$ and $\gamma(w,\delta) - \gamma(w,0)$ vanish for $\delta = 0$, there exists a locally Lipschitz function $\rho \in \mathcal{K}$ such that

$$|\tilde{z}(t,j+1)| \le \mu|\tilde{z}(t,j)| + \rho(|\delta(t,j)|).$$

Pick $(r,N) \in (\mathbb{R}_+)^2$ arbitrary an define the virtual clock system

$$\left\{\begin{array}{rcll} \dot{\tau} & = & 1 & \tau \in [0,N] \\ \tau^+ & = & \max\{0, \tau - r\} & \tau \in [0,N] \end{array}\right. . \tag{4.18}$$

Then in view of (Cai et al., 2008, Prop. 1.2) a solution $(w,z,\delta) \in \mathcal{S}_{\mathcal{H}_{cl}^{ls}}$ is in $\mathcal{E}_{radt}^{r,N}$ if and only if (w,z,δ) is a solution pair of the extended system (4.15), (4.18). Pick $(w,z,\delta) \in \mathcal{E}_{radt}^{r,N}$ and, with

$$\omega \in \big(0, \log \mu^{-1}\big) \qquad\qquad k \in \left(0, \frac{\log \mu^{-1} - \omega}{r}\right)$$

consider the function

$$V(\tilde{z},\tau) := \exp(-k\tau)|\tilde{z}|.$$

then point 2.a) of the requirement holds with $\underline{\sigma} = \exp(-kN)$ and $\bar{\sigma} = 1$. For all $(t,j) \in \mathcal{I}(w,z)$, we have

$$D^+V(\tilde{z},\tau) = -kV(\tilde{z},\tau),$$

while for all $(t,j) \in \Gamma(w,z)$, noting that $\tau^+ \geq \tau - r$, we obtain

$$V(\tilde{z},\tau)^+ \leq \mu e^{kr} V(\tilde{a},\tau) + \rho(|\delta|).$$

as by definition we have

$$\mu e^{kr} \leq \mu e^{\log \mu^{-1} - \omega} \leq e^{-\omega},$$

then points 2.b) and 2.c) of the identifier requirement follow with $\nu := \min\{k,\omega\}$ and ρ defined as above.

Point 1) of the identifier requirements follows from the fact that, in view of (4.16) $\theta^\star(t,j) = (R^\star(t,j)+\Omega)^\dagger \zeta^\star(t,j)$ is in $\vartheta_w^\circ(t,j)$. Point 3.a) follows directly by the fact that the state is constant during flows and, hence, $\dot{h}(z) = 0$ in $\mathcal{I}(z)$. It thus remains to prove point 3.b). We can write (we omit the time dependency)

$$\begin{aligned}|\theta - \theta^\star| &\leq |(R+\Omega)^\dagger \zeta - (R^\star+\Omega)^\dagger \zeta^\star| \\ &\leq |(R+\Omega)^\dagger - (R^\star+\Omega)^\dagger||\zeta^\star| + |(R^\star+\Omega)^\dagger||\zeta - \zeta^\star|\end{aligned}$$

In view of (Campbell and Meyer, 2009, Thm. 10.4.5), we have

$$|(R+\Omega)^\dagger - (R^\star+\Omega)^\dagger| \leq 3\max\{|(R+\Omega)^\dagger|^2, |(R^\star+\Omega)^\dagger|^2\}|R - R^\star|. \tag{4.19}$$

If $(w,z,\delta) \in \mathcal{E}_{PE}$, then for some $(J,\epsilon) \in \mathbb{N} \times \mathbb{R}_+^*$ we have $\mathrm{msv}(R+\Omega) > \epsilon$ and $\mathrm{msv}(R^\star+\Omega) \geq \epsilon$, and thus (4.19) implies

$$|\theta - \theta^\star| \leq \big(3\max\{1/\epsilon^2, 1/\epsilon'^2\}|\zeta^\star| + 1/\epsilon\big)|z - z^\star|,$$

for all $(t,j) \in \mathrm{dom}(w,z)|_{\geq T}$ with $T := J + t_J$. Thus, noting that

$$|\zeta^\star| \leq \frac{1}{1-\mu} \sup_{w\in W} |\gamma(w,0)|,$$

we obtain point 3.b) of the identifier requirement with κ_2 linear, and this concludes the proof. ■

4.3 Continuous-Time Least-Squares Identifiers

In this section we present a continuous-time version of the discrete-time identifier developed in the previous Section 4.2. Again we stick to the scalar case, i.e. we assume $\beta \in \mathbb{R}$, and we recall that a multivariable identifier can be obtained straightforwardly either as the composition of p single-variable identifiers or, as in Sections 6.4 and 7.3, by properly augmenting the dimension of the regressor. We assume to have fixed the order $d \in \mathbb{N}$ of the model and a we pick the same model structure of (4.10), i.e.

$$\phi(\cdot, \theta) = \theta^T \sigma(\cdot)\,,$$

with $\sigma : \mathcal{A} \to \mathbb{R}$ a known locally Lipschitz function. We consider a continuous-time least-squares functional, obtained by letting in (4.8): $d_{i,j} = 0$ for all $i, j \in \mathbb{N}$, $\omega(\theta) = \theta^T \Omega \theta$, for some positive semi-definite $\Omega \in \mathbb{R}^{d\times d}$, and

$$c_{s,t}(\cdot) := \exp(-\lambda(t-s))|\cdot|^2,$$

for some $\lambda > 0$ that plays the role of a forgetting factor (i.e. λ is the continuous-time analogous of μ of Section 4.2). We thus obtain the cost functional

$$J_w(\theta)(t,j) = \lambda \int_0^t e^{-\lambda(t-s)} \big|\beta^\star(w(t_i,i)) - \theta^T \sigma(\alpha^\star(w(t_i,i)))\big|^2 ds + \theta^T \Omega \theta. \tag{4.20}$$

The continuous-time analogous of (4.12)-(4.13) is obtained by letting in (4.4)-(4.5), $\mathcal{Z} := \mathbb{R}^{d\times d} \times \mathbb{R}^d$, by partitioning the state as $z = (R, \zeta)$, with $R \in \mathbb{R}^{d\times d}$ and $\zeta \in \mathbb{R}^d$, and by letting

$$\begin{cases} \dot{R} &= -\lambda R + \lambda \sigma(\alpha)\sigma(\alpha)^T \\ \dot{\zeta} &= -\lambda \zeta + \lambda \sigma(\alpha)\beta \end{cases} \qquad \begin{cases} R^+ &= 0 \\ \zeta^+ &= 0 \end{cases} \tag{4.21}$$

with output

$$\theta = (R + \Omega)^\dagger \zeta \tag{4.22}$$

where again $\cdot^\dagger$ denotes the Moore-Penrose pseudoinverse. We endow $\mathcal{Z}$ with the norm $|z| = |R| + |\zeta|$. Then, with $\Sigma(\cdot)$ and $\gamma(\cdot)$ defined as in (4.14), the overall system (4.1), (4.2), (4.21), (4.22) reads as

$$\mathcal{H}_{cl}^{ls}: \begin{cases} \dot{w} & \in & S(w) \\ \dot{R} & = & -\lambda R + \lambda\Sigma(w,\delta) \\ \dot{\zeta} & = & -\lambda\zeta + \lambda\gamma(w,\delta) \end{cases} \qquad \begin{cases} w^+ & \in & R(w) \\ R^+ & = & 0 \\ \zeta^+ & = & 0 \end{cases} \tag{4.23}$$
$$(w,z,\delta) \in \mathcal{C} \qquad\qquad (w,z,\delta) \in \mathcal{D}$$

where $\mathcal{C} := \mathcal{C}_w \times \mathcal{Z} \times \mathcal{A} \times \mathcal{B}$ and $\mathcal{D} := \mathcal{D}_w \times \mathcal{Z} \times \mathcal{A} \times \mathcal{B}$.

We introduce a persistence of excitation condition that is the continuous-time analogous of Definition 4.2:

Definition 4.3. *With $(T,\epsilon) \in \mathbb{R}_+ \times \mathbb{R}_+^*$, a complete hybrid arc $\alpha : \operatorname{dom}\alpha \to \mathcal{A}$ is said to be (T,ϵ)-persistently exciting if*

$$\operatorname{msv}\left(\int_0^t e^{-\lambda(t-s)}\Sigma(\alpha(s,j_s),0)ds + \Omega\right) \geq \epsilon, \qquad \forall t \geq T.$$

For $(T,\epsilon) \in \mathbb{R}_+ \times \mathbb{R}_+^*$, we denote by $\mathcal{E}_{PE}^{(T,\epsilon)} \subset \mathcal{S}_{\mathcal{H}_{cl}^{ls}}$ the set of solution pairs (w,z,δ) of (4.23) for which $\alpha^\star(w)$ and $\alpha^\star(w) + \delta_\alpha$ are both (T,ϵ)-PE. We also observe that, whenever $T' \leq T$ and $\epsilon' \geq \epsilon$, $\mathcal{E}_{PE}^{(T',\epsilon')} \subset \mathcal{E}_{PE}^{(T,\epsilon)}$. We also underline that the continuous-time analogous of remarks 4.1, (4.2) and (4.3) can be stated in this case.

With $(\nu,N) \in \mathbb{R}_+ \times \mathbb{N}$ we define the restriction $\mathcal{E}_{adt}^{\nu,N}$ on $\mathcal{H}_{cl}^{ls}$ as the set of solution pairs in $\mathcal{S}_{\mathcal{H}_{cl}^{ls}}$ fulfilling the reverse average dwell-time condition (A.2) with parameters (ν,N). Then the following result holds.

Proposition 4.2. *With $\mathcal{H}_w$ given by* (4.1), (4.2) *and J_w by* (4.20)*, suppose that Assumption 4.1 holds and pick arbitrary $(\nu,N) \in \mathbb{R}_+^* \times \mathbb{N}^*$ and $(T,\epsilon) \in \mathbb{R}_+ \times \mathbb{R}_+^*$. Then the identifier* (4.21), (4.22) *fulfills the identifier requirement relatively to the identification problem $(\mathcal{H}_w, J_w)$ with restriction $\mathcal{E}_{PE}^{T,\epsilon} \cap \mathcal{E}_{adt}^{\nu,N}$ and with any κ_2, with $\underline{\sigma}, \bar{\sigma}$ and κ_1 linear and ρ locally Lipschitz.*

We underline that, also in this case, the existence of the steady state $z^\star$ such that point 1 of the identifier requirement holds is guaranteed by construction. Point 2 requires instead the time domain of the solutions to satisfy (A.2) and the persistence of excitation condition is only needed to obtain the regularity requirement.

Proof of Proposition 4.2. Well-definiteness in the sense of Definition 4.1 is clear. Pick a solution $(w, z) \in \mathcal{S}_{\mathcal{H}_{cl}^{ls}}$. Then, at each $(t, j) \in \operatorname{dom}(w, z)$, the set (4.9) of minimizers of (4.20) is given by the set of θ that annihilate its gradient, i.e.

$$\vartheta_w^\circ(t, j) = \left\{\theta \in \mathbb{R}^d \,:\, (R^\star(t, j) + \Omega)\theta = \zeta^\star(t, j)\right\} \tag{4.24}$$

where

$$\begin{aligned} R^\star(t, j) &:= \lambda \int_0^t e^{-\lambda(t-s)} \Sigma(w(s, j_s), 0) ds \\ \zeta^\star(t, j) &:= \lambda \int_0^t e^{-\lambda(t-s)} \gamma(w(s, j_s), 0) ds. \end{aligned}$$

Let $z^\star : \operatorname{dom} z^\star = \operatorname{dom}(w, z) \to \mathcal{Z}$ be defined as $\zeta^\star(t, j) := (R^\star(t, j), \zeta^\star(t, j))$. By construction $R^\star$ and $\zeta^\star$ satisfy $R^\star(t, j+1) = R^\star(t, j)$ and $\zeta^\star(t, j+1) = \zeta^\star(t, j)$ for all $(t, j) \in \Gamma(w, z)$ and

$$\begin{aligned} \dot{R}^\star(t, j) &= \lambda R^\star(t, j) + \Sigma(w(t, j), 0) \\ \dot{\zeta}^\star(t, j) &= \mu \zeta^\star(t, j) + \gamma(w(t, j), 0) \end{aligned} \tag{4.25}$$

for all $(t, j) \in \mathcal{I}(w, z)$. Let $\tilde{z} = (\tilde{R}, \tilde{z}) := z - z^\star$, then, in view of (4.23), (4.25), for all $(t, j) \in \mathcal{I}(\tilde{z})$, $\tilde{z}$ satisfies

$$\begin{aligned} D^+|\tilde{z}(t, j)| &= D^+\left(|\tilde{R}(t, j)| + |\tilde{z}(t, j)|\right) \\ &\leq -\lambda|\tilde{z}(t, j)| + |\Sigma(w(t, j), \delta(t, j)) - \Sigma(w(t, j), 0)| \\ &\quad + |\gamma(w(t, j), \delta(t, j)) - \gamma(w(t, j), 0)| \end{aligned}$$

As Σ and σ are locally Lipschitz and for all $w \in \mathcal{W}$ the quantities $\Sigma(w, \delta) - \Sigma(w, 0)$ and $\gamma(w, \delta) - \gamma(w, 0)$ vanish for $\delta = 0$, there exists a locally Lipschitz function $\rho_0 \in \mathcal{K}$ such that

$$D^+|\tilde{z}(t, j)| \leq -\lambda|\tilde{z}(t, j)| + \rho_0(|\delta(t, j)|).$$

Pick $(\nu, N) \in \mathbb{R}_+ \times \mathbb{N}$ arbitrary an define the virtual clock system

$$\begin{cases} \dot{\tau} & \in & F_\tau(\tau) & \tau \in [0, N] \\ \tau^+ & = & \tau - 1 & \tau \in [1, N] \end{cases} . \tag{4.26}$$

with

$$F_\tau(\tau) := \begin{cases} \nu & \tau \in [0, N_0) \\ [0, \nu] & \tau = N_0 \end{cases}$$

Then in view of (Cai et al., 2008, Prop. 1.1) a solution $(w, z, \delta) \in \mathcal{S}_{\mathcal{H}_{cl}^{ls}}$ is in $\mathcal{E}_{adt}^{\nu,N}$ if and only if (w, z, δ) is a solution pair of the extended system (4.23), (4.26). Pick $(w, z, \delta) \in \mathcal{E}_{adt}^{\nu,N}$ and, with

$$\omega \in (0, \lambda) \qquad\qquad k \in \left(0, \frac{\lambda - \omega}{\nu}\right)$$

consider the function

$$V(\tilde{z}, \tau) := \exp(k\tau)|\tilde{z}|.$$

then point 2.a) of the requirement holds with $\underline{\sigma} = 1$ and $\bar{\sigma} = e^{kN}$. Moreover, for all $(t, j) \in \mathcal{I}(w, z)$, we have

$$\begin{aligned} D^+V(\tilde{z}, \tau) &\leq (k\nu - \lambda)V(\tilde{z}, \tau) + e^{kN}\rho_0(|\delta|) \\ &\leq -\omega V(\tilde{z}, \tau) + e^{kN}\rho_0(|\delta(t, j)|), \end{aligned}$$

while, for all $(t, j) \in \Gamma(w, z)$, noting that $\tau^+ = \tau - 1$, we obtain

$$V(\tilde{z}, \tau)^+ = e^{-k}V(\tilde{a}, \tau).$$

and hence it follows that there exists a locally Lipschitz $\rho \in \mathcal{K}$ such that then points 2.b) and 2.c) of the identifier requirement follow with $\nu := \min\{k, \omega\}$.

Point 1) of the identifier requirements follows from the fact that, in view of (4.24) $\theta^\star(t, j) = (R^\star(t, j) + \Omega)^\dagger \zeta^\star(t, j)$ is in $\vartheta_w^\circ(t, j)$. Point 3.b) follows directly by the fact that the state is constant during jumps. It thus remains to prove point 3.a) that, however, follows from the same arguments used in the proof of Proposition 4.1. ■

4.4 Mini-Batch Identifiers for Nonlinear Parametrizations

In this section we consider a class of discrete-time identification algorithms that work on a "moving window" of prescribed size. More precisely, with $M \in \mathbb{N}$ an arbitrary number denoting the size of the moving window and with φ a hybrid arc taking values in an Euclidean space $\mathcal{X}$, we define the window operator ω_M on φ as

$$\omega_M(\varphi)(t,j) := \begin{pmatrix} \varphi(t^{j-M}, j-M) \\ \varphi(t^{j-M+1}, j-M+1) \\ \dots \\ \varphi(t^{j-2}, j-2) \\ \varphi(t^{j-1}, j-1) \end{pmatrix} \in \mathcal{X}^M.$$

With reference to the framework of Section 4.1, suppose that for some arbitrary $m, p \in \mathbb{N}$, with $p \leq m$, $\mathcal{A} = \mathbb{R}^m$ and $\mathcal{B} = \mathbb{R}^p$, and we assume that we are given:

- An exosystem $\mathcal{H}_w$ of the form (4.1)-(4.2) that generates the data $\alpha^\star(w)$ and $\beta^\star(w)$.
- A model order $d \in \mathbb{N}$ and a model set $\mathcal{M}$ containing prediction models of the form $\Phi(\theta, \cdot)$ with θ that ranges in $\mathbb{R}^d$.
- A cost function J_w of the form (4.8) obtained with $\mathrm{c}_{s,t} = 0$ for all $s, t \in \mathbb{R}$ and with $\mathrm{d}_{i,j} = 0$ for all $i, j \in \mathbb{N}$ such that $i < j - M$, i.e.

$$J_w(\theta)(t,j) := \sum_{i=\max\{0,j-M\}}^{j-1} \mathrm{d}_{i,j}\big(\varepsilon(w(t^i,i),\theta)\big), \tag{4.27}$$

  where we recall that (see (4.7))

$$\varepsilon(w,\theta) := \beta^\star(w) - \Phi(\theta, \alpha^\star(w)). \tag{4.28}$$

- A continuous function $\mathcal{G} : \mathcal{A}^M \times \mathcal{B}^M \to \mathbb{R}^d$ that satisfies

$$\mathcal{G}\big(\omega_M(\alpha^\star(w))(t,j), \omega_M(\beta^\star(w))(t,j)\big) \in \arg\inf_{\theta \in \mathbb{R}^d} J_w(\theta)(t,j)\,. \tag{4.29}$$

Remark 4.4. The class of algorithms that we consider here include, for instance, algorithms that are originally defined in a batch (or off-line) context and that can be used online by running them at each new sample over the fixed amount of data in the moving window. Many algorithms for nonlinear parametrizations are of this kind; for instance the *gradient descent* techniques can be seen as mini-batch algorithms of this kind working on windows of dimension $M = 1$. $\triangle$

Remark 4.5. The condition (4.29) is motivated by the fact that in view of (4.28), the cost functional J_w given in (4.27) at a given $(t, j) \in \operatorname{dom} w$ is only a function of the windowed quantities $\omega_M(\alpha^\star(w))(t, j)$ and $\omega_M(\beta^\star(w))(t, j)$. $\triangle$

We construct an identifier of the form (4.4)-(4.5), fitting in the framework of Section 4.1, by letting $\mathcal{Z} := \mathcal{A}^M \times \mathcal{B}^M$, by decomposing the state as $z := (\chi, \xi)$, with $\chi := (\chi_1, \dots, \chi_M) \in \mathcal{A}^M$, $\chi_i \in \mathcal{A}$, and $\xi := (\xi_1, \dots, \xi_M) \in \mathcal{B}^M$, $\xi_i \in \mathcal{B}$, and by choosing F and G such that the following equations holds:

$$\begin{cases} \dot{\chi} &= 0 \\ \dot{\xi} &= 0 \end{cases} \quad \begin{cases} \chi_i^+ &= \chi_{i+1}, & i = 1, \dots, M-1 \\ \chi_M^+ &= \alpha \\ \xi_i^+ &= \xi_{i+1}, & i = 1, \dots, M-1 \\ \xi_M^+ &= \beta \end{cases} \tag{4.30}$$

with flow and jump sets given by $\mathcal{Z} \times \mathcal{A} \times \mathcal{B}$ and with output

$$\theta = \mathcal{G}(\chi, \xi). \tag{4.31}$$

We interconnect the identifier (4.30)-(4.31) with an exosystem of the form (4.1)-(4.2), by letting $\alpha := \alpha^\star(w) + \delta_\alpha$ and $\beta := \beta^\star(w) + \delta_\beta$, with $\delta = (\delta_\alpha, \delta_\beta) \in \mathcal{A} \times \mathcal{B}$ a hybrid input, obtaining the composite system:

$$
\mathcal{H}_{cl}: \begin{cases} \begin{cases} \dot{w} \in S(w) \\ \dot{\chi} = 0 \\ \dot{\xi} = 0 \end{cases} & (w,z,\delta) \in \mathcal{C}_w \times \mathcal{Z} \times \mathcal{A} \times \mathcal{B} \\ \begin{cases} w^+ \in R(w) \\ \chi_i^+ = \chi_{i+1}, \quad i = 1,\dots,M-1 \\ \chi_M^+ = \alpha \\ \xi_i^+ = \xi_{i+1}, \quad i = 1,\dots,M-1 \\ \xi_M^+ = \beta \end{cases} & (w,z,\delta) \in \mathcal{C}_w \times \mathcal{Z} \times \mathcal{A} \times \mathcal{B} \end{cases} \tag{4.32}
$$

For $(r, N) \in (\mathbb{R}_+)^2$, we let $\mathcal{E}_{radt}^{r,N}$ be the set for solution pairs of (4.32) that satisfy the reverse average dwell-time condition (A.3). Then the following result holds:

Proposition 4.3. *With $\mathcal{H}_w$ given by (4.1)-(4.2) and J_w by (4.27), suppose that Assumption 4.1 holds and pick arbitrarily $(r, N) \in (\mathbb{R}_+^*)^2$. Then there exist $\kappa_2 \in \mathcal{K}$ and linear $\underline{\sigma}, \bar{\sigma}, \rho \in \mathcal{K}_\infty$ such that identifier (4.30)-(4.31) satisfies the identifier requirement relative to the identification problem $(\mathcal{H}_w, J_w)$ with restriction $\mathcal{E}_{radt}^{r,N}$ and with any $\kappa_1 \in \mathcal{K}$.*

Proof. The jump equation of the identifier (4.30) can be written in the compact form as

$$
z^+ = Az + B \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \tag{4.33}
$$

where A and B are suitably defined. Consider the following lemma, that is proved at the end of this proof.

Lemma 4.1. *Let $A \in \mathbb{R}^n$, $n \in \mathbb{N}$, be Schur. Then for any $Q \in \mathbb{R}^n$ symmetric and positive definite there exists $\mu \in (0, 1)$ and an unique $P \in \mathbb{R}^n$ symmetric and positive definite such that*

$$
A^T PA - \mu P = -Q. \tag{4.34}
$$

Pick $Q = Q^T > 0$ arbitrarily and let $\mu \in (0, 1)$ and $P = P^T > 0$ be produced by Lemma 4.1, with A the same as in (4.33). Then we endow $\mathcal{Z}$ with the norm

$$
|z|_P := \sqrt{z^T P z}.
$$

For any $w \in \mathcal{S}_{\mathcal{H}^w}$ and let $z^\star$ be given by $z^\star := (\chi^\star, \xi^\star)$ with

$$\chi^\star(t,j) := \omega_M(\alpha^\star(w))(t,j), \qquad\qquad \xi^\star(t,j) := \omega_M(\beta^\star(w))(t,j).$$

Point 1 of the identifier requirement then follows directly by (4.29). Moreover, we observer that $z^\star$ satisfies

$$z^{\star+} = Az^\star + B\begin{pmatrix}\alpha^\star(w)\\ \beta^\star(w)\end{pmatrix}$$

with A the same as in (4.32). Pick $(w, z, \delta) \in \mathcal{S}_{\mathcal{H}^{cl}}$ and, with $z^\star$ defined as above, let $\tilde{z} := z - z^\star$. Then as $P > 0$, there exist $\underline{c}$ and $\bar{c}$ positive such that

$$\underline{c}|\tilde{z}| \leq |\tilde{z}|_P \leq \bar{c}|\tilde{z}|. \tag{4.35}$$

Moreover, for all $(t,j) \in \Gamma(\tilde{z})$ we have $\tilde{z}^+ = A\tilde{z} + B\delta$, so as, in view of (4.34), we obtain

$$\begin{aligned}(|\tilde{z}|_P^+)^2 &= \tilde{z}^T A^T P A \tilde{z} + \delta^T B P B \delta + 2\tilde{z}^T A^T P B \delta\\ &\leq \mu|\tilde{z}|_P^2 + |B|^2|P||\delta|^2 + |2\tilde{z}^T A^T P B\delta|.\end{aligned}$$

With $\gamma \in (\mu, 1)$, pick

$$\epsilon := \underline{c}^2(\gamma - \mu).$$

Then in view of (4.35), the Young's inequality gives

$$|2\tilde{z}^T A^T P B\delta| \leq \epsilon|\tilde{z}|^2 + \frac{|A^T P B|^2}{\epsilon}|\delta|^2 \leq (\gamma-\mu)|\tilde{z}|_P^2 + \frac{|A^T P B|}{\underline{c}^2(\gamma-\mu)}|\delta|^2,$$

so as for some constant $c_1 > 0$, we obtain

$$|\tilde{z}|_P^+ \leq \gamma|\tilde{z}|_P + c_1|\delta|. \tag{4.36}$$

On the other hand, for $(t,j) \in \mathcal{I}(\tilde{z})$, we have

$$D^+|\tilde{z}|_P = 0,$$

and the point 2 of the requirement for the solution pairs inside $\mathcal{E}_{PE}^{r,N}$ for some fixed $(r, N) \in (\mathbb{R}_+^*)^2$ follows from the same arguments used in Proposition 4.1. Finally, the point 3.a of the identifier requirement holds for any κ_1 by the fact that $\dot{\theta} = 0$ in $\mathcal{I}(w, z, \delta)$, while point 3.b is a consequence of the continuity of $\mathcal{G}$. ■

Proof of Lemma 4.1. Let $\bar{\lambda}_A$ be the eigenvalue of A with largest modulus. As A is Schur, $|\bar{\lambda}_A| \in [0, 1)$. Pick

$$\mu \in \big(|\bar{\lambda}_A|^2, 1\big)$$

and let $F := A/\sqrt{\mu}$. Then

$$\sigma(F) = \frac{1}{\sqrt{\mu}}\sigma(A)$$

and, hence, for all $\lambda_F \in \sigma(F)$, we have

$$|\lambda| \leq |\bar{\lambda}_A|/\sqrt{\mu} < 1,$$

i.e. F is Schur.

Pick arbitrarily $Q \in \mathbb{R}^{n\times n}$ such that $Q = Q^T > 0$ and let $Q_0 := Q\mu$. As F is Schur, there exists unique $P = P^T > 0$ such that the Lyapunov equation

$$F^T PF - P = -Q_0$$

holds. Thus

$$A^T PA - \mu P = \mu\big(F^T PF - \mu P\big) = -\mu Q_0 = -Q$$

which is the claim. ■

4.5 Wavelet Identifiers for Multiresolution Identification

In this section we construct an identifier that performs a wavelet expansion of the prediction model. The identifier is composed of a cascade of an arbitrary number of least squares identifiers of the kind presented in Section 4.2. The cascade structure of the identifier reflects the multiresolution nature of the wavelet expansion: the first least squares stage captures the best representation at the starting *scale*. All the other stages encode the information corresponding to the "detail" that is missing to the precedent stage to obtain a representation of the prediction model at a finer scale. We suppose the reader to be familiar with the wavelet theory. An essential review of the main tools used in this section is reported in Appendix B. For the basic concepts on wavelets, Riesz bases and biorthogonality we remind to (Daubechies, 1992; Walnut, 2002; Strang and Nguyen, 1996; Christensen, 2008). For convenience, we consider again the single variable case (i.e. $\beta^\star(w) \in \mathcal{B} = \mathbb{R}$ in (4.2)), though we recall that an extension to multivariable cases can be obtained by the composition of multiple single variable identifiers. In the following we also let in (4.2) $\mathcal{A} = \mathbb{R}^m$, with $m \in \mathbb{N}$ arbitrary.

To support the subsequent construction, we assume the existence of a (virtual) "true" model[2] $\phi \in L_2(\mathbb{R}^m)$ relating the signals $\alpha^\star(w)$ and $\beta^\star(w)$, i.e. we suppose we can write $\beta^\star(w) = \phi(\alpha^\star(w))$. As the functions $\alpha^\star$ and $\beta^\star$ are locally bounded, Assumption 4.1 justifies restricting ϕ to the set $\mathcal{C}_c(\mathbb{R}^m)$ of compactly supported continuous functions $\mathbb{R}^m \to \mathbb{R}$ inside $L_2(\mathbb{R}^m)$. As a first step, we fix a GMRA $(\mathbf{V}_i)_i$ associated to *compactly supported* scaling function Υ and wavelet functions Ψ^h, $h = 1, \dots, 2^m - 1$ (see Appendix B for details on the notation) and we choose a "starting scale" $i_0 \in \mathbb{Z}$. In the usual interpretation, the *orthogonal projection* $\mathrm{P}_{i_0}\phi$ of ϕ onto $\mathbf{V}_{i_0}$ represents the approximation of ϕ at scale i_0. In this construction i_0 is the largest scale and, hence, $\mathrm{P}_{i_0}\phi$ represents the *coarsest* attainable approximation of ϕ. According to (B.11), we can expand $\mathrm{P}_{i_0}\phi$ as

$$\mathrm{P}_{i_0}\phi = \sum_{\mathbf{k} \in \mathbf{H}_{i_0}} a_{i_0,\mathbf{k}} \Upsilon_{i_0,\mathbf{k}} \tag{4.37}$$

for some $a_{i_0,\mathbf{k}} \in \mathbb{R}$ and where $\mathbf{H}_{i_0} \subset \mathbb{Z}^m$ is a *finite* (ϕ is compactly supported) set such that[3] $\operatorname{supp}\phi \subset \cup_{\mathbf{k}\in\mathbf{H}_{i_0}} \operatorname{supp}\Upsilon_{i_0,\mathbf{k}}$ and $\operatorname{supp}\Upsilon_{i_0,\mathbf{k}} \cap \operatorname{supp}\phi = \emptyset$ whenever $\mathbf{k} \notin \mathbf{H}_{i_0}$. The coefficients $a_{i_0,\mathbf{k}}$ are, however, not known and in general they cannot be computed as they would require the computation of the scalar product of ϕ with the dual scaling function $\tilde{\Upsilon}_{i,\mathbf{k}}$. Nevertheless, equation (4.37) is recognized to be a linear regression fitting into the least squares framework introduced in Section 4.2. As a consequence, a first least squares identifier of the form (4.12)-(4.13), denoted by z_0, can be used to find the *best* guess of the coefficients $a_{i_0,\mathbf{k}}$.

[2]We stress though that ϕ needs not to actually exists and it is only needed to guide the identifier construction.

[3]$\operatorname{supp} f$ denotes the support of a function f.

In constructing z_0, we consider the symbolic relation

$$\beta^\star = \sum_{\mathbf{k}\in\mathbf{H}_{i_0}} a_{i_0,\mathbf{k}} \Upsilon_{i_0,\mathbf{k}}(\alpha^\star),$$

that suggests to choose in (4.10) the regressor $\sigma_0 := \mathrm{col}\,(\Upsilon_{i_0,\mathbf{k}} \,:\, \mathbf{k}\in\mathbf{H}_{i_0})$ and to design the identifier's functions (to which we append the subscript 0 to distinguish from the successive stages) accordingly. Consistently with the framework of Section 4.1, the identifier z_0 works on the perturbed inputs

$$\alpha_0 := \alpha^\star + \delta_\alpha, \quad \beta_0 := \beta^\star + \delta_\beta, \tag{4.38}$$

with $\delta = (\delta_\alpha, \delta_\beta)$ an unmodeled disturbance. With $d_0 := |\mathbf{H}_{i_0}|$, we write the dynamics of z_0 in the compact form

$$\mathcal{H}_0 : \begin{cases} \dot{z}_0 &= 0 \\ z_0^+ &= F_0(z_0, \alpha_0, \beta_0) \end{cases} \tag{4.39}$$

with state space $\mathcal{Z}_0 := \mathbb{R}^{d_0\times d_0} \times \mathbb{R}^{d_0}$, flow and jump sets given by $\mathcal{Z}_0 \times \mathbb{R}^m \times \mathbb{R}$ and with output

$$\theta_0 = h_0(z_0),$$

where $F_0(z_0, \alpha_0, \beta_0) := \mu_0 z_0 + (\sigma_0(\alpha_0)\sigma_0(\alpha_0)^T,\, \sigma_0(\alpha_0)\beta_0)$ and h_0 defined according to (4.13), with $\mu_0 \in (0,1)$ and $\Omega_0 \in \mathbb{R}^{d_0\times d_0}$ suitably chosen. The parameter θ_0 produced by z_0 corresponds thus to the best estimate of the coefficients $a_{i_0,\mathbf{k}}$ in (4.37), and we can associate to z_0 the prediction $\hat{\beta}_{\mathbf{i_0}}$ of $\beta^\star$ given by

$$\hat{\beta}_{\mathbf{i_0}} := \Phi_{i_0}(\theta_0, \alpha_0), \quad \text{with} \quad \Phi_{i_0}(\theta_0, \cdot) := \theta_0^T \sigma_0(\cdot) \in \mathbf{V}_{i_0}$$

to which we refer as the *best prediction at scale* i_0. The error of the prediction $\hat{\beta}_{\mathbf{i_0}}$, called "prediction error at scale i_0", is defined as

$$\hat{\varepsilon}_{\mathbf{i_0}} := \beta - \hat{\beta}_{\mathbf{i_0}},$$

which can be seen as the best guess of the output $\varepsilon_{\mathbf{i_0}}$ of the "scale-i_0 error model"

$$\mathrm{E}_{i_0} := \phi - \mathrm{P}_{i_0}\phi. \tag{4.40}$$

To increase the resolution of the representation, we add a second least squares stage, that we call z_1, to obtain an approximation of ϕ at the finer scale $i_0 - 1$. We proceed by considering the orthogonal projection of the scale-i_0 error model E_{i_0} onto the subspace $\mathbf{W}_{i_0} = \overline{\mathrm{span}}\{\Psi^h_{i_0,\mathbf{k}} : \mathbf{k} \in \mathbb{Z}^m, h = 1, \dots, 2^{m-1}\}$, obtaining the *detail* $\mathrm{Q}_{i_0}\phi = \mathrm{P}_{i_0-1}\phi - \mathrm{P}_{i_0}\phi$ that is missing from $\mathrm{P}_{i_0}\phi$ to have the approximation $\mathrm{P}_{i_0-1}\phi$ of ϕ at the finer scale $i_0 - 1$. According to (B.11), we can write

$$\mathrm{Q}_{i_0-1}\phi = \sum_{h=1}^{2^m-1} \sum_{\mathbf{k}\in\mathbf{K}_{i_0}} b^h_{i_0,\mathbf{k}} \Psi^h_{i_0,\mathbf{k}} \tag{4.41}$$

for some $b^h_{i_0,\mathbf{k}} \in \mathbb{R}$ and with $\mathbf{K}_{i_0} \subset \mathbb{Z}^m$ a finite set such that $\operatorname{supp}\phi \subset \cup_{h=1}^{2^m-1} \cup_{\mathbf{k}\in\mathbf{K}_{i_0}} \operatorname{supp}\Psi^h_{i_0,\mathbf{k}}$ and $\operatorname{supp}\Psi^h_{i_0,\mathbf{k}} \cap \operatorname{supp}\phi = \emptyset$ for all $\mathbf{k} \notin \mathbf{K}_{i_0}$ and all $h = 1, \dots, 2^m - 1$. The design of the second least-square stage z_1 is done by looking at (4.41) as a linear regression of the form (4.10), with regressor $\sigma_1 := \mathrm{col}(\Psi^h_{\ell,\mathbf{k}} : \mathbf{k} \in \mathbf{K}_{i_0}, h = 1, \dots, 2^{m-1})$. As the model (4.41) relates the input $\alpha^\star$ to the error attained by the approximation $\mathrm{P}_{i_0}\phi$ at scale i_0, instead of (4.38) the identifier z_1 processes the inputs

$$\alpha_1 := \alpha^\star + \delta_\alpha, \qquad\qquad \beta_1 := \varepsilon_{\mathbf{i_0}} + \delta_\beta = \beta^\star - \hat{\beta}_{\mathbf{i_0}} + \delta_\beta,$$

with $\delta = (\delta_\alpha, \delta_\beta)$ the same as in (4.38) The identifier z_1 is defined on the space $\mathcal{Z}_1 := \mathbb{R}^{d_1\times d_1} \times \mathbb{R}^{d_1}$, with $d_1 := 2^{m-1}|\mathbf{K}_{i_0}|$, and its dynamics is described by

$$\mathcal{H}_1 : \begin{cases} \dot{z}_1 &= 0 \\ z_1^+ &= F_1(z_0, z_1, \alpha_1, \beta_1) \end{cases}$$

with flow and jump set given by $\mathcal{Z}_0 \times \mathcal{Z}_1 \times \mathbb{R}^m \times \mathbb{R}$, with state $z_1 \in \mathcal{Z}_1$ and output

$$\theta_1 = h_1(z_1),$$

where, according to (4.12), $F_1(z_0, z_1, \alpha_1, \beta_1) := \mu_1 z_1 + (\sigma_1(\alpha_1)\sigma_1(\alpha_1)^T, \sigma_1(\alpha_1)\beta_1)$ and h_1 is defined as in (4.13), with $\mu_1 \in (0, 1)$ and $\Omega_1 \in \mathbb{R}^{d_1\times d_1}$ suitably chosen.

With $\theta^{\mathbf{i_0-1}} := \mathrm{col}(\theta_0, \theta_1)$, we associate to the cascade of z_0 and z_1 the *scale-*$(i_0 - 1)$ *prediction model*

$$\Phi_{i_0-1}(\theta^{i_0-1}, \cdot) := \theta_0^T\sigma_0(\cdot) + \theta_1^T\sigma_1(\cdot) \in \mathbf{V}_{i_0-1}, \tag{4.42}$$

which is composed of two contributions: $\theta_0^T \sigma_0(\alpha)$ gives as output $\hat{\beta}_{\mathbf{i_0}}$, that is the best guess of the output of the scale-i_0 model $\mathrm{P}_{i_0}\phi$, while $\theta_1^T \sigma_1(\alpha)$ gives as output the best guess $\hat{\varepsilon}_{\mathbf{i_0}}$ of the error $\hat{\varepsilon}_{\mathbf{i_0}}$ attained by $\mathrm{P}_{i_0}\phi$. In other words, z_0 projects the "real model" ϕ onto $\mathbf{V}_{i_0}$, while z_1 projects the error $\phi(\cdot) - \Phi_{i_0}(\theta_0, \cdot)$ (which is itself an approximation of E_{i_0} in (4.40)) onto $\mathbf{W}_{i_0}$. As a consequence of the (bi)orthogonal properties of the subspaces $\mathbf{V}_{i_0}$ and $\mathbf{W}_{i_0}$, the prediction of (4.42), that is

$$\hat{\beta}_{\mathbf{i_0-1}} := \hat{\beta}_{\mathbf{i_0}} + \hat{\varepsilon}_{\mathbf{i_0}},$$

also represents the best guess of the output of the scale-$(i_0 - 1)$ model $\mathrm{P}_{i_0-1}\phi$.

This construction procedure generalizes to arbitrary scale $i_0 - \ell$, with $\ell \in \mathbb{N}$. Once the identifier z_ℓ has been fixed to provide the best prediction $\hat{\beta}_{\mathbf{i_0}-\ell}$ at scale $i_0 - \ell$ given by

$$\hat{\beta}_{\mathbf{i_0}-\ell} := \Phi_{i_0-\ell}(\theta^{\mathbf{i_0}-\ell}, \alpha),$$

with

$$\Phi_{i_0-\ell}(\theta^{\mathbf{i_0}-\ell}, \cdot) := \sum_{i=0}^{\ell} \theta_i^T \sigma_i(\cdot) \ \in \ \mathbf{V}_{i_0-\ell}$$

and with $\theta^{\mathbf{i_0}-\ell} := \mathrm{col}(\theta_0, \ldots, \theta_\ell)$, if more resolution is needed, a further least squares stage $z_{\ell+1}$ can be added as above: we consider the scale-$(i_0 - \ell)$ error model

$$\mathrm{E}_{i_0-\ell} := \phi - \mathrm{P}_{i_0-\ell}\phi,$$

where $\mathrm{P}_{i_0-\ell}\phi$ is approximated at best by the prediction model $\Phi_{i_0-\ell}(\theta^{\mathbf{i_0}-\ell}, \cdot)$, and we project $\mathrm{E}_{i_0-\ell}$ onto the detail space $\mathbf{W}_{i_0-\ell}$, which coincides with span of the scale-$(i_0 - \ell)$ wavelets $\Psi^h_{i_0-\ell,\mathbf{k}}$. In this way we obtain the detail $\mathrm{Q}_{i_0-\ell}\phi$ at scale $i_0 - \ell$ that is the information missing to obtain a finer approximation $\mathrm{P}_{i_0-\ell-1}\phi$ at scale $i_0 - \ell - 1$. The expansion (4.41) generalizes to scale $i_0 - \ell$ as follows

$$\mathrm{Q}_{i_0-\ell}\phi = \sum_{h=1}^{2^m-1} \sum_{\mathbf{k}\in\mathbf{K}_{i_0-\ell}} b^h_{i_0-\ell,\mathbf{k}} \Psi^h_{i_0-\ell,\mathbf{k}}, \tag{4.43}$$

with $\mathbf{K}_{i_0-\ell} \subset \mathbb{Z}^m$ a finite set such that $\mathrm{supp}\,\phi \subset \cup_{h=1}^{2^m-1} \cup_{\mathbf{k}\in\mathbf{K}_{i_0-\ell}} \mathrm{supp}\,\Psi^h_{i_0-\ell,\mathbf{k}}$ and $\mathrm{supp}\,\phi \cap \sup \Psi^h_{i_0-\ell,\mathbf{k}} = \emptyset$ for all $\mathbf{k} \notin \mathbf{K}_{i_0-\ell}$ and all $h = 1, \ldots, 2^m - 1$. As (4.43) is recognized to be a linear regression of the form (4.10), with regressor $\sigma_{\ell+1} := \mathrm{col}(\Psi^h_{i_0-\ell,\mathbf{k}} \ : \ \mathbf{k} \in \mathbf{K}_{i_0-\ell})$, we define $z_{\ell+1}$ as an identifier of the form (4.12)-(4.13)

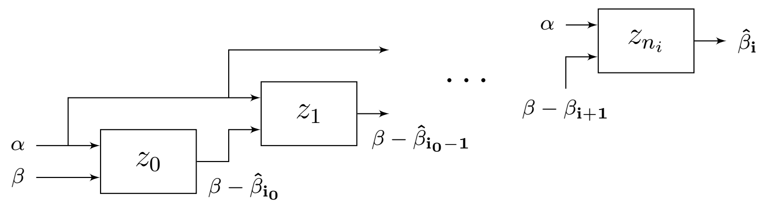

Figure 4.1: Block-diagram of the cascade structure of the scale-i wavelet identifier.

working on the inputs

$$\alpha_{\ell+1} := \alpha^\star + \delta_\alpha, \quad \beta_{\ell+1} := \beta^\star - \hat{\beta}_{\mathbf{i_0}-\ell} + \delta_\beta, \tag{4.44}$$

with $\delta = (\delta_\alpha, \delta_\beta)$ the same as in (4.38), and defined on the state space $\mathcal{Z}_{\ell+1} := \mathbb{R}^{d_{\ell+1}\times d_{\ell+1}} \times \mathbb{R}^{d_{\ell+1}}$, being $d_{\ell+1} := 2^{m-1}|\mathbf{K}_{i_0-\ell}|$. More precisely, by letting $\mathbf{z}_{\ell+1} := \mathrm{col}(z_0, \dots, z_{\ell+1})$, we define $z_{\ell+1}$ as

$$\mathcal{H}_{\ell+1} : \begin{cases} \dot{z}_{\ell+1} &= 0 \\ z^+_{\ell+1} &= F_{\ell+1}(\mathbf{z}_{\ell+1}, \alpha_{\ell+1}, \beta_{\ell+1}) \end{cases}$$

with flow and jump set given by $\mathcal{Z}_0 \times \cdots \times \mathcal{Z}_{\ell+1} \times \mathbb{R}^m \times \mathbb{R}$ and output

$$\theta_{\ell+1} = h_{\ell+1}(z_{\ell+1}),$$

being $F_{\ell+1}(\mathbf{z}_{\ell+1}, \alpha_{\ell+1}, \beta_{\ell+1}) := \mu_{\ell+1} z_{\ell+1} + (\sigma_{\ell+1}(\alpha_{\ell+1})\sigma_{\ell+1}(\alpha_{\ell+1})^T, \sigma_{\ell+1}(\alpha_{\ell+1})\beta_{\ell+1})$, with $\mu_{\ell+1} \in (0,1)$ and $\Omega_{\ell+1} \in \mathbb{R}^{d_{\ell+1}\times d_{\ell+1}}$ appropriately chosen.

In this way we can add an arbitrarily large number of least squares stages, obtaining an identifier that is able to approximate ϕ at arbitrarily fine scale. By adding $n_i \in \mathbb{N}$ stages to the initial scale-i_0 identifier z_0 we obtain a scale-i ($i = i_0 - n_i$) identifier having the form:

$$\mathcal{H}^i : \begin{cases} \dot{z}_\ell &= 0 \\ z^+_\ell &= F_\ell(\mathbf{z}_\ell, \alpha_\ell, \beta_\ell) \end{cases}, \qquad \ell = 0, \dots, n_i \tag{4.45}$$

with state $\mathbf{z}_{n_i} := (z_0, \dots, z_{n_i})$ defined on the state space $\mathcal{Z} := \mathcal{Z}_0 \times \cdots \times \mathcal{Z}_{n_i}$, with flow and jump sets given by $\mathcal{Z} \times \mathbb{R}^m \times \mathbb{R}$, and output

$$\theta^{\mathbf{i}} := \mathrm{col}\left(h_1(z_1), \dots, h_{n_i}(z_{n_i})\right). \tag{4.46}$$

Figure 4.1 depicts a block-diagram representation of the cascade structure of the identifier (4.45). We associate to $\mathcal{H}^i$ the corresponding scale-i prediction model

$$\Phi_i(\theta^{\mathbf{i}}, \cdot) := \sum_{\ell=0}^{n_i} \theta_\ell^T \sigma_\ell(\cdot) \ \in \ \mathbf{V}_i. \tag{4.47}$$

We denote by $\mathcal{H}_i^{cl}$ the overall system obtained by the interconnection of the exosystem (4.1)-(4.2) and the identifier (4.45), with $\alpha = \alpha^\star(w) + \delta_\alpha$ and $\beta = \beta^\star(w) + \delta_\beta$. In the same way as in Section 4.2, with $(r, N) \in (\mathbb{R}_+)^2$ arbitrarily chosen, we define the restriction $\mathcal{E}_{radt}^{r,N}$ on $\mathcal{H}^{cl}$ to be the set of solution pairs in $\mathcal{S}_{\mathcal{H}^{cl}}$ that satisfy the reverse average dwell-time condition (A.3) with parameters (r, N). Moreover, for arbitrary $(J, \epsilon) \in \mathbb{N} \times \mathbb{R}_+^*$, we let $\mathcal{E}_{PE}^{J,\epsilon} \subset \mathcal{S}_{\mathcal{H}_i^{cl}}$ be the set of solution pairs of $\mathcal{H}_i^{cl}$ such that the inputs $\alpha^\star(w)$ and $\alpha^\star(w) + \delta_\alpha$ are (J, ϵ)-PE for each stage z_ℓ, $\ell = 0, \ldots, n_i$. Then, the following holds:

Lemma 4.2. *with $\mathcal{H}_w$ given by* (4.1), (4.2), *suppose that Assumption 4.1 holds and pick $(r, N) \in (\mathbb{R}_+)^2$ and $(J, \epsilon) \in \mathbb{N} \times \mathbb{R}_+^*$ arbitrarily. Then for each $i \leq i_0$ and each $w \in \mathcal{S}_{\mathcal{H}_w}(W)$ there exists unique $\mathbf{z}_{n_i}^\star : \mathbb{R} \to \mathcal{Z}$ such that the identifier* (4.45) *fulfills the point 2 of the identifier requirement with restriction $\mathcal{E}_{PE}^{J,\epsilon} \cap \mathcal{E}_{radt}^{r,N}$, with $\underline{\sigma}, \bar{\sigma}$ linear and ρ locally Lipschitz.*

The proof of Lemma 4.2 follows from quite standard inductive arguments used to study cascade interconnections of stable systems and it is thus omitted. The main observation behind the proof is that if the lemma holds for a given $\ell \in \{0, \ldots, n_i - 1\}$, then in view of (4.44) the stage $z_{\ell+1}$ works on the inputs

$$\alpha_{\ell+1} = \alpha^\star + \delta_\alpha, \qquad\qquad \beta_{\ell+1} = \beta_{\ell+1}^\star + \delta_\beta^{\ell+1}$$

having defined

$$\begin{aligned} \beta_{\ell+1}^\star &:= \beta^\star - \Phi_{i_0-\ell}(\theta^{\star \mathbf{i_0}-\ell}, \alpha^\star) \\ \delta_\beta^\ell &:= \delta_\beta + \Phi_{i_0-\ell}(\theta^{\star \mathbf{i_0}-\ell}, \alpha^\star) - \Phi_{i_0-\ell}(\theta^{\mathbf{i_0}-\ell}, \alpha^\star + \delta_\alpha), \end{aligned}$$

with $\theta^{\star \mathbf{i_0}-\ell} := \operatorname{col}(h_1(z_1^\star), \ldots, h_\ell(z_\ell^\star))$. From Proposition 4.1 it follows that, for each solution pair in $(w, \mathbf{z}_{n_i}, \delta) \in \mathcal{E}_{PE}^{(J,\epsilon)}$ and all $(t, j) \in \operatorname{dom}(w, \mathbf{z}_\ell, \delta)|_{\geq J + t_J}$ and $\ell = 1, \ldots, n_i - 1$ we can bound δ_β^ℓ as a function of $|\delta|$ and $|\mathbf{z}_{\ell-1} - \mathbf{z}_{\ell-1}^\star|$ only. Thus, the overall cascade is a series interconnection of stable systems and the lemma is proved by induction. Lemma 4.2 is used to identify, for each $w \in \mathcal{S}_{\mathcal{H}^w}$, a steady state trajectory $z_\ell^\star$ for each stage $\ell = 0, \dots, n_i$. From Proposition 4.1 it follows that the system $\mathcal{H}^0$, given by (4.45)-(4.46) with $i = 0$, also satisfies the point 1 of the identifier requirement relatively to the cost functional

$$\begin{aligned} J_w^{i_0}(\theta_0)(t,j) &= \sum_{i=0}^{j} \mu^{j-i} |\varepsilon_{i_0}^\star(w(t_i,i),\theta_0)|^2 + \theta_0^T \Omega_0 \theta_0 \\ \varepsilon_{i_0}^\star(w,\theta_0) &:= \beta^\star(w) - \Phi_{i_0}(\theta_0, \alpha^\star(w)). \end{aligned} \tag{4.48}$$

Let $\theta_0^\star := h_0(z^\star)$. Then from Proposition 4.1 also follows that $\mathcal{H}_1$ (obtained by letting $i = 1$ in (4.45)-(4.46)) satisfies the point 1 of the identifier requirement relatively to the cost functional

$$J_w^{i_0-1}(\theta^{\mathbf{i_0-1}})(t,j) = J_w^{i_0}(\theta_0)(t,j) + \sum_{i=0}^{j} \mu^{j-i} |\varepsilon_{i_0-1}^\star(w(t_i,i),\theta_0^\star,\theta_1)|^2 + \theta_1^T \Omega_1 \theta_1$$

$$\varepsilon_{i_0-1}^\star(w,\theta_0,\theta_1) := \varepsilon_{i_0}^\star(w,\theta_0) - \theta_1^T \phi_1(\alpha^\star(w)).$$

In general, for $\ell = 0, \dots, n_i$, we can define the recursion

$$\begin{aligned} J_w^{i_0-\ell}(\theta^{\mathbf{i_0}-\ell})(t,j) &= J_w^{i_0-\ell+1}(\theta^{\mathbf{i_0}-\ell+\mathbf{1}})(t,j) \\ &\quad + \sum_{\nu=0}^{j} \mu_\ell^{j-\nu} |\varepsilon_{i_0-\ell}^\star(w(t_\nu,\nu),(\theta_0^\star,\dots,\theta_{\ell-1}^\star,\theta_\ell))|^2 + \theta_\ell^T \Omega_\ell \theta_\ell \end{aligned} \tag{4.49}$$

with $\theta_k^\star := h(z_k^\star)$ for $k = 0, \dots, \ell - 1$, with

$$\varepsilon_{i_0-\ell}^\star(w,\theta_0,\dots,\theta_\ell) := \sum_{k=i_0}^{i_0-\ell+1} \varepsilon_k(w,\theta^{\mathbf{k}}) - \theta_\ell^T \phi_\ell(\alpha^\star(w))$$

and originating from $J_w^{i_0}$ (i.e. with $\ell = 0$) that is given by (4.48). By inductive arguments, and in view of Lemma 4.2, it is thus possible to conclude the following:

Proposition 4.4. *Pick* $(r, N) \in (\mathbb{R}_+)^2$, $(J, \epsilon) \in \mathbb{N} \times \mathbb{R}_+^*$ *and* $i \in \mathbb{Z}_{\le i_0}$ *arbitrarily and, with* $\mathcal{H}_w$ *given by* (4.1), (4.2), *suppose that Assumption 4.1 holds and define* J_w^i *by letting* $\ell = n_i$ *in* (4.49). *Then the identifier* (4.45)-(4.46) *fulfills the identifier requirement relatively to the identification problem* $(\mathcal{H}_w, J_w^i)$ *with restriction* $\mathcal{E}_{PE}^{J,\epsilon} \cap \mathcal{E}_{radt}^{r,N}$ *with any* κ_1, *with* $\underline{\sigma}, \bar{\sigma}$ *and* κ *linear and* ρ *locally Lipschitz.*

Remark 4.6. It is worth noting that the same approximation attainable by the identifier (4.45)-(4.46) at scale i could be in principle obtained by a single least squares stage of the form (4.39) working on the linear regression (4.37) simply by letting $i_0 = i$ (i.e. by directly starting from a finer scale). Nevertheless, the cascade identifier (4.45) permits to add or remove detail stages without affecting the state of the coarser stages (if not indirectly by inducing a transitory in the system), and to separate the "learning dynamics" of each successive detail, that are parametrized by μ_ℓ. Changing resolution when a single stage is used means instead to perform a completely new experiment, thus directly inducing a new transitory in the parameter estimation. We also note that we can pass from the coefficients of the cascade identifier (4.45) to those of the corresponding single stage identifier and vice-versa by means of the *forward* and *inverse discrete wavelet transforms* (for further detail see e.g. (Strang and Nguyen, 1996; Walnut, 2002)). △

Remark 4.7. We also observe that there is a "natural" ordering in the choice of the forgetting factors μ_ℓ, $\ell = 0, \ldots, n_i$, that consists of taking $\mu_\ell \geq \mu_{\ell+1}$. In fact, coarser scales (i.e. lower ℓ) are usually associated to rougher, yet essential, traits; finer scales correspond instead to more "volatile" details. Hence, the forgetting factors μ_ℓ of coarser scales are naturally chosen larger than those at finer scales, as the *learning* of fundamental rough traits, associated to long-term memory, is slower to acquire and forget than details, associated instead to a short-term memory. △

5

Adaptive High-Gain Observers via System Identification Techniques

In this chapter we consider the problem of adaptive observers design by approaching adaptation as a system identification problem. We aim to define a co-design strategy for the observers and the identification algorithms yielding adaptive solutions that make sense in a broad data-driven context and that can be used for uncertainties more general than the usual parametric ones treated in canonical adaptive control frameworks. We seek a design in which adaptation can be cast and solved by different identification techniques, rather than by ad hoc algorithms, and in which performances can be evaluated in a system identification sense, without the conceptual need of a "true model" or a "true parameter". The attention for adaptive observation is strongly motivated by the fact that, in a given time-scale and in given coordinates, output regulation for nonlinear system is essentially an observation problem. Thus results achieved in the simpler context of observation are then easily portable to a regulation context.

The problem of designing adaptive observers for uncertain nonlinear systems boasts decades of active research. Most contributions developed in the 90s and early 2000s have focused on single-input-single-output systems possessing "canonical" forms, with the uncertainty that is concentrated into a set of finite parameters of known dimension entering linearly in the system equations (see e.g. (Bastin and Gevers, 1988; Marino and Tomei, 1992, 1995; Marino et al., 2001)). Related extensions to multivariable systems and more general forms appeared in (Besançon, 2000; Zhang, 2002). Adaptive observers designs for systems presenting a nonlinear parametrization in the uncertain parameters started to appear only quite recently (see (Farza et al., 2009; Tyukin et al., 2013; Afri et al., 2017; Besançon and Ţiclea, 2017) and the references therein). In particular, in (Farza et al., 2009) a general class of high-gain observers (Gauthier and Kupka, 2001) was enriched with an adaptation mechanism of the kind of those proposed in (Zhang, 2002); in (Tyukin et al., 2013), a fairly more general class of nonlinear parametrization was considered for uniformly observable systems; in (Afri et al., 2017) the theory of nonlinear Luenberger observers (Andrieu and Praly, 2006) was applied to estimate the state and parameters of uncertain linear systems and, in (Besançon and Ţiclea, 2017), more general system exhibiting nonlinearities both in the states and parameters are dealt with by using the same arguments of (Besançon and Ţiclea, 2007) in dealing with non-uniformly observable systems.

All the aforementioned approaches are strongly based on a classical "adaptive control" perspective, in which all the uncertainty is concentrated in a finite set of parameters with known dimension, whose knowledge would result in the knowledge of the *true system* to be observed. In line with the certainty equivalence principle, the uncertainty is usually dealt with by using an estimate of the true uncertain parameter, whose adaptation is carried out by ad hoc adaptation laws induced by a Lyapunov analysis or by immersion arguments. When linear systems are considered, parametric uncertainties reflect into variations of the plant's matrices in the usual Euclidean topology. The interest on this kind of perturbations is further motivated by the fact that such Euclidean topology is equivalent to the (weak) C^1 topology (Hirsch, 1994) in the space of linear maps. When nonlinear systems are considered, instead, restricting the focus to parametric uncertainties is conceptually a strong limitation (Bin et al., 2018a) yielding a quite non-standard topology in the space of the functions that define the plant.

Here we approach the problem of adaptive observer design under a different perspective: instead of assuming that the state to be estimated is generated by a process with known parametrization and unknown parameters, i.e. that there is a *true parameter* to be found, we look at adaptation of the system's model as a *system identification* problem, in which a model relating the input-output signals is inferred by the observations and adapted by using well known system identification techniques (Ljung, 1999; Ljung and Söderström, 1985; Sjöberg et al., 1995). The design of the observer follows a canonical construction, and it is designed so that the asymptotic observation error results to be directly related to the *prediction error* of the identifier used along the observed system's trajectories.

We specifically consider the observer design problem for continuous-time nonlinear systems of the form[1]

$$\begin{aligned} \dot{x}_i &= x_{i+1} \qquad i = 1, \dots, n-1 \\ \dot{x}_n &= \phi(x) \\ y &= x_1 \end{aligned} \tag{5.1}$$

with state $x \in \mathbb{R}^n$, $n \in \mathbb{N}$, output $y \in \mathbb{R}$ and with $\phi : \mathbb{R}^n \to \mathbb{R}$ an unknown function. Our goal is to reconstruct at best the state x of (5.1) by processing the output y and without the full knowledge of ϕ. For simplicity of exposition, we limit to single-output systems in the simple form (5.1). Nevertheless, we notice that the same arguments extend straightforwardly to more complex high-gain constructions such as (Astolfi and Marconi, 2015; Astolfi et al., 2018) and to multiple outputs. We construct a systematic framework in which identification schemes and *high-gain observers* (see e.g. (Gauthier and Kupka, 2001; Hammouri, 2007)) can be co-designed to solve general instances of the adaptive observation problem presented above. The chapter concludes with an example (Section 5.3), showing how the different parts of the observer can be tuned and how the wavelet multiresolution approach developed in Section 4.5 can be effective in approximating functions with unknown structure.

[1] We recall that any n-dimensional system that is *uniformly observable* in the sense of (Hammouri, 2007, Thm. 2.2.1) is locally diffeomorphic to (5.1).

5.1 The Observer's Structure

We construct an adaptive observer for the system (5.1) by joining together a high-gain observer and an identifier that fulfills the identifier requirement detailed in Chapter 4. As a standing assumption on (5.1), we suppose the following:

Assumption 5.1. *The initial conditions of* (5.1) *range in a compact invariant set* $X \subset \mathbb{R}^m$ *and* ϕ *is* C^2 *on an open set containing* X*.*

We consider here the following class of identifiers

$$\begin{cases} \dot{z} &= F(z,\alpha,\beta) \qquad (x,\alpha,\beta) \in \mathcal{Z} \times \mathbb{R}^n \times \mathbb{R} \\ z^+ &= G(z,\alpha,\beta) \qquad (z,\alpha,\beta) \in \mathcal{Z} \times \mathbb{R}^n \times \mathbb{R} \end{cases} \tag{5.2}$$

with $\mathcal{Z}$ an Euclidean space and output $\theta \in \mathbb{R}^d$ given by

$$\theta = h(z), \tag{5.3}$$

that is obtained from (4.4)-(4.5) by letting $\mathcal{A} = \mathbb{R}^n$, $\mathcal{B} = \mathbb{R}$ and F and G be single valued maps. To (5.2)-(5.3) we associate a prediction model $\Phi(\theta, \cdot)$ that we assume to be C^2. If ϕ were perfectly known, a high-gain observer would be sufficient to have an asymptotic exact estimate of x. The role of the identifier is to adapt a guess of ϕ on the basis of the available information. The relation

$$\dot{x}_n = \phi(x)$$

in (5.1) is seen as a prediction error model, where x plays the role of the regressor, ϕ of the "true model" and $\dot{x}_n$ of its output. The identifier (5.2)-(5.3) is designed by assuming to have available some measurements of the quantities $\alpha^\star = x$ and $\beta^\star = \dot{x}_n$ corrupted by an additive disturbance δ. In terms of Section 4.1, the observed system (5.1) is seen as an exosystem of the form (4.1), with state $w := x \in \mathbb{R}^n =: \mathcal{W}$ and output

$$(\alpha^\star(w), \beta^\star(w)) := (x, \phi(x)) = (x, \dot{x}_n). \tag{5.4}$$

The prediction error (4.7) assumes the expression

$$\varepsilon^\star(w, \theta) = \phi(x) - \Phi(\theta, x). \tag{5.5}$$

The a priori knowledge of the possible class of functions to which ϕ and x belong to may considerably help in finding an appropriate model set $\mathcal{M}$ and an appropriate parametrization $\Phi(\theta, \cdot)$ of the elements of $\mathcal{M}$. This qualitative a-priori information also guides the designer in the definition of the cost functional (4.8) and on the particular choice of the identification algorithm. For instance, if ϕ is known to be linear, then a model set composed of prediction models of the kind $\theta^T x$, with $\theta \in \mathbb{R}^d$ and $d = n$, and a least squares algorithm of the kind of those proposed in sections 4.2 of (4.2) are natural choices. If very few information is known about ϕ, an *universal approximator* approach of the kind of the wavelet identifier proposed in Section 4.5 can be used. In any case, from now on we assume that the identifier (5.2)-(5.3), the prediction model $\Phi(\theta, \cdot)$ and a cost functional J_w of the form (4.8) are fixed, and we denote by $\mathcal{H}_{(w,z)}$ the system composed by the exosystem $\mathcal{H}_w$ given by (5.1) and the identifier (5.2)-(5.3) obtained by letting $(\alpha, \beta) = (\alpha^\star(w), \beta^\star(w)) + \delta$, with $(\alpha^\star, \beta^\star)$ given by (5.4) and δ a hybrid input with values in $\mathbb{R}^n \times \mathbb{R}$. Then we assume the following:

Assumption 5.2. *There exist non-empty* $\mathcal{E}_0 \subset \mathcal{S}_{\mathcal{H}_{(w,z)}}$*,* $\nu > 0$*, locally Lipschitz* $\kappa_1, \kappa_2, \rho \in \mathcal{K}$ *and locally linear*[2] *functions* $\underline{\sigma}$ *and* $\bar{\sigma}$ *such that the identifier* (5.2)-(5.3) *satisfies the identifier requirement relatively to the identification problem* $(\mathcal{H}_w, J_w)$ *with restriction* $\mathcal{E}_0$*.*

Assumption 5.3. *With* X *the set for which Assumption 5.1 holds, there exist* $T > 0$ *and a compact set* $Z \subset \mathcal{Z}$ *such that, for each solution* x *of* (5.1) *originating in* X*, the corresponding trajectory* $z^\star$ *for which the identifier requirement is satisfied fulfils* $z^\star(t, j) \in Z$ *for all* $(t, j) \in \operatorname{dom} z^\star|_{\geq T}$*.*

The observer subsystem is a hybrid version of a classical high-gain observer, extended by one state. It is defined over the state space $\mathbb{R}^{n+1}$ and its equations read as follows:

$$
\begin{cases}
\dot{\hat{x}}_i &= \hat{x}_{i+1} + c_i g^i (y - \hat{x}_1), \qquad i = 1, \dots, n \\
\dot{\hat{x}}_{n+1} &= \psi(z, \hat{x}) + c_{n+1} g^{n+1} (y - \hat{x}_1)
\end{cases}
$$
$$
\begin{cases}
\hat{x}_i^+ &= \hat{x}_i, \qquad i = 1, \dots, n \\
\hat{x}_{n+1}^+ &= \ell(z, \hat{x})
\end{cases}
$$

with flow and jump set given by $\mathbb{R}^{n+1} \times \mathcal{Z}$. The coefficients c_i's are chosen so

[2] Namely that can be taken linear on each compact set.

that the roots of $\lambda^{n+1} + c_{n+1}\lambda^n + \cdots + c_2\lambda + c_1$ have negative real part and $g > 0$ is a control parameter to be fixed later. The intuitive reason to use an observer extended of one state is that in this way $\hat{x}$ gives a proxy for the data $(x, \dot{x}_n)$ that are needed by the identifier; an observer of order n could not produce a proxy variable for $\dot{x}_n$.

The functions ψ and ℓ are defined as follows: let $\Phi', \Phi^+ : \mathcal{Z} \times \mathbb{R}^n \times \mathbb{R}$ be defined as

$$\begin{aligned}
\Phi'(z, \alpha, \beta) &:= \frac{\partial \Phi(h(z), \alpha)}{\partial \alpha} A \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \frac{\partial \Phi(h(z), \alpha)}{\partial z} F(z, \alpha, \beta) \\
\Phi^+(z, \alpha, \beta) &= \Phi\big(h(G(z, \alpha, \beta)), \alpha\big)
\end{aligned}$$

where $A \in \mathbb{R}^{n \times (n+1)}$ is the "shift matrix", defined by letting $A_{i,i+1} = 1$ for $i = 1, \ldots, n$ and zero otherwise. With $\widehat{X} \subset \mathbb{R}^{n+1}$ and $\widehat{Z} \subset \mathcal{Z}$ arbitrary compact subsets such that the sets X and Z defined in assumptions 5.1 and 5.3 satisfy $X \subset \widehat{X}$ and $Z \subset \widehat{Z}$, we define the functions Φ'_s and Φ^+_s to be any bounded functions that agree respectively with Φ' and Φ^+ on $\widehat{Z} \times \widehat{X}$. In particular, for $\star \in \{', +\}$ we ask for the existence of $M > 0$ such that

$$|\Phi^\star_s(\nu)| \leq M \qquad \nu \in \mathcal{Z} \times \mathbb{R}^n \times \mathbb{R}. \tag{5.6}$$

The functions ψ and ℓ are thus chosen as

$$\begin{aligned}
\psi(z, \hat{x}) &:= \Phi'_s(z, \hat{x}_{[1,n]}, \hat{x}_{n+1}), \\
\ell(z, \hat{x}) &:= \Phi^+_s(z, \hat{x}_{[1,n]}, \hat{x}_{n+1}),
\end{aligned}$$

being $\hat{x}_{[1,n]} := \mathrm{col}(\hat{x}_1, \ldots, \hat{x}_n)$.

We complete the design by letting in (5.2)

$$\alpha := \hat{x}_{[1,n]}, \qquad\qquad \beta := \hat{x}_{n+1},$$

obtaining the overall system

$$
\left\{\begin{array}{lcl}
\dot{x}_i & = & x_{i+1}, \qquad\qquad\qquad i = 1, \dots, n-1 \\
\dot{x}_n & = & \phi(x) \\
\dot{\hat{x}}_i & = & \hat{x}_{i+1} + c_i g^i (x_1 - \hat{x}_1), \quad i = 1, \dots, n \\
\dot{\hat{x}}_{n+1} & = & \psi(z, \hat{x}) + c_{n+1} g^{n+1} (x_1 - \hat{x}_1) \\
\dot{z} & = & F(z, \hat{x}_{[1,n]}, \hat{x}_{n+1})
\end{array}\right.
\left\{\begin{array}{lcl}
x^+ & = & x \\
\hat{x}_i^+ & = & \hat{x}_i, \qquad i = 1, \dots, n \\
\hat{x}_{n+1}^+ & = & \ell(z, \hat{x}) \\
z^+ & = & G(z, \hat{x}_{[1,n]}, \hat{x}_{n+1})
\end{array}\right. \tag{5.7}
$$

with flow and jump set given by $\mathbb{R}^{n+1} \times \mathbb{R}^{n+1} \times \mathcal{Z}$. For simplicity, in the following we let $\mathbf{x} := (x, \hat{x}, z)$ and $\mathcal{X} := \mathbb{R}^{n+1} \times \mathbb{R}^{n+1} \times \mathcal{Z}$. With $(\lambda, N_0) \in \mathbb{R}_+ \times \mathbb{N}$, we let $\mathcal{E}_{adt}^{\lambda, N_0}$ be the set of solutions of (5.7) that satisfy the average dwell-time condition (A.2) with parameters (λ, N_0).

5.2 Asymptotic Properties

According to A5.2, to each solution pair $(x, z, \delta) \in \mathcal{E}_0$ is associated an optimal trajectory $z^\star$ for the identifier z such that the corresponding output $\theta^\star$ minimizes pointwise J_w. In the overall system (5.7) the identifier can be seen as fed by the ideal input $(x, \phi(x))$ plus a disturbance given by $\delta = (\hat{x}_{[1,n]} - x, \hat{x}_{n+1} - \phi(x))$. Thus, for each solution $\mathbf{x}$ of the overall system (5.7) such that $(x, z, \delta) \in \mathcal{E}_0$, the quantity

$$
\varepsilon^\star(x(t,j), \theta^\star(t,j)) = \phi(x(t,j)) - \Phi(\theta^\star(t,j), x(t,j)), \tag{5.8}
$$

derived from (5.5), is well defined and corresponds to the optimal prediction error of the identifier when acting on the unperturbed input x and when the state estimate mismatch δ is zero. We let $\mathcal{E}_1$ be the set of solutions of (5.7) such that, with $\delta = (\hat{x}_{[1,n]} - x, \hat{x}_{n+1} - \phi(x))$, $(x, z, \delta) \in \mathcal{E}_0$. Then the following result holds:

Proposition 5.1. *Under assumptions 5.1, 5.2 and 5.3 there exist $g^\star > 0$, $\lambda^\star(g) > 0$ and for each $N_0 \in \mathbb{N}^*$, a $c > 0$, such that for every $g > g^\star$ and every $\lambda \in (0, \lambda^\star(g))$, all* *the solutions in* $\mathcal{E}_1 \cap \mathcal{E}_{adt}^{\lambda, N_0}$ *are bounded and satisfy*

$$\limsup |x - \hat{x}_{[1,n]}| \leq c \limsup |\varepsilon^\star(x, \theta^\star)|.$$

Proposition 5.1 states that if the control parameter g is suitably chosen, then for any solution of (5.7) along which the identifier satisfies the identifier requirement and the flows are persistent, the asymptotic estimation error is linearly bounded by the prediction capabilities of the identifier *evaluated along the observed system's trajectories*. We observe how the ideas of the theory of dual control (Feldbaum, 1960) emerge under this simple design: the control parameter g of the high-gain observer is chosen large enough to make sure that, despite the initial error in the approximation of ϕ, the state estimate $\hat{x}_{[1,n]}$ gets close enough to the actual plant's state x, allowing the identifier to work on data (α, β) that are *close enough to the ideal quantities* $(\alpha^\star, \beta^\star)$. This in turn allows the identifier to identify the function ϕ at meaningful points, so as the identified dynamics reflects the actual underlying movements of the real plant. The strong stability and regularity properties of the identifier requirements make sure that small estimation errors of the state reflect into small differences in the identified model and the loop is closed by further asking the control parameter g to induce a contraction in of the nested relationship linking the identifier performances and the quality of the state estimate.

Proof of Proposition 5.1. Boundedness of $(\hat{x}, z)$ for all the solutions in $\mathcal{E}_1$ follows by standard high-gain arguments in view of the boundedness property (5.6) of ψ and ℓ and because of A5.3 and the stability property (point 2) of the identifier requirement. In particular the fact that ψ and ℓ are bounded allows us to conclude that the trajectories of $\hat{x}$ are uniformly ultimately bounded, i.e. there exist a compact set $\widehat{X}' \subset \mathbb{R}^{n+1}$ containing X (with X the set of A5.1) and, for each compact set $\widehat{X}_0 \subset \mathbb{R}^{n+1}$, a $T > 0$, such that every solution of the overall system (5.7) with $g > 0$ and originating in $X \times \widehat{X}_0 \times \mathbb{Z}$ satisfies $\hat{x}(t,j) \in \widehat{X}'$ for all $(t,j) \in \operatorname{dom} \hat{x}|_{\geq T}$. On the other hand, we can see the identifier z as driven by the ideal input $(\alpha^\star, \beta^\star) = (x, \phi(x))$ perturbed by the disturbance

$$\delta = \begin{pmatrix} \hat{x}_{[1,n]} - x \\ \hat{x}_{n+1} - \phi(x) \end{pmatrix}. \tag{5.9}$$

Then Assumption 5.2 ensures the existence, for every solution in $\mathcal{E}_1$, of a $z^\star$ such that the identifier requirement holds. As δ is eventually bounded, this and Assumption 5.3 imply the existence of compact sets $\widetilde{X}' \subset \mathbb{R}^{n+1}$ and $Z', \widetilde{Z}' \subset \mathcal{Z}$ such that

$$\begin{aligned} &\forall \mathbf{x} \in \mathcal{E}_1,\ \exists T > 0 \text{ s.t. } \forall t + j \geq T \\ &\qquad \hat{x}(t,j) \in \widehat{X}', \quad \delta(t,j) \in \widetilde{X}', \quad z(t,j) \in Z', \quad z(t,j) - z^\star(t,j) \in \widetilde{Z}' \end{aligned} \tag{5.10}$$

with δ given by (5.9) and where T can be taken the same for all $\mathbf{x} \in \mathcal{E}_1$ that originates in the same compact set.

Pick $\mathbf{x} \in \mathcal{E}_1$, let $\theta^\star = h(z^\star)$ and change of variables as

$$\begin{aligned} \tilde{\chi}_i &:= g^{1-i}(\hat{x}_i - x_i), \qquad i = 1, \ldots, n \\ \tilde{\chi}_{n+1} &:= g^{-n}(\hat{x}_{n+1} - \Phi(\theta^\star, x)) \\ \tilde{z} &:= z - z^\star. \end{aligned} \tag{5.11}$$

Let us denote $\tilde{\mathbf{x}} := (x, \tilde{\chi}, \tilde{z})$. For $i \in \{1, \ldots, n-1\}$ we have

$$\dot{\tilde{\chi}}_i = g(\tilde{\chi}_{i+1} + c_i \tilde{\chi}_1).$$

For $i = n$, adding and subtracting $g^{1-n}\Phi(\theta^\star, x)$ yields

$$\begin{aligned} \dot{\tilde{\chi}}_n &= g^{1-n}\big(\hat{x}_{n+1} + c_n g^n \tilde{\chi}_1 - \phi(x)\big) \\ &= g\tilde{\chi}_{n+1} + c_i g \tilde{\chi}_1 - g^{1-n}\varepsilon^\star(x, \theta^\star). \end{aligned}$$

with $\varepsilon^\star$ given by (5.8). Finally, for $i = n+1$, we obtain

$$\dot{\tilde{\chi}}_{n+1} = g c_{n+1} \tilde{\chi}_1 + g^{-n}\big(\psi(z, \hat{x}) - \Phi'(z^\star, x, \phi(x))\big).$$

We observe that the new variables fulfill the following bounds

$$\begin{aligned} |\hat{x}_{[1,n]} - x| &\leq g^{n-1}|\tilde{\chi}| \\ |\hat{x}_{n+1} - \phi(x)| &= |g^n \tilde{\chi}_{n+1} + \Phi(\theta^\star, x) - \phi(x)| \\ &\leq g^n |\tilde{\chi}| + |\varepsilon^\star(x, \theta^\star)|. \end{aligned} \tag{5.12}$$

Furthermore, we can assume without loss of generality that the number T for which (5.10) holds is the same for which the point 3 of the identifier requirement holds. Then, as κ_1 is locally Lipschitz and $\tilde{z} \in \widetilde{Z}'$ in $\mathcal{I}(\tilde{\mathbf{x}})|_{\geq T}$, the boundedness (5.6) of ψ implies the existence of $L > 0$ such that, for all $(t,j) \in \mathcal{I}(\tilde{\mathbf{x}})|_{\geq T}$

$$\begin{aligned}
&|\psi(z,\hat{x}) - \Phi'(z^\star, x, \phi(x))| \\
&\quad = |\Phi'_s(z, \hat{x}_{[1,n]}, \hat{x}_{n+1}) - \Phi'(z^\star, x, \phi(x))| \\
&\quad \leq L\big(|z - z^\star| + |\hat{x}_{[1,n]} - x| + |\hat{x}_{n+1} - \phi(x)|\big) \\
&\quad \leq L\big(|\tilde{z}| + g^n|\tilde{\chi}| + |\varepsilon^\star(x,\theta^\star)|\big).
\end{aligned} \tag{5.13}$$

During jumps, instead, we have $\tilde{\chi}_i^+ = \tilde{\chi}_i$ for $i = 1, \ldots, n$, while for $i = n+1$

$$\tilde{\chi}_{n+1}^+ = g^{-n}\big(\ell(z,\hat{x}) - \Phi^+(z,x,\phi(x))\big).$$

As κ_2 is locally Lipschitz, the same arguments used above show that, for some $L > 0$ that we take without loss of generality equal to those used in (5.13) (otherwise change it with the maximum of the two), the following bound holds

$$|\ell(z,\hat{x}) - \Phi^+(z,x,\phi(x))| \leq L\big(|\tilde{z}| + |\varepsilon^\star(x,\theta^\star)|\big) + g^n L|\tilde{\chi}|. \tag{5.14}$$

for all $(t,j) \in \Gamma(\tilde{\mathbf{x}})|_{\geq T}$.

We can thus write the dynamics of $\tilde{\chi}$ using the following equations (for compactness we omit the argument of $\varepsilon^\star$)

$$\begin{aligned}
\dot{\tilde{\chi}} &= gH\tilde{\chi} + g^{-n}B(\psi(z,\hat{x}) - \Phi'(z^\star, \phi(x))) + D(g)\varepsilon^\star \\
\tilde{\chi}^+ &= R\tilde{\chi} + g^{-n}B(\ell(z,\hat{x}) - \Phi^+(z,x,\phi(x)))
\end{aligned}$$

with H a Hurwitz matrix (because of the choice of the coefficients c_i), $B := \mathrm{col}(0,\ldots,0,1) \in \mathbb{R}^{n+1}$, $R := \mathrm{diag}(I_n, 0)$, $D(g) := \mathrm{col}(0,\ldots,g^{1-n},g^{-n})$ and with the terms $\psi(z,\hat{x}) - \Phi'(z^\star,\phi(x))$ and $\ell(z,\hat{x}) - \Phi^+(z,x,\phi(x))$ that satisfy the bounds (5.13) and (5.14) in $\operatorname{dom}\tilde{\mathbf{x}}|_{\geq T}$. Consider the function

$$U(\tilde{\chi}) := \sqrt{\tilde{\chi}^T P \tilde{\chi}}$$

where $P = P^T > 0$ is the unique solution to the Lyapunov equation

$$H^T P + PH = -I.$$

Then there exists constants $a_1, a_2 > 0$ such that

$$a_1|\tilde{\chi}| \leq U(\tilde{\chi}) \leq a_2|\tilde{\chi}|, \quad |x - \hat{x}_{[1,n]}| \leq g^{n-1}|\tilde{\chi}|. \tag{5.15}$$

Furthermore, for any solution $\mathbf{x} \in \mathcal{E}_1$ and all $(t, j) \in \mathcal{I}(\mathbf{x})$

$$D^+U(\tilde{\chi}) = \frac{1}{2U(\tilde{\chi})}\Big(- g|\tilde{\chi}|^2 + 2g^{-n}\tilde{\chi}^T PB(\psi(z, \hat{x}) - \Phi'(z^\star, \phi(x))) + 2\tilde{\chi}^T PD(g)\varepsilon^\star\Big).$$

Using (5.13) and (5.15) yields, for all $(t, j) \in \mathcal{I}(\tilde{\mathbf{x}})|_{\geq T}$,

$$D^+U(\tilde{\chi}) \leq (b_2 - gb_1)U(\tilde{\chi}) + b_2\big(g^{-n}|\tilde{z}| + g^{1-n}|\varepsilon^\star|\big),$$

with $b_1 := 1/(2a_2)$ and $b_2 := L|P|\sqrt{(}2)/a_1$. Hence, in view of (5.15), choosing $g > g_1^\star := 2b_1b_2$ yields

$$D^+U(\tilde{\chi}) \leq -gb_3U(\tilde{\chi}) + b_2\big(g^{-n}|\tilde{z}| + g^{1-n}|\varepsilon^\star|\big), \tag{5.16}$$

with $b_3 := b_1/2$. On the other hand, using (5.14) and (5.15) yields

$$\begin{aligned} U(\tilde{\chi})^+ &:= U(R\tilde{\chi} + g^{-n}B(\ell(z, \hat{x}) - \Phi^+(z, x, \phi(x)))) \\ &\leq \sqrt{2|P|}\big(\|\tilde{\chi}| + g^{-n}|\ell(z, \hat{x}) - \Phi^+(z, x, \phi(x))|\big) \\ &\leq b_4U(\tilde{\chi}) + b_5g^{-n}\big(|\tilde{z}| + |\varepsilon^\star|\big), \end{aligned} \tag{5.17}$$

for all $(t, j) \in \Gamma(\tilde{\mathbf{x}})|_{\geq T}$ and with $b_4 := \sqrt{2|P|}L/a_1$ and $b_5 := b_4a_1$.

In view of Assumption 5.2, the functions $\underline{\sigma}$, $\bar{\sigma}$ and ρ of point 2 of the identifier requirement can be taken linear on $\widetilde{X}'$ and $\widetilde{Z}'$ respectively. Therefore, (5.16) and (5.17), give the existence of constants $h_1, h_2, h_3 > 0$ such that, for all $(t, j) \in \mathcal{I}(\tilde{\mathbf{x}})|_{\geq T}$

$$\begin{aligned} U(\tilde{\chi}) \geq h_1 \max\big\{g^{-(n+1)}V_z(\tilde{z}), g^{-n}|\varepsilon^\star|\big\} \\ \implies D^+U(\tilde{\chi}) \leq -gh_2U(\tilde{\chi}), \end{aligned} \tag{5.18}$$

and for all $(t, j) \in \Gamma(\tilde{\mathbf{x}})|_{\geq T}$,

$$U(\tilde{\chi})^+ \leq h_3 \max\big\{U(\tilde{\chi}), g^{-n}V_z(\tilde{z}), g^{-n}|\varepsilon^\star|\big\}. \tag{5.19}$$

Moreover, equations (5.12) and (5.15) imply that, δ defined in (5.9), satisfies

$$|\delta| \le g^n|\tilde{\chi}| + |\varepsilon^\star| \le g^n U(\tilde{\chi})/a_1 + |\varepsilon^\star|,$$

so that, under Assumption 5.2, point 2 of the identifier requirement implies that, for some $h_4 > 0$ and for all $(t,j) \in \mathcal{I}(\tilde{\mathbf{x}})|_{\ge T}$,

$$\begin{aligned} V_z(\tilde{z}) \ge h_4 \max\{g^n U(\tilde{\chi}), |\varepsilon^\star|\} \\ \implies D^+ V_z(\tilde{z}) \le -\nu V_z(\tilde{z}), \end{aligned} \tag{5.20}$$

while for some $h_5 > 0$ and all $(t,j) \in \Gamma(\tilde{\mathbf{x}})|_{\ge T}$

$$V_z(\tilde{z})^+ \le \max\{e^{-\nu} V_z(\tilde{z}), h_5 g^n U(\tilde{\chi}), h_5|\varepsilon^\star|\}. \tag{5.21}$$

Let define the function

$$W(\tilde{\chi}, \tilde{z}) := \max\left\{U(\tilde{\chi}), \frac{V_z(\tilde{z})}{h_4 g^n}\right\},$$

and pick

$$g > g^\star := \max\{g_1^\star,\ h_1 h_4\}. \tag{5.22}$$

Let $\mathcal{I}_z(\tilde{\mathbf{x}})$ be the set of $(t,j) \in \mathcal{I}(\tilde{\mathbf{x}})|_{\ge T}$ such that $V_z(\tilde{z}(t,j)) \ge h_4 g^n U(\tilde{\chi}(t,j))$ and let $\mathcal{I}_U(\tilde{\mathbf{x}}) := \mathcal{I}(\tilde{\mathbf{x}})|_{\ge T} \setminus \mathcal{I}_z(\tilde{\mathbf{x}})$. Pick $(t,j) \in \mathcal{I}_z(\tilde{\mathbf{x}})$, then $W(\tilde{\chi}, \tilde{z}) = V_z(\tilde{z})/(h_4 g^n)$ and, thus, the relation

$$W(\tilde{\chi}, \tilde{z}) \ge g^{-n}|\varepsilon^\star|$$

implies $V_z(\tilde{z}) \ge h_4 \max\{g^n U(\tilde{\chi}), |\varepsilon^\star|\}$. Hence (5.20) yields

$$D^+ W(\tilde{\chi}, \tilde{z}) \le -\nu W(\tilde{\chi}, \tilde{z}).$$

If instead $(t,j) \in \mathcal{I}_U(\tilde{\mathbf{x}})$, then $W(\tilde{\chi}, \tilde{z}) = U(\tilde{\chi})$, and (5.22) implies that $1/h_4 > h_1/g$, that in turn yields

$$U(\tilde{\chi}) > \frac{V_z(\tilde{z})}{h_4 g^n} > h_1 g^{-(n+1)} V_z(\tilde{z}).$$

Hence, if

$$W(\tilde{\chi}, \tilde{z}) \ge h_1 g^{-n}|\varepsilon^\star|,$$

it holds that $U(\tilde{\chi}) \geq h_1 \max\{g^{-(n+1)} V_z(\tilde{z}), g^{-n}|\varepsilon^\star|\}$, and thus (5.18) implies

$$D^+ W(\tilde{\chi}, \tilde{z}) \leq -g h_2 W(\tilde{\chi}, \tilde{z}).$$

By putting all together, we thus conclude that for all $(t, j) \in \mathcal{I}(\tilde{\mathbf{x}})|_{\geq T}$ and with $\bar{h} := \max\{1, h_3\}$ and $\bar{\nu}(g) := \min\{\nu, g h_2\}$, it holds that

$$W(\tilde{\chi}, \tilde{z}) \geq \bar{h} g^{-n} |\varepsilon^\star| \implies D^+ W(\tilde{\chi}, \tilde{z}) \leq -\bar{\nu}(g) W(\tilde{\chi}, \tilde{z}). \tag{5.23}$$

On the other hand, in view of (5.19) and (5.21), we obtain

$$\begin{aligned} W(\tilde{\chi}, \tilde{z})^+ \leq \max\Big\{& h_3 U(\tilde{\chi}), h_3 h_4 \frac{V_z(\tilde{z})}{h_4 g^n}, h_3 g^{-n}|\varepsilon^\star|, \\ & e^{-\nu} \frac{V_z(\tilde{z})}{h_4 g^n}, (h_5/h_4) U(\tilde{\chi}), (h_5/h_4) g^{-n} |\varepsilon^\star| \Big\} \\ \leq\ & \bar{q} \max\left\{ W(\tilde{\chi}, \tilde{z}), g^{-n}|\varepsilon^\star| \right\} \end{aligned}$$

for all $(t, j) \in \Gamma(\tilde{\mathbf{x}})|_{\geq T}$ and with $\bar{q} := \max\{h_3, h_3 h_4, e^{-\nu}, h_5/h_4\}$.

We observe that, as $\bar{q}$ is not necessarily in $[0, 1)$, W is not an ISS-Lyapunov function (Cai and Teel, 2009). Nevertheless, by following (Cai et al., 2008), we can augment the system (5.7) with a virtual clock subsystem with state $\tau \in \mathbb{R}$ satisfying

$$\begin{cases} \dot{\tau} & \in & F_\tau(\tau) & \tau & \in [0, N_0] \\ \tau^+ & = & \tau - 1 & \tau & \in [1, N_0] \end{cases} \tag{5.24}$$

with

$$F_\tau(\tau) := \begin{cases} \lambda & \tau \in [0, N_0) \\ [0, \lambda] & \tau = N_0. \end{cases}$$

Then, by (Cai et al., 2008, Prop. 1.1), for any $(\lambda, N_0) \in \mathbb{R}_+ \times \mathbb{N}$, every solution $\mathbf{x}$ of the overall system (5.7) is in $\mathcal{E}_{adt}^{\lambda, N_0}$ if and only if $(\mathbf{x}, \tau)$ is a solution of (5.7), (5.24). Pick $\upsilon \in (0, \bar{\nu}(g))$ and let k and $\lambda^\star(g)$ be such that

$$k \geq \upsilon + \log \bar{q}, \qquad \lambda^\star(g) \in \left(0, \frac{\bar{\nu}(g) - \upsilon}{k}\right].$$

Pick $\lambda \in (0, \lambda^\star(g)]$, $N_0 \in \mathbb{N}$ and define the function

$$V(\tilde{\chi}, \tilde{z}, \tau) := \exp(k\tau) W(\tilde{\chi}, \tilde{z}).$$

Let $\mathcal{E}$ be the set of solutions of the extended system (5.7), (5.24) such that, for each $(\mathbf{x}, \tau) \in \mathcal{E}$, $\mathbf{x} \in \mathcal{E}_1 \cap \mathcal{E}_{adt}^{(\lambda, N_0)}$. Let $\tilde{\mathbf{x}}$ be derived by $\mathbf{x}$ as before according to (5.11), and let $\widetilde{\mathcal{E}}$ be the set of solutions of the form $(\tilde{\mathbf{x}}, \tau)$ such that $(\mathbf{x}, \tau) \in \mathcal{E}$. Then, for all $(\tilde{\mathbf{x}}, \tau) \in \mathcal{E}$ an for all $(t, j) \in \mathcal{I}(\tilde{\mathbf{x}}, \tau)|_{\geq T}$, (5.23) implies

$$
\begin{aligned}
V(\tilde{\chi}, \tilde{z}, \tau) \geq e^{kN_0} \bar{h} g^{-n} |\varepsilon^\star| &\implies W(\tilde{\chi}, \tilde{z}) \geq \bar{h} g^{-n} |\varepsilon^\star| \\
&\implies D^+ V(\tilde{\chi}, \tilde{z}, \tau) \leq (k\lambda - \bar{\nu}(g)) V(\tilde{\chi}, \tilde{z}, \tau) \\
&\implies D^+ V(\tilde{\chi}, \tilde{z}, \tau) \leq -\upsilon V(\tilde{\chi}, \tilde{z}, \tau).
\end{aligned}
$$

Furthermore, for all $(t, j) \in \Gamma(\tilde{\mathbf{x}}, \tau)|_{\geq T}$,

$$
\begin{aligned}
V(\tilde{\chi}, \tilde{z}, \tau)^+ &\leq \bar{q} e^{k(\tau - 1)} \max\{W(\tilde{\chi}, \tilde{z}), g^{-n} |\varepsilon^\star|\} \\
&\leq \max\{\bar{q} e^{-\upsilon - \log \bar{q}} V(\tilde{\chi}, \tilde{z}, \tau), \bar{q} e^{kN_0} g^{-n} |\varepsilon^\star|\} \\
&\leq \max\{e^{-\upsilon} V(\tilde{\chi}, \tilde{z}, \tau), \bar{q} e^{k(N_0 - 1)} g^{-n} |\varepsilon^\star|\}.
\end{aligned}
$$

From these latter two relations, we thus conclude (see (Cai and Teel, 2009, Thm. 3.1)) that there exists $\bar{c} > 0$ such that

$$
\limsup V(\tilde{\chi}, \tilde{z}, \tau) \leq \bar{c} g^{-n} \limsup |\varepsilon^\star|.
$$

Since

$$
|x - \hat{x}_{[1,n]}| \leq g^{n-1} |\tilde{\chi}| \leq g^{n-1} U(\tilde{\chi})/a_1 \leq g^{n-1} W(\tilde{\chi}, \tilde{z})/a_1 \leq g^{n-1} V(\tilde{\chi}, \tilde{z}, \tau)/a_1,
$$

then

$$
\limsup |x - \hat{x}_{[1,n]}| \leq \bar{c}/(a_1 g) \limsup |\varepsilon^\star| \leq c \limsup |\varepsilon^\star|
$$

with $c := \bar{c}/a_1$, and the claim follows. ■

5.3 An Example

We present here an example in which the adaptive observer developed in Section 5.1 is used to estimate the state of the following system

$$
\begin{aligned}
\dot{x}_1 &= x_2 \\
\dot{x}_2 &= 5 \sin(2x_1) + 2x_1 - x_1^3.
\end{aligned}
\tag{5.25}
$$

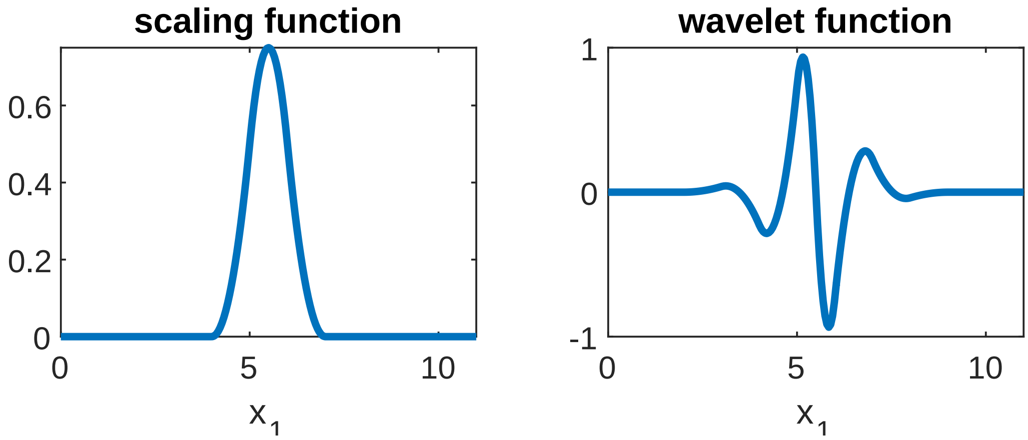

Figure 5.1: Scaling function and wavelet function.

It can be shown that for each initial condition $x(0)$, there exists a compact set X for which Assumption 5.1 holds. System (5.25) has the form (5.1) with $n = 2$ and

$$\phi(x) = \phi(x_1) = 5\sin(2x_1) + 2x_1 - x_1^3.$$

As a first step we construct and adaptive observer of the kind presented in Section 5.1 with the identifier that performs a wavelet expansion of ϕ. As ϕ depends only on x_1, for simplicity of exposition we will use 1-dimensional wavelets, i.e. in the identifier presented in Section 4.5 we let $m = 1$ and we consider functions inside $L_2(\mathbb{R})$ that only depend on x_1. We choose a biorthogonal B-spline construction[3] for the mother scaling and wavelet functions, that are represented in Figure 5.1.

The results reported below are obtained along a solution of (5.25) originating in $x(0) = (-2.5, 3)$, with the observer parameters $g = 30$, $M = 1000$ and, to underline how additional details reflect into the state estimation error, we have implemented a wavelet identifier with growing resolution. In particular, for the first 60 seconds no identifier is present, and the observer is implemented with $\psi = 0$ and $\ell(z, \hat{x}) = \hat{x}_3$. At time $60s$ the identifier switches to a single least-squares stage z_0 working at the initial scale $i_0 = 4$, and the functions ψ and ℓ are chosen accordingly. At time $120s$ a second stage, z_1, is added to obtain a representation at scale $i_1 = i_0 - 1 = 3$. At time $180s$ a third stage is added to reach a representation at scale $i_2 = i_0 - 2 = 2$. At time $240s$ a last stage is added to reach a representation

[3]They can be obtained in MATLAB by using the command `wavefun` of the Wavelet Toolbox and passing as argument the name ’`bior3.5`’.

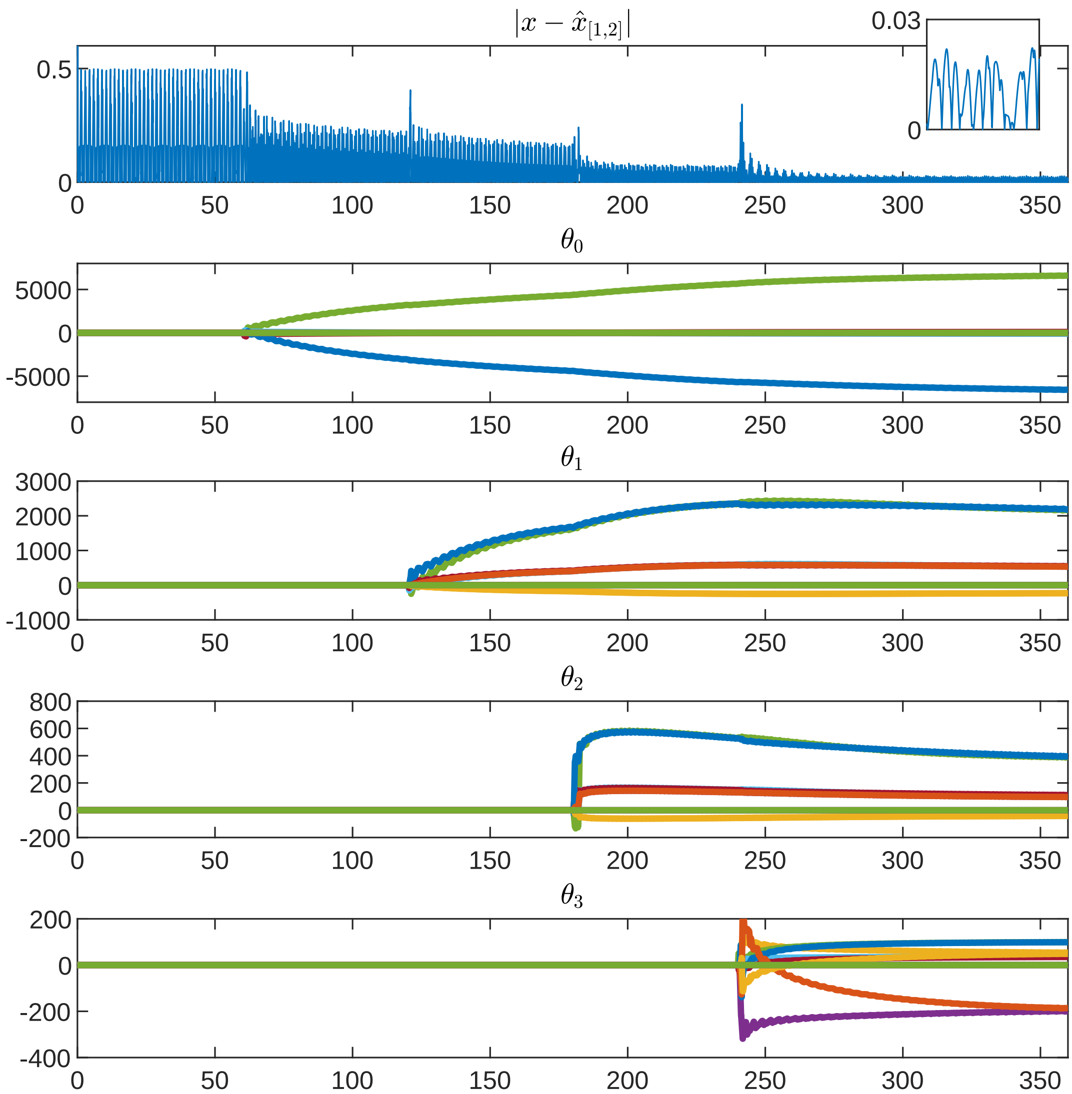

Figure 5.2: The first plot shows the time behavior of the state estimation error. The other four plots show the time evolution of the outputs θ_ℓ of the least-squares stages of the wavelet identifier.

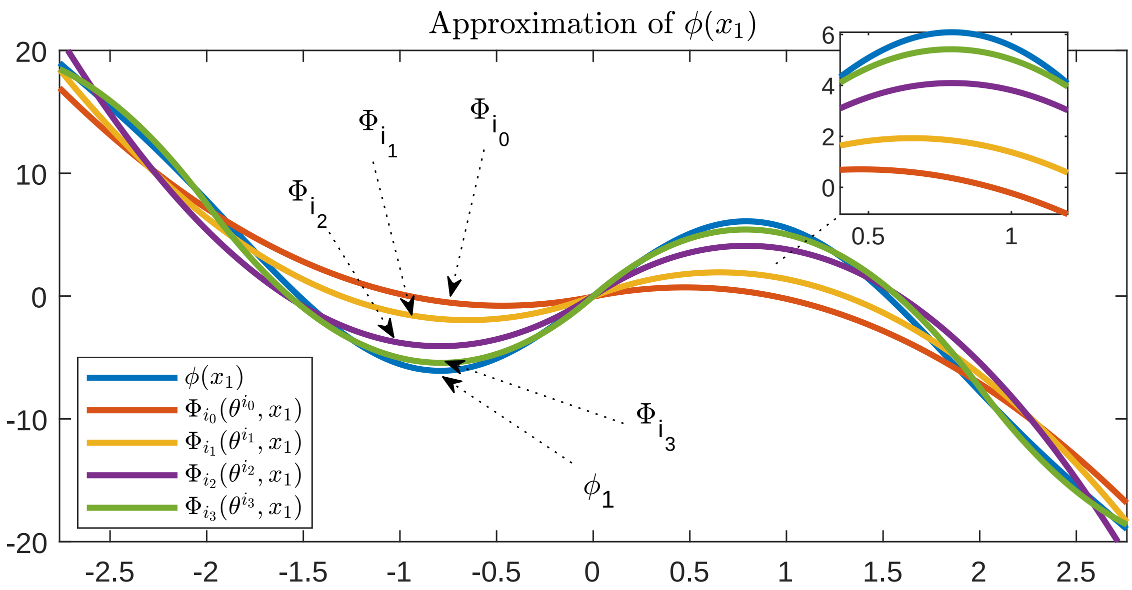

Figure 5.3: Approximation of $\phi(x_1)$ on the interval $[-2.7, 2.7]$ in which $x_1(t)$ is confined.

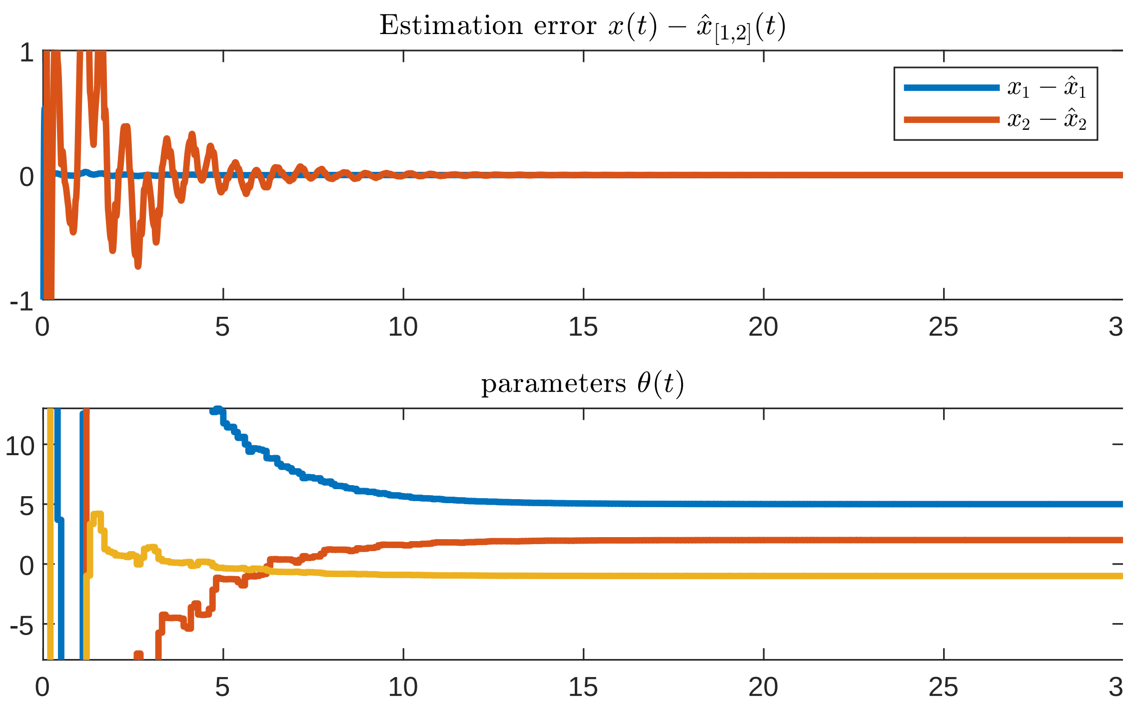

Figure 5.4: Estimation error and parameters evolution with a least square identifier.

at scale $i_3 = i_0 - 3 = 1$. The identifier samples and updates every $T = 0.1$ seconds. To ensure the PE property of each stage we set the regularization matrices as $\Omega_\ell = 10^{-4} I_{d_\ell}$ for $\ell = 0, \dots, 3$. The forgetting factors μ_ℓ have been set to $\mu_0 = 0.999$, $\mu_1 = 0.998$, $\mu_2 = 0.997$ and $\mu_3 = 0.996$. Figure 5.2 shows the time evolution of the norm of the estimation error $x(t) - \hat{x}_{[1,2]}(t)$ and of the outputs $\theta_\ell(t)$ of each stage. We underline how different learning dynamics arise naturally between two successive scales also with similar values of μ_ℓ. Figure 5.3 shows instead the approximation of the function $\phi(x_1)$ obtained with the prediction models at the different scales $i_\ell = 4, 3, 2, 1$. Each approximation is obtained by computing the scale-i_ℓ prediction model (4.47) with the most recent value of $\theta^{\mathbf{i}_\ell}$.

In the second simulation we suppose that all the uncertainty we have on ϕ is concentrated to the coefficients multiplying the terms $\sin(2x_1)$, x_1 and x_1^3. We thus use a single least square identifier to estimate the uncertain parameters, obtained by letting $\sigma(x_1) := \mathrm{col}(\sin(2x_1), x_1, x_1^3)$ in (4.10), $d = 3$ and $\mu = 0.9$ in (4.12) and $\Omega = 0$ in (4.13). Figure 5.4 depicts the time evolution of the estimation error $x - \hat{x}_{[1,2]}$ and of the parameters θ obtained with this identifier, showing ho, in this simpler case in which the uncertainty is concentrated in the parameters, an asymptotically exact state estimation is achieved.

Conclusion

This second part was dedicated to the development of a theory of adaptation in which adaptation is seen as a system identification problem, and approached in a hybrid system framework. Chapter 4 proposed a framework in which identification schemes can be described in a system theoretical envelope, and the underlying optimization properties can be characterized in terms of a strong stability requirement (the identifier requirement). Different cases of identifiers fitting in the framework have been presented, form least squares algorithms to nonlinear mini batch procedures and non-parametric wavelet decomposition.

Chapter 5 presented an approach to the problem of adaptive observers design that draws inspiration from the ideas of the literature of identification for control and in which adaptation is approached in the framework of Chapter 4. Differently from canonical adaptive observer designs, here we do not assume a particular structure of the uncertainty and we do not propose an ad hoc adaptation mechanism, rather we allow for different parametric and non-parametric system identification techniques to be applied. Differently from the works on identification for control, we propose here a more system theory oriented treatise of identifiers and we consider a nonlinear observation problem rather than a linear robust stabilization one.

From the viewpoint of observation, the results presented in this chapter are quite limited in scope, in the sense that they strongly rely on the structure of the high-gain observers. Nevertheless, the underlying idea of studying the identification schemes as hybrid systems and to use identifiers to perform adaptation in adaptive schemes represents an interesting approach that might lead to interesting extensions and that is worth investigate further.

The focus on observers is motivated by the similarities that observation theory shares with output regulation, and the material presented in these chapters represent also a first brick towards the development of a theory of adaptive output regulation that relies on system identification for adaptation. The application to output regulation of the ideas presented so far is indeed the subject of the forthcoming chapters.

Part III

Adaptive Output Regulation

6

Adaptive Output Regulation of Nonlinear Systems

In the first part of the thesis we discussed how a chicken-egg dilemma necessarily arises in the construction of nonlinear post-processing regulators. We highlighted how complementary pre-processing paradigms can be used to bypass the chicken-egg dilemma and to define regulators that do not suffer from the intertwining between the internal model unit and the stabilizer, at the price, however, of consistently reducing the class of systems that can be dealt with. On the other hand, we observed that also the existing post-processing solutions are far to give a definite answer to the problem and that, typically, in post-processing regulators the need of avoiding the chicken-egg dilemma yields restrictive assumptions on the exosystem or on the admissible steady state trajectories, and leads to sacrifice asymptotic regulation for an approximate result. Nevertheless, the nice intuition that can be drawn from the existing post-processing solutions, and that coincides with the interpretation given in sections 2.2 and 3.1 of the linear regulator, is that the internal model may be *fixed a priori* on the basis of the *expected class* of signals that it will have to generate at the steady state. The stabilizer is then fixed at a second phase to stabilize the overall system, with the (usually fragile) property that, if the ideal steady state of the internal model unit actually is in the expected class, then asymptotic regulation can be claimed.

In this chapter we propose a framework based on a post-processing structure leveraging on the aforementioned intuition, namely that *the internal model can be fixed in advance on the basis of the expected class of functions $\eta^\star$ resulting after considering all the possible plant's and exosystem's uncertainties and all the possible choices of the stabilizer inside a prescribed set*. We further observe that, from the internal model's viewpoint, the uncertainty on choice of the stabilizer (that is present at the moment in which the structure of the internal model is fixed) has the same effect of the uncertainties in the plant's or exosystem's model, as all of them will anyway result in potential deviations from the "expected class" of the actual ideal steady state that the internal model has to generate. This "expected class" (from now on denoted by $\mathcal{C}^\star_\eta$) of steady-state functions $\eta^\star$ results thus from an overall assessment about the knowledge of the plant and the exosystem, the expected uncertainty in their models and the expected set of stabilizers that will be adopted, all treated equally.

In the linear case $\mathcal{C}^\star_\eta$ coincides with the set of solutions of a system that includes the modes of the exosystem, i.e. all the systems generating the possible $\eta^\star \in \mathcal{C}^\star_\eta$ are immersed into the linear regulator (1.13). This "immersion assumption" is also what lies under *all* the nonlinear approaches claiming a certain degree of "robustness' (it is the case, for instance, of the "structurally stable" framework of (Byrnes et al., 1997a,b) and of all the subsequent nonlinear extensions of (Isidori et al., 2012; Forte et al., 2013; Bin et al., 2016), included the approach presented in Section 3.2). Instead of assuming that we know a system whose set of solutions include $\mathcal{C}^\star_\eta$, the idea is to use *adaptation* to take care of the overall uncertainty characterizing $\mathcal{C}^\star_\eta$, in the same way as we did in chapters 4 and 5. The key point of the proposed approach is that adaptation is cast as a *system identification* problem defined on the closed-loop system trajectories. This, indeed, permits us to "shift" the problem of dealing with the uncertainty of $\eta^\star$ to the identification phase where, however, we can rely on well-known identification algorithms that are naturally able to handle properly wide classes of signals, thus making their application a perfect fit.

As a necessary compromise, though, the proposed approach is structurally *approximate*, since from the identification viewpoint the assumption of the existence of a "true model" (and, by analogy, of asymptotic regulation) is quite pointless. Consistently, our main result aims to relate the performances on the regulation side with the performances of the corresponding identified model, expressed in terms of the *prediction error* evaluated along the ideal error-zeroing steady state. Asymptotic regulation, in turn, will follow only when a right model exists that is in the "range" of the identifier used, and in this sense we observe how this design philosophy matches with the properties mentioned before of the other post-processing approaches.

In Section 6.1 we propose a general framework where post-processing adaptive internal model-based regulators can be constructed for multivariable nonlinear systems. The treatise in Section 6.1, and the related results, are deliberately kept general enough to embrace a large spectrum of problems, at the cost of bordering the tautology. In Sections 6.2, 6.3 and 6.4 we present some design examples showing how the general guidelines of Section 6.1 can be applied to relevant classes of problems. In particular, we provide a systematic design procedure for non-square minimum-phase normal forms (Section 6.2), we show how additional non-vanishing outputs can be naturally handled in the framework (Sections 6.3.1 and 6.3.2) and we provide an adaptive solution for general multivariable linear systems (Section 6.4). For simplicity, all the results given here refer to continuous-time systems. We underline, however, that this is just for ease of exposition and all what is said here can be proved to hold if the identifier is a hybrid system of the kind fitting in the framework of Chapter 4. The content of this chapter was the subject of the paper (Bin and Marconi, 2018a), currently under review.

6.1 A Framework for Adaptive Regulation

In this section we deal with a general class of multivariable nonlinear systems of the form

$$\begin{aligned} \dot{x} &= f(w,x,u) \\ y &= h(w,x) \end{aligned} \tag{6.1}$$

with state $x \in \mathbb{R}^n$, control input $u \in \mathbb{R}^m$, measured outputs $y \in \mathbb{R}^q$ and with $w \in \mathbb{R}^{n_w}$ produced by an exosystem of the form

$$\dot{w} = s(w), \tag{6.2}$$

with initial conditions that are constrained to a compact invariant set $W \subset \mathbb{R}^{n_w}$. Associated to (6.1), there is a set of $p > 0$ *regulation errors* defined as

$$e = h_e(w, x) \tag{6.3}$$

with $h_e : W \times \mathbb{R}^n \to \mathbb{R}^p$. As customary, we assume e to belong to the set of measurable outputs, i.e. we suppose that $q \geq p$ and that $h(w, x) = \mathrm{col}(h_e(w, x), h_{\mathrm{m}}(w, x))$, where $y_{\mathrm{m}} = h_{\mathrm{m}}(w, x)$ represents some additional measurements that are not required to vanish in steady state.

We build the design procedure within the non-equilibrium framework of (Byrnes and Isidori, 2003) by assuming, for each w solution of (6.2) with $w(0) \in W$, the existence of a unique continuously differentiable functions $x^\star : \mathbb{R} \to \mathbb{R}^n$ and an integrable function $u^\star : \mathbb{R} \to \mathbb{R}^m$ solutions to the *regulator equations*

$$\begin{array}{rcl} \dot{w} & = & s(w) \\ \dot{x}^\star & = & f(w, x^\star, u^\star) \\ 0 & = & h_e(w, x^\star). \end{array} \tag{6.4}$$

In general, $(x^\star, u^\star)$ are uncertain and strongly dependent on the regulated dynamics. We thus aim to develop a design paradigm not substantially relying on their knowledge, by just assuming that the designer has some insight, better specified later, on the structure of (6.4) to be able to calibrate the regulator. The resulting framework leads necessarily to a regulation that is, in general, “approximate" with the asymptotic bound on the regulation error that is also related to the amount of information available on $(x^\star, u^\star)$.

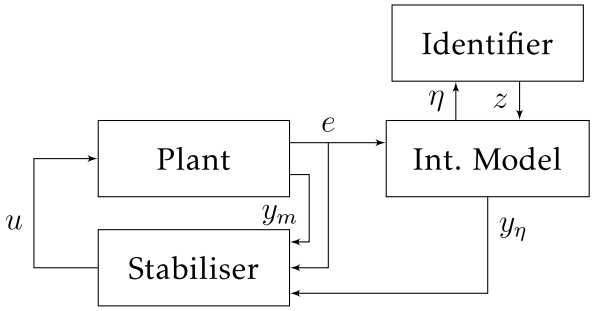

Figure 6.1: Block-diagram of the regulator.

6.1.1 The Control Structure

The proposed post-processing control structure is sketched in Figure 6.1. The internal model unit is a system of the form (compare with (1.13))

$$\dot{\eta} = \Phi(\eta, z) + Ge, \qquad \eta \in \mathbb{R}^{dp} \tag{6.5}$$

with a virtual output

$$y_\eta = \Gamma(\eta), \qquad y_\eta \in \mathbb{R}^{p_\eta} \tag{6.6}$$

where $d, p_\eta \in \mathbb{N}$, $\Gamma : \mathbb{R}^{pd} \to \mathbb{R}^{p_\eta}$ and $\Phi(\eta, z)$ and G have the following structure

$$\Phi(\eta, z) = \begin{pmatrix} 0 & I_p & 0 & \cdots & 0 \\ 0 & 0 & I_p & \cdots & 0 \\ . & . & . & \ddots & . \\ 0 & 0 & 0 & \cdots & I_p \\ & & \psi(\eta, z) & & \end{pmatrix}, \qquad G = \begin{pmatrix} G_1 \\ G_2 \\ \vdots \\ G_d \end{pmatrix}$$

with $\psi : \mathbb{R}^{dp} \times \mathcal{Z} \to \mathbb{R}^{p \times dp}$ and $G_i \in \mathbb{R}^{p \times p}$, $i = 1, \ldots, d$, being $\mathcal{Z}$ a finite-dimensional normed vector space. In referring to the state η of (6.5), we will often use the partition $\eta = (\eta_1, \ldots, \eta_d)$, with $\eta_i \in \mathbb{R}^p$. The internal model is parametrized by z that is the state of the *identifier* described by

$$\dot{z} = \mu(z, \eta), \tag{6.7}$$

in which $z : \mathcal{Z} \times \mathbb{R}^{dp} \to \mathcal{Z}$, and whose role is detailed later. Finally, the stabilizer is a system of the form

$$\begin{aligned} \dot{\xi} &= \varphi(\xi, y, y_\eta) \\ u &= \gamma(\xi, y, y_\eta), \end{aligned} \tag{6.8}$$

with $\xi \in \mathbb{R}^{n_\xi}$, $n_\xi \in \mathbb{N}$, $\varphi : \mathbb{R}^{n_\xi} \times \mathbb{R}^q \times \mathbb{R}^{p_\eta} \to \mathbb{R}^{n_\xi}$ and $\gamma : \mathbb{R}^{n_\xi} \times \mathbb{R}^q \times \mathbb{R}^{p_\eta} \to \mathbb{R}^m$.

The specific choice of the previous systems will be detailed in the next sections. For the time being, we just assume that the stabilizer and the internal model, regardless their specific design, fulfill a *steady state left-invertibility* condition. As for (6.8), in particular, we assume that for each $w : \mathbb{R} \to W$, $x^\star : \mathbb{R} \to \mathbb{R}^n$ and $u^\star : \mathbb{R} \to \mathbb{R}^m$ solution of the regulator equation (6.4), and with $y^\star := h(w, x^\star)$, there exist unique $y_\eta^\star : \mathbb{R} \to \mathbb{R}^{p_\eta}$ and $\xi^\star : \mathbb{R} \to \mathbb{R}^{n_\xi}$ solution of

$$\begin{aligned} \dot{\xi}^\star &= \varphi(\xi^\star, y^\star, y_\eta^\star) \\ u^\star &= \gamma(\xi^\star, y^\star, y_\eta^\star)\,. \end{aligned} \tag{6.9}$$

Similarly, as far as the internal model is concerned, we assume that, given $y_\eta^\star$, there exist a unique $\eta^\star : \mathbb{R} \to \mathbb{R}^{pd}$ fulfilling

$$y_\eta^\star = \Gamma(\eta^\star), \qquad\qquad \eta_i^\star = \dot{\eta}_{i-1}^\star, \quad i = 2, \dots, d\,. \tag{6.10}$$

These left-invertibility assumptions guarantee the existence of an ideal steady state $(\xi^\star, \eta^\star)$ for the stabilizer and the internal model that is compatible with the regulation requirement $e = 0$. Consistently with the linear case, we observe that, in principle, the previous equations could be solved with $y_\eta^\star = 0$ and $\eta^\star = 0$, namely with a vanishing steady state contribution of the internal model. It is worth also remarking that $(\xi^\star, \eta^\star)$ clearly depend on $(\varphi, \gamma, \Gamma)$ and, also with $(\varphi, \gamma, \Gamma)$ known and fixed, they are, in general, unknown functions as so are $(x^\star, u^\star)$.

For $\eta^\star$ to be a (steady-state) trajectory of the internal model in the closed-loop structure, the function $\psi(\cdot,\cdot)$ and $\mu(\cdot,\cdot)$ in (6.5), (6.7) should be ideally chosen so that

$$\dot{\eta}_d^\star = \psi(\eta^\star, z^\star), \tag{6.11}$$

for some $z^\star : \mathbb{R} \to \mathcal{Z}$ solution of

$$\dot{z}^\star = \mu(z^\star, \eta^\star)\,.$$

This, in fact, would guarantee that $(x^\star(t), \xi^\star(t), \eta^\star(t), z^\star(t))$ is a trajectory of the closed-loop system associated to an identically zero regulation error. The design of ψ and μ along this direction, however, hides the chicken egg-dilemma, as their design is clearly affected by $\eta^\star$, which depends on $(\varphi, \gamma, \Gamma)$ that, in turn, depend on (ψ, μ) themselves.

6.1.2 A "Class-Type" Internal Model

Our design strategy pivots around the idea that (6.11) can be seen as a "prediction model" relating the "next derivative" $\dot{\eta}_d^\star$ to the "previous derivatives" $\eta^\star$ of the ideal steady state of the internal model, and that the design the identifier (6.7) can be cast as an identification problem aimed to find the model that fits at best those signals. Approaching the problem in this way, though, hides a number of problems. First, the signal $\eta^\star$ involved in (6.11) is not known and thus it is not clear on which data the identifier should work with. This problem will be tackled in the next section by feeding the identifier with η as proxy variable of $\eta^\star$. Furthermore, even if $\eta^\star$ were known, the "next" derivative $\dot{\eta}^\star$ is not available and there is not a clear way of expressing it as combination of known state variables without leading to an algebraic loop. This issue will be tackled by setting up the identification problem not on ψ, but rather "one integrator away". More in details, we consider an auxiliary system with state $\bar{\eta} \in \mathbb{R}^{(d-1)p}$ (i.e. reduced by a block of p components with respect to (6.5)), reading as

$$\begin{array}{rcll} \dot{\bar{\eta}}_i & = & \bar{\eta}_{i+1} + G_i e, & i = 1, \dots, d-2 \\ \dot{\bar{\eta}}_{d-1} & = & \phi(\bar{\eta}, \theta) + G_{d-1} e & \end{array} \tag{6.12}$$

where the matrices G_i are the same as in (6.5), $\phi : \mathbb{R}^{p(d-1)} \times \Theta \to \mathbb{R}^p$ is a C^1 function to be fixed (with Θ a normed vector space of finite dimension), and

$$\theta = \omega(z) \tag{6.13}$$

is a virtual output associated to the identifier (6.7), defined by a C^1 map $\omega : \mathcal{Z} \to \Theta$ to be fixed. The analysis of the previous section shows that, if (6.12) were used in place of (6.5), then $(x^\star, \xi^\star, \bar{\eta}^\star, z^\star)$, in which $\bar{\eta}^\star = \eta^\star_{[1,d-1]}$ where $\eta^\star$ is the same as in (6.10), would be a trajectory of the closed-loop system associated to a regulation error identically zero *provided that*

$$\eta_d^\star(t) = \phi(\eta_{[1,d-1]}^\star(t), \theta^\star(t)) \qquad \forall\, t \geq 0 \tag{6.14}$$

for some $\theta^\star : \mathbb{R} \to \Theta$ such that $\theta^\star = \omega(z^\star)$. The idea that is followed is then to define the identification problem on the equation (6.14), instead of (6.11), and to design the identifier (6.7), (6.13) to produce the $\theta^\star$ that guarantees the "best" attainable prediction of $\eta_d^\star$ on the basis of $\eta_{[1,d-1]}^\star$. The clear advantage of (6.14) over (6.11), in fact, is that the former does not involve the knowledge of $\dot{\eta}^\star$. The design of the internal model unit (6.5) is then completed by defining ψ as

$$\psi(\eta, z) = \frac{\partial \phi(\eta_{[1,d-1]}, \theta)}{\partial \eta_{[1,d-1]}} \eta_{[2,d]} + \frac{\partial \phi(\eta_{[1,d-1]}, \theta)}{\partial \theta} \frac{\partial \omega(z)}{\partial z} \mu(z, \eta) \tag{6.15}$$

so that $t \mapsto \psi(\eta(t), z(t))$ equals the time derivative of $t \mapsto \phi(\eta_{[1,d-1]}(t), \theta(t))$. In this way, indeed, the virtual system (6.12) is immersed in the implemented internal model (6.5) that, thus, can generate each of its solutions.

Along this direction, and borrowing the notation typically adopted in the identification literature (Ljung, 1999), we refer to the map $\phi(\cdot, \theta)$ as the *prediction model* relating the "input data" $\eta_{[1,d-1]}^\star$ to the "output" $\eta_d^\star$, and to the set $\mathcal{M} := \{\phi(\cdot, \theta) \ : \ \theta \in \Theta\}$ of all the obtainable candidate models as the corresponding *model set*. The selection of $(d, \mathcal{M})$, in turn, is where the "chicken egg-dilemma" arises, as it clearly relies on the a priori information about the class $\mathcal{C}_\eta^\star$ of the signals $\eta^\star$ which, in turn, depends on the choice of the stabilizer. To break the loop and solve the dilemma, here we assume a priori that $\eta^\star$ belongs to a class $\mathcal{C}_\eta^\star$ of functions, so as we can fix $(d, \mathcal{M})$ and the adaptation algorithm. This will result in a "class-type" internal model, which can lead to asymptotic regulation only if a "real model" relating $\eta_d^\star$ and $\eta_{[1,d-1]}^\star$ exists and lies in $\mathcal{M}$. Possible examples of the class $\mathcal{C}_\eta^\star$ are linear functions of $w(t)$ (which is the case of linear systems, see Remark 2), or polynomial functions of $w(t)$, or simply integrable/differentiable functions of time. In a general multivariable nonlinear context the fact that d and $\mathcal{M}$ can be taken so that (6.14) is fulfilled for some member $\phi(\cdot, \theta)$ of $\mathcal{M}$ is hard to be assumed due to the uncertainties in the $(x^\star, u^\star)$, highly uncertain, and in the stabilizer, yet to be fixed. Furthermore, even in the fortunate case in which the ideal relation (6.14) can be fulfilled, this could require an unacceptable complexity, and an approximated model with a possibly lower d would be preferable. Overall, the a-priori guess of the class C_η, of d and $\mathcal{M}$ is where the knowledge of the system and "the touch" of the designer come into play, as better highlighted in Section 6.1.5.

6.1.3 The Design of the Identifier

By the previous section we assume that the dimension d of the internal model and the model set $\mathcal{M}$ are fixed, and we shift our attention on the design of the identifier. We introduce the *"prediction error"* corresponding to the model $\phi(\cdot,\theta)$ as

$$\varepsilon(t,\theta) := \eta_d^\star(t) - \phi(\eta_{[1,d-1]}^\star(t),\theta)$$

and we look for an identifier (6.7), (6.13) able to select the best θ, say $\theta^\star$, whose corresponding model $\phi(\cdot,\theta^\star(t))$ is, at each t, the "best" model in $\mathcal{M}$ relating $\eta_d^\star$ and $\eta_{[1,d-1]}^\star$, minimizing in some sense ε. As customary in system identification, the meaning of "best" in the model selection is based on the definition of a *fitness criteria* assigning to each model $\phi(\cdot,\theta) \in \mathcal{M}$ a suitable and comparable value. In particular, with $C^0(\Theta,\mathbb{R}_+)$ the space of continuous functions $\Theta \to \mathbb{R}_+$, to each function $\eta^\star : \mathbb{R} \to \mathbb{R}^{pd}$ we associate the map $\mathcal{J}_{\eta^\star} : \mathbb{R}_+ \to C^0(\Theta,\mathbb{R}_+)$ given by

$$\big(\mathcal{J}_{\eta^\star}(\theta)\big)(t) := \int_0^t c_\varepsilon\Big(t,s,|\varepsilon(s,\theta)|\Big)ds + c_r(\theta), \qquad (6.16)$$

with $c_\varepsilon : \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ and $c_r : \Theta \to \mathbb{R}_+$ some user-defined positive functions characterizing the particular overlying identification problem[1]. To $\mathcal{J}_{\eta^\star}$ we associate the set-valued map $\vartheta_{\eta^\star}^\circ : \mathbb{R}_+ \rightrightarrows \Theta$ defined as

$$\vartheta_{\eta^\star}^\circ(t) := \operatorname*{argmin}_{\theta\in\Theta} (\mathcal{J}_{\eta^\star}(\theta))(t),$$

and we introduce the following requirement.

Requirement 6.1 (Identifier basic requirement)**.** *The identifier* (6.7) *is said to satisfy the* identifier basic requirement *if for any integrable function $\eta^\star : \mathbb{R} \to \mathbb{R}^d$ there exists a unique function $z^\star : \mathbb{R} \to \mathcal{Z}$ solution of*

$$\dot z^\star = \mu(z^\star,\eta^\star), \qquad z(0) = z^\star(0)$$

[1]More precisely, the integral term of (6.16) measures how well a given choice of θ fits the historical data, while $c_r(\theta)$ plays the role of a regularization factor.

such that, with $\theta^\star(t) := \omega(z^\star(t))$, the following holds

$$\theta^\star(t) \in \vartheta^\circ_{\eta^\star}(t).$$

The identifier basic requirement represents the elementary property that (6.7) must have to be consistent with the system identification viewpoint of the problem and with the underlying optimization characterization. Further stability properties can be added to the requirement if needed; in Section 6.2, for instance, we will ask an additional strong stability property of the ideal steady-state $z^\star$ expressed in terms of an input-to-state stability (ISS) requirement. It is worth noting that the identifier basic requirement is implied by the identifier requirement of Section 4.1 (Requirement 4.1), namely each identifier fitting in the framework of Chapter 4 fulfills the identifier basic requirement.

6.1.4 Selection of the Stabiliser and a Structural Result

The previous sections left open the design of the "innovation terms" G_i, $i = 1, \dots, d$, of the internal model and of the stabilizer, while identifying a class of possible ideal steady states $(x^\star, \xi^\star, \eta^\star, z^\star)$ associated to the dynamic blocks of Figure 6.1, with the last three functions that are still floating, as dependent on the particular instance of the stabilizer. We observe that, in general, this ideal steady state *is not* a trajectory of the closed-loop system due to the mismatch between $\eta^\star_d$ and $\phi(\eta^\star_{[1,d-1]}, \theta^\star)$. In this respect it seems reasonable to look for a choice of G_i and of the stabilizer to steer the closed-loop trajectories "close" to the above steady state, where "how close" is related to the optimal value of the prediction error. Towards this end, change variables as $(x, \xi, \eta, z) \mapsto (\tilde{x}, \tilde{\xi}, \tilde{z}, \tilde{\eta})$, with

$$\begin{aligned} &\tilde{x} := x - x^\star, \quad \tilde{\xi} := \xi - \xi^\star, \\ &\tilde{z} := z - z^\star, \quad \tilde{\eta} := \eta - \mathrm{col}\left(\eta^\star_{[1,d-1]}, \phi(\eta^\star_{[1,d-1]}, \theta^\star)\right) \end{aligned} \tag{6.17}$$

so that, by letting

$$\varepsilon^\star := \eta^\star_d - \phi(\eta^\star_{[1,d-1]}, \theta^\star)$$

be the optimal prediction error achieved by the identifier, system (6.1), (6.5), (6.7) and (6.8) in new coordinates read as

$$\begin{aligned} \dot{\tilde{x}} &= \tilde{f}(w, \tilde{x}, \tilde{\xi}, \tilde{\eta}, \eta^\star, \xi^\star) \\ \dot{\tilde{\xi}} &= \tilde{\varphi}(w, \tilde{x}, \tilde{\xi}, \tilde{\eta}, \eta^\star, \xi^\star) \end{aligned} \tag{6.18}$$

and

$$\begin{aligned} \dot{\tilde{\eta}}_i &= \tilde{\eta}_{i+1} + G_i e \qquad\qquad i = 1, \ldots, d-2 \\ \dot{\tilde{\eta}}_{d-1} &= \tilde{\eta}_d + G_{d-1} e + \varepsilon^\star \\ \dot{\tilde{\eta}}_d &= \tilde{\psi}(\tilde{\eta}, \tilde{z}, \eta^\star, z^\star, \varepsilon^\star) + G_d e \end{aligned} \tag{6.19}$$

and

$$\dot{\tilde{z}} = \tilde{\mu}(\tilde{z}, \tilde{\eta}, \eta^\star, z^\star, \varepsilon^\star), \tag{6.20}$$

where, with $E := \mathrm{col}(0_p, \ldots, 0_p, -I_p)$, we let

$$\begin{aligned} &\tilde{\psi}(\cdot) := \psi(\tilde{\eta} + \eta^\star + E\varepsilon^\star, \tilde{z} + z^\star) - \dot{\phi}(\eta^\star, z^\star) \\ &\tilde{\mu}(\cdot) := \mu(\tilde{z} + z^\star, \tilde{\eta} + \eta^\star + E\varepsilon^\star) - \mu(z^\star, \eta^\star) \\ &\tilde{f}(\cdot) := f(w, \tilde{x} + x^\star, \gamma(\tilde{\xi} + \xi^\star, h(w, \tilde{x} + x^\star), \tilde{\eta}_1 + \eta_1^\star)) - f(w, x^\star, u^\star) \\ &\tilde{\varphi}(\cdot) := \varphi(\tilde{\xi} + \xi^\star, h(w, \tilde{x} + x^\star), \tilde{\eta}_1 + \eta_1^\star) \ - \varphi(\xi^\star, y^\star, \eta_1^\star) \end{aligned}$$

and $e = \tilde{h}_e(w, \tilde{x}) := h_e(w, \tilde{x} + x^\star)$. As emphasized by the notation, we observe that the system dynamics depends on $(\xi^\star, \eta^\star, z^\star)$ which is floating with the stabilizer. However, we also observe that, by using (6.4) and (6.15), all the previous functions are vanishing at $(\tilde{x}, \tilde{\xi}, \tilde{\eta}, \tilde{z}) = (0, 0, 0, 0)$ and $\varepsilon^\star = 0$ for all $(\xi^\star, \eta^\star, z^\star)$. This system is thus regarded as a system with state $(\tilde{x}, \tilde{\xi}, \tilde{\eta}, \tilde{z})$ perturbed by the input $\varepsilon^\star$. An ISS property with respect to the disturbance $\varepsilon^\star$ is thus the natural requirement for the design of the G_i's and of the stabilizer. For compactness we let $\tilde{\mathbf{x}} := \mathrm{col}(\tilde{x}, \tilde{\xi}, \tilde{\eta}, \tilde{z})$.

Requirement 6.2 (Stability requirement)**.** *We say that the stabilizer* (6.8) *and the matrices G_i of* (6.5) *satisfy the* stability requirement *if system* (6.18), (6.19), (6.20) *is practically ISS with respect to the input $\varepsilon^\star$ with possible restrictions on the initial conditions. Namely, there exist a set $\mathcal{O} \subseteq \mathbb{R}^n \times \mathbb{R}^{n_\xi} \times \mathbb{R}^{dp} \times \mathcal{Z}$, functions $\beta_{\mathrm{s}} \in \mathcal{KL}$ and $\rho_{\mathrm{s}} \in \mathcal{K}$, and a positive ν_{s} such that for all initial conditions* $(x(0), \xi(0), \eta(0), z(0)) \in \mathcal{O}$

the trajectories of (6.18), (6.19), (6.20) *satisfy*

$$|\tilde{\mathbf{x}}(t)| \le \beta_{\rm s}(|\tilde{\mathbf{x}}(0)|, t) + \rho_{\rm s}(|\varepsilon^\star(t)|_{[0,t)}) + \nu_{\rm s}\,.$$

Constructive designs that show how this requirement can be fulfilled are postponed to the next sections, while the following theorem formalizes a direct consequence of two requirements.

Theorem 6.1. *Let the internal model, the identifier and the stabilizer be designed to fulfill the requirements 6.1 and 6.2. Then there exists a class-$\mathcal{K}$ function $\rho_{\rm e}$ such that the closed-loop trajectories originating from $\mathcal{O}$ are bounded and*

$$\limsup_{t\to\infty} |e(t)| \le \rho_{\rm e}\Big(\limsup_{t\to\infty} |\varepsilon^\star(t)| + \nu_{\rm s}\Big)$$

with $\mathcal{O}$ and $\nu_{\rm s}$ introduced in Definition 6.2.

Theorem 6.1 states that a regulator constructed to satisfy the two requirements structurally achieves approximate regulation. The proof of the theorem is a straightforward consequence of regularity of the function defining the plant's data and it is thus omitted. The requirement and the theorem are deliberately formulated in a quite general way, as their aim is just to formalize the structural properties of the regulators constructed by following the procedure detailed above. The presence of $\nu_{\rm s}$ in the requirement makes it equivalent to a requirement of ultimate uniform boundedness of the trajectories resulting, by continuity, in the expected ultimate bounded of Theorem 6.1 on the error. The constant $\nu_{\rm s}$ was introduced to fit in the framework also "practical" stabilizers that, even if the internal model and identifier are able to attain $\varepsilon^\star = 0$, thus making $(\xi^\star, \eta^\star, z^\star)$ a possible trajectory of the system, might not be able to steer the closed-loop system to that ideal steady state. The performance of the regulator is thus given by two factors: the performance of the internal model/identifier (responsible for $\varepsilon^\star$) and those of the stabilizer (responsible for $\nu_{\rm s}$). Asymptotic regulation is achieved if the identifier is able to reach a null prediction error and the stabilizer is able to make the ideal state state $(\xi^\star, \eta^\star, z^\star)$ attractive. The general requirement of Definition 6.2 will be supported in Section 6.2 by a constructive high-gain design paradigm for the class of minimum-phase systems.

6.1.5 Remarks on the Framework

We underline how the a priori knowledge on the system and its possible variations play a crucial role in the whole procedure, especially in the design guidelines of the internal model unit and the identifier detailed in Section 6.1.2 and 6.1.3. As a matter of fact, the more one knows about $(x^\star, u^\star)$ and the stabilizer, the more knows about the class $\mathcal{C}_\eta^\star$, the better is the attainable prediction error and thus, according to Theorem 6.1, the lower is the steady state regulation error. The selection of ϕ, in turn, comes from an overall assessment about the plant, the exosystem, and the stabilizer without indeed relying on the perfect knowledge of none of those systems. We emphasize how, in the actual perspective, also the perfect knowledge of the exosystem dynamics loses importance in the design of the internal model as, indeed, it is the knowledge of the whole system that is necessary to extract information about the class $\mathcal{C}_\eta^\star$ to which $\eta^\star$ shall belong to. In this respect, it is worth remarking that the lack of knowledge of $\eta^\star$ is tightly connected to the chicken-egg dilemma, since $\eta^\star$ depends on the stabilizer not yet fixed, and to the need of designing regulators that are not too "friend-centric" according to the discussion in Section 2.2. We further emphasize that the "touch" of the designer and her/his knowledge about the overall system (and potential stabilizers) play a crucial role in identifying the right class $\mathcal{C}_\eta^\star$ and consequently fixing the right model set for ϕ. We also emphasize that the same dimension d of the internal model is a crucial degree-of-freedom that can be played at this stage, by trading off between asymptotic regulation requirements, which would suggest large values of d, and issues related to the implementation and the reduction of the complexity, typically pushing for low values of d.

6.2 A High-Gain Strategy for the Stability Requirement

6.2.1 Design of the Adaptive Internal Model Unit

In this section we present a high-gain strategy for the design of the matrices G_i and of the stabilizer to fulfill the ISS property of Theorem 6.1. We approach the problem by first studying the interconnection (6.19)-(6.20) between the internal model and the identifier, seen as a system with state $(\tilde{\eta}, \tilde{z})$ and with inputs $e = \tilde{h}_e(w, \tilde{x})$ and $\varepsilon^\star$. As a first step we reinforce the identifier basic requirement by asking an additional strong stability property of the ideal steady-state $z^\star$.

Requirement 6.3 (Identifier strong stability requirement)**.** *The identifier* (6.7) *is said to satisfy the* identifier strong stability requirement *if it satisfies the identifier basic requirement and, in addition, there exist $\beta_z \in \mathcal{KL}$ and $\rho_z \in \mathcal{K}$ such that, for every integrable $\delta : \mathbb{R} \to \mathbb{R}^{pd}$, all the solutions of the system* (6.7) *with input $\eta^\star + \delta$ satisfy*

$$|z(t) - z^\star(t)| \le \beta_z(|z(0) - z^\star(0)|, t) + \rho_z(|\delta|_{[0,t)})$$

for all $t \in \mathbb{R}_+$.

We observe how, for continuous-time identifiers, the identifier strong stability requirement is essentially the same as the Requirement 4.1 of Chapter 4. Hence, again, every identifier that fits in the framework of Section 4.1 also satisfies the identifier strong stability requirement.

The next Lemma claims that for an appropriate choice of the quantities G_i, and with $\psi(\cdot, \cdot)$ fulfilling certain properties, the system (6.19)-(6.20) can be robustly stabilized by the input-output pair $(e, \tilde{\eta}_1)$. The design of the G_i is done following a standard high-gain paradigm (Gauthier and Kupka, 2001; Byrnes and Isidori, 2004) by letting

$$G_i := h_i g^i I_p \,, \qquad i = 1, \dots, d \tag{6.21}$$

where $g \ge 1$ is a tuning parameter to be fixed and $h_1, \dots, h_d \in \mathbb{R}$ are such that the roots of the polynomial $p(s) := s^d + h_1 s^{d-1} + \dots + h_{d-1} s + h_d$ have negative real part.

Lemma 6.1. *Suppose that there exists a $L_\psi > 0$ such that*

$$|\tilde{\psi}(\tilde{\eta}, \tilde{z}, \eta^\star, z^\star, \varepsilon^\star)| \le L_\psi\big(|\tilde{\eta}| + |\tilde{z}| + |\varepsilon^\star|\big)$$

for all $(\eta^\star, z^\star)$ and $(\tilde{\eta}, \tilde{z}, \eta^\star)$. Assume that the identifier fulfills the identifier strong stability requirement. Then there exist $g^\star > 0$, $\rho_{\rm i} > 0$, $\pi_{\rm i} > 0$ and $\beta_{\rm i} \in \mathcal{KL}$, such that for all $g \ge g^\star$ the trajectories of the system (6.19)-(6.20) *with the G_i's fixed as in*

(6.21) *and with the input* e *chosen as*

$$e = \tilde{e} - \tilde{\eta}_1 \tag{6.22}$$

where $\tilde{e}$ *is an auxiliary input, satisfy*

$$|(\tilde{z}(t), \tilde{\eta}(t))| \leq \beta_{\rm i}(|(\tilde{z}(0), \tilde{\eta}(0))|, t) + \rho_{\rm i}|\tilde{e}|_{[0,t)} + \pi_{\rm i}|\varepsilon^\star|_{[0,t)}$$

for all $t \in \mathbb{R}_+$.

The proof of the lemma follows from the results of Section 5.1 and it is thus omitted. The high value of g is chosen in order to decrease the asymptotic gain between the input $\tilde{z}$ and the state $\tilde{\eta}$ of the system (6.19) and thus to impose a small gain condition in the interconnection with (6.20). Furthermore, we observe that the globally Lipschitz property of $\tilde{\psi}$ required in the Lemma can be simply obtained by assuming a locally Lipschitz property and saturating the right-hand side of (6.15) with a saturation level fixed according to the set where $\dot{\phi}(\eta^\star, z^\star)$ is expected to range. This, however, requires an a priori knowledge of bounds of $(\eta^\star, z^\star)$, which can be seen as *quantitative* manifestation of the chicken-egg dilemma. We now shift the attention to system (6.18), regarded as a system with input $\tilde{\eta}$ and, with an eye to (6.22), output

$$\tilde{e} = h_e(w, \tilde{x} + x^\star) + \tilde{\eta}_1 \,.$$

The overall closed-loop system is thus given by the interconnection of system (6.19)-(6.20), with input $(\tilde{e}, \varepsilon^\star)$ and output $\tilde{\eta}$, and system (6.18), with input $\tilde{\eta}$ and output $\tilde{e}$. It comes thus natural to design the stabilizer to induce a small-gain condition in the aforementioned interconnection. In this direction we state the forthcoming proposition, where we make reference to a set $\widetilde{X}_0 \times \widetilde{\Xi}_0$ of initial conditions for $(\tilde{x}, \tilde{\xi})$, in order to take into account local or semiglobal (relatively to the error-zeroing manifold) contexts. The sets $\widetilde{X}_0$ and $\widetilde{\Xi}_0$ are obtained by first fixing a set of initial conditions for (x, ξ) of the form $X_0 \times \Xi_0$, and then taking $\widetilde{X}_0 \times \widetilde{\Xi}_0$ be the union of all the points of the form $(x(0) - x^\star(0), \xi(0) - \xi^\star(0))$ where $(x(0), \xi(0)) \in X_0 \times \Xi_0$ and $(x^\star(0), \xi^\star(0))$ are obtained by any pair of functions $(x^\star, \xi^\star)$ coming from the regulator equations (6.4) and the invertibility condition (6.9), with $w(\cdot)$ that ranges in the set of the solutions of the exosystem originating in W.

Proposition 6.1. *Let the matrices G_i be fixed according to Lemma 6.1 and let $X_0 \times \Xi_0 \subset \mathbb{R}^n \times \mathbb{R}^{n_\xi}$. Suppose that the stabilizer is fixed so that the trajectories of* (6.18) *originating from $\widetilde{X}_0 \times \widetilde{\Xi}_0$ satisfy, for all $t \in \mathbb{R}_+$, the practical ISS condition*

$$|(\tilde{x}(t), \tilde{\xi}(t))| \le \beta_{\rm s}'\big(|(\tilde{x}(0), \tilde{\xi}(0))|, t) + \rho_{\rm s}'(|\tilde{\eta}|_{[0,t)}\big) + \nu$$

for some $\beta_{\rm s} \in \mathcal{KL}$, $\rho_{\rm s}' \in \mathcal{K}$, and positive ν, and, moreover,

$$\limsup_{t\to\infty} |\tilde{e}(t)| \le \rho_{\rm s}'' \limsup_{t\to\infty} |\tilde{\eta}(t)| + \nu$$

for some positive $\rho_{\rm s}''$ such that $\rho_{\rm s}''\rho_{\rm i} < 1$. Then, the stabilizer and the matrices G_i fulfil the stability requirement with $\mathcal{O} = X_0 \times \Xi_0 \times \mathbb{R}^{dp} \times \mathcal{Z}$.

The proof of the theorem follows by classical small gain arguments and it thus omitted.

6.2.2 Design of a Stabilizer for Minimum-Phase Normal Forms

In this section we show how for the class of minimum-phase systems that possess a normal form a stabilizer can be constructed to fulfil the assumptions of Proposition 6.1. We consider a subclass of systems (6.1) with state $x = \mathrm{col}(x_0, \chi, \zeta)$ satisfying the following equations

$$\dot{x}_0 = f_0(w, x) + b(w, x)u \tag{6.23a}$$
$$\dot{\chi} = F\chi + H\zeta \tag{6.23b}$$
$$\dot{\zeta} = q(w, x) + \Omega u \tag{6.23c}$$

and with regulation errors given by

$$e = C\chi\,,$$

where $\zeta \in \mathbb{R}^p$, $\chi = \mathrm{col}(\chi^1, \dots, \chi^p)$ with $\chi^i \in \mathbb{R}^{n_\chi^i}$, $i = 1, \dots, p$ (where $n_\chi^1 + \dots + n_\chi^p = n_\chi$), $F := \mathrm{diag}(F_1, \dots, F_p) \in \mathbb{R}^{n_\chi \times n_\chi}$ and $H := \mathrm{diag}(H_1, \dots, H_p) \in \mathbb{R}^{n_\chi \times p}$ with

$$F_i := \begin{pmatrix} 0_{n_\chi^i - 1} & I_{n_\chi^i - 1} \\ 0 & 0_{1\times(n_\chi^i - 1)} \end{pmatrix}, \qquad H_i := \begin{pmatrix} 0_{(n_\chi^i - 1)\times 1} \\ 1 \end{pmatrix}.$$

and $C := \mathrm{diag}(C_1, \dots, C_p)$, with

$$C_i := \begin{pmatrix} 1 & 0_{1\times(n_\chi^i - 1)} \end{pmatrix} .$$

The χ subsystem is described by p chain of integrators with ζ entering on the last equation and regulation errors given by the first component of each chain χ^i. The control input u takes values in $\mathbb{R}^m$, with $m \geq p$, the functions f_0, b and q are sufficiently smooth functions and $\Omega \in \mathbb{R}^{p\times m}$, denoting the so-called "high-frequency matrix", is full row-rank. The form (6.23) is indeed representative of many frameworks addressed in literature. For instance, systems having a well-defined vector relative degree with respect to the input-output pair (u, e) and admitting a *canonical normal form* fit in the proposed framework. In this case the x_0 dynamics in (6.23) does not depend on u and, when $e = 0$, it represents the *zero dynamics* of the system relative to the indicated input-output pair. On the other hand (6.23), with a different structure of χ and of the matrices F and H, is also representative of systems that are "just" (globally) strongly invertible in the sense of (Hirschorn, 1979; Singh, 1981) and feedback linearizable with respect to the input-output pair (u, e) and, as such, can be transformed in *partial normal form*, see (Wang et al., 2015a). In this case the dynamics (6.23b)-(6.23c) are the partial normal form of the systems and the subsystem (6.23a) is indeed the whole systems (i.e. $x = x_0$). In the following we assume that $y_{\mathrm{m}} = \mathrm{col}(\chi, \zeta)$, namely, as χ and ζ are linear combinations of the error and its time derivatives, and we look for a *partial state feedback* solution. A pure error feedback regulator only processing e can be obtained by replacing the time derivatives with appropriate estimates by using state standard high-gain techniques (see Teel and Praly, 1995) not here presented. We consider system (6.2), (6.23) under the following assumptions.

Assumption 6.1. *For each solution $w : \mathbb{R} \to W$ to (6.2) with $w(0) \in W$, there exist $x_0^\star : \mathbb{R}_+ \to \mathbb{R}^n$ and $u^\star : \mathbb{R}_+ \to \mathbb{R}^m$ such that, with $x^\star(t) := (x_0^\star(t), 0, 0)$, the following hold*

$$\begin{aligned} \dot{x}_0^\star &= f_0(w, x^\star) + b(w, x^\star)u^\star \\ 0 &= q(w, x^\star) + \Omega u^\star . \end{aligned}$$

Assumption 6.2. *There exist a locally Lipschitz $\rho_0 \in \mathcal{K}$ and $\beta_0 \in \mathcal{KL}$, such that, for any solution $(w, x) : \mathbb{R}_+ \to W \times \mathbb{R}^n$ to (6.2), (6.23) with $w(0) \in W$, the following* *estimate holds*

$$|x_0(t) - x_0^\star(t)| \leq \beta_0\big(|x_0(0) - x_0^\star(0)|, t\big) + \rho_0\big(|(\chi, \zeta)|_{[0,t)}\big)$$

for all $t \geq 0$ and for all locally bounded $u(\cdot)$.

Assumption 6.3. *There exists a full-rank matrix $\mathcal{L} \in \mathbb{R}^{m\times p}$ satisfying*

$$\mathcal{L}^T \Omega^T + \Omega \mathcal{L} \geq I_p \,.$$

The equations of $(x^\star, u^\star)$ in Assumption 6.1 are the specialization of the regulator equations (6.4) to this particular class of systems, Assumption 6.1 is thus necessary according to (Byrnes and Isidori, 2003). Assumption 6.2, on the other hand, asks for *uniform detectability* of (6.23), with the adjective "uniform" that refers to the fact that condition (6.2) is required to hold for all possible u (see Liberzon et al., 2002). In case of systems with canonical normal form in which (6.23a) does not depend on u, this assumption boils down to a conventional *minimum-phase* requirement, typically assumed in the pertinent literature. Finally, Assumption 6.3 is a robust stabilizability requirement, easily generalizable as in Section 2.3 whenever the high-frequency matrix Ω is state dependent (not done here for ease of exposition). As a first step, we let in (6.6)

$$\Gamma(\eta) := \eta_1$$

and, as customary in the context of minimum-phase systems, we look for a semiglobal stabilization strategy based on high-gain techniques. In particular, with $X \subset \mathbb{R}^n$ an arbitrary compact set, we consider the class of linear static stabilizers

$$u = \mathcal{L}\Big(\mathcal{K}_1\chi + \mathcal{K}_2\zeta + \mathcal{K}_3\eta_1\Big)\,, \tag{6.24}$$

with $\mathcal{K}_1 \in \mathbb{R}^{p\times(n_\chi - p)}$, $\mathcal{K}_2 \in \mathbb{R}^{p\times p}$, $\mathcal{K}_3 \in \mathbb{R}^{p\times p}$ gains to be fixed and with $\mathcal{L} \in \mathbb{R}^{m\times p}$ fulfilling Assumption 6.3. For a given $w(t) \in W$, let $x^\star(t)$ denote the corresponding function defined by Assumption 6.1. In view of Assumption 6.2 and from (6.24) and the choice of Γ above, $y_\eta^\star = \Gamma(\eta^\star)$ must satisfy

$$-\,q(w, x^\star) = \Omega\mathcal{L}\mathcal{K}_3 y_\eta^\star. \tag{6.25}$$

In order to obtain differentiability of $y_\eta^\star$, we constraint $\mathcal{K}_3$ to be non singular, so that $\Omega\mathcal{L}\mathcal{K}_3$ is invertible (the invertibility of $\Omega\mathcal{L}$ is implied by Assumption 6.3). By merging (6.25) with the actual knowledge about $x^\star$ and q, a class $\mathcal{C}_\eta^\star$ for the $\eta^\star$ can be guessed and, according to Section (6.1.2), an appropriate choice of the dimension d of the internal model and of a model set $\mathcal{M}$ (and hence a structure for ϕ) can be derived. An identifier of the form (6.7) that satisfies the strong stability requirement can be then fixed, so as the internal model can be fixed according to Lemma 6.1. As for the stabilizer inside the class (6.24), it turns out that the gain matrices $\mathcal{K}_1$, $\mathcal{K}_2$ and $\mathcal{K}_3$ can be fixed so that the ISS requirement of Proposition 6.1 is fulfilled with $\nu = 0$. This is formalized in the following proposition.

Proposition 6.2. *Suppose that assumptions 6.1, 6.2 and 6.3 hold and let $X \subset \mathbb{R}^n$ be compact. Then there exist $\mathcal{K}_1$, $\mathcal{K}_2$ and an invertible $\mathcal{K}_3$ such that the hypotheses of Proposition 6.1 hold with $\nu = 0$ along all the solutions satisfying $x(t) \in X$.*

We observe that the result holds only as long as the trajectories of the plant remain in the (arbitrary) compact set X. Standard high-gain arguments typically used in the semiglobal stabilization literature, here omitted for reasons of space, can be used to show that the control parameters in (6.24) can be chosen to ensure such a boundedness property, thus completing the result.

Proof of Proposition 6.2. For each $i = 1, \dots, p$, consider the change of variables

$$\begin{aligned} \chi_1^i \mapsto \tilde{\chi}_1^i &:= \chi_1^i + \tilde{\eta}_1^i, \\ \chi_j^i \mapsto \tilde{\chi}_j^i &:= \chi_j^i, \qquad j = 2, \dots, n_\chi^i \end{aligned} \tag{6.26}$$

and let $\tilde{\chi} := \mathrm{col}(\tilde{\chi}^1, \dots, \tilde{\chi}^p)$, with $\tilde{\chi}^i := \mathrm{col}(\tilde{\chi}_1^i, \dots, \tilde{\chi}_{n_\chi^i}^i)$. In the new variables, we have

$$\tilde{e} := e - \tilde{\eta}_1 = C\tilde{\chi}\,.$$

Moreover, from (6.19) we obtain

$$\dot{\tilde{\chi}} = (F + gh_1 C^T C)\tilde{\chi} + H\zeta + C^T(\tilde{\eta}_2 - gh_1\tilde{\eta}_1)\,. \tag{6.27}$$

The following result is an adaptation of Lemma 2.1.

Lemma 6.2. *For any $\epsilon > 0$, there exist $K \in \mathbb{R}^{p\times n_\chi}$, with KCC^T invertible, $\beta_\chi, \tilde{\beta}_e \in \mathcal{KL}$ and $a_\chi > 0$ such that (6.27) with $\zeta = \tilde{\zeta} + K\tilde{\chi}$, being $\tilde{\zeta} \in \mathbb{R}^p$ an auxiliary input,* *satisfies*

$$
\begin{aligned}
|\tilde{\chi}(t)| &\le \beta_\chi(|\tilde{\chi}(0)|, t) + a_\chi\Big(|\tilde{\zeta}|_{[0,t]} + |\tilde{\eta}|_{[0,t]}\Big)\\
|\tilde{e}(t)| &\le \tilde{\beta}_e(|\tilde{\chi}(0)|, t) + \epsilon|\tilde{\zeta}|_{[0,t]} + \epsilon|\tilde{\eta}|_{[0,t]}\,.
\end{aligned}
$$

We keep ϵ as a degree of freedom for now, and, with K produced by Lemma 6.2, we change variables according to

$$
\zeta \mapsto \tilde{\zeta} := \zeta - K\tilde{\chi}\,. \tag{6.28}
$$

In view of (6.23c), $\tilde{\zeta}(t)$ fulfills

$$
\dot{\tilde{\zeta}} = \delta(\tilde{\eta}, \tilde{\chi}, \tilde{\zeta}) + q(w, x) + \Omega u \tag{6.29}
$$

where

$$
\begin{aligned}
\delta(\tilde{\eta}, \tilde{\chi}, \tilde{\zeta}) := -K\Big((F + gh_1C^TC + HK)\tilde{\chi} + H\tilde{\zeta}\\
+ gC^T(\tilde{\eta}_2 - h_1\tilde{\eta}_1)\Big)\,.
\end{aligned}
$$

With $\ell > 0$ a design parameter to be fixed, in (6.24), let

$$
\mathcal{K}_1 := \ell K, \qquad \mathcal{K}_2 := -\ell I_p, \qquad \mathcal{K}_3 := \ell KCC^T\,.
$$

In view of Lemma 6.2, $\mathcal{K}_3$ is invertible. Moreover exists $\ell_1^\star > 0$ such that for all $\ell > \ell_1^\star$ $|\mathcal{K}_3^{-1}| \le 1$. Developing (6.24) yields

$$
u = -\ell\mathcal{L}\big(\zeta - K\chi - KCC^T\eta_1\big) = -\ell\mathcal{L}\tilde{\zeta} + \mathcal{L}\mathcal{K}_3 y_\eta^\star
$$

In view of (6.25), substituting this latter relation in (6.29) yields

$$
\dot{\tilde{\zeta}} = \delta(\tilde{\eta}, \tilde{\chi}, \tilde{\zeta}) + \tilde{q}(w, x, x^\star) - \ell\Omega\mathcal{L}\tilde{\zeta}\,.
$$

where $\tilde{q}(w, x, x^\star) := q(w, x) - q(w, x^\star)$. We will fix ℓ later according to the following Lemma, that is adapted from Lemma 2.2.

Lemma 6.3. *There exist $\beta_\zeta \in \mathcal{KL}$, $a_\zeta > 0$ $\ell_2^\star > \ell_1^\star$ such that, for all $\ell > \ell_2^\star$ and as long* *as $x(t) \in X$, the following holds*

$$|\tilde{\zeta}(t)| \leq \beta_\zeta(|\tilde{\zeta}(0)|, t) + \frac{a_\zeta}{\ell}\big(|\tilde{x}_0|_{[0,t)} + |\tilde{\chi}|_{[0,t)} + |\tilde{\eta}|_{[0,t)}\big)\,. \qquad (6.30)$$

As $(w, x) \in W \times X$, there exists $a_{01} > 0$ for which the function $\rho_0(\cdot)$ of Assumption 6.2 fulfills $\rho_0(|(\chi, \zeta)|_{[0,t)}) \leq a_{01}(|\chi|_{[0,t)} + |\zeta|_{[0,t)})$. By letting $a_{02} := (1 + |K|)a_{01}$, in view of (6.26), (6.28), Assumption 6.2 yields

$$|\tilde{x}_0(t)| \leq \beta_0(|\tilde{x}_0(0)|, t) + a_{02}|\tilde{\chi}|_{[0,t)} + a_{01}\big(|\tilde{\eta}|_{[0,t)} + |\tilde{\zeta}|_{[0,t)}\big)$$

and, as a trivial small-gain condition between $\tilde{\chi}$ and $\tilde{x}$ holds, as long as $x(t) \in X$, we obtain

$$\begin{aligned} |(\tilde{x}_0, \tilde{\chi})| &\leq \beta_{0\chi}(|(\tilde{x}_0(0), \tilde{\chi}(0))|, t) + a_{03}\big(|\zeta|_{[0,t)} + |\tilde{\eta}|_{[0,t)}\big) \\ |\tilde{e}| &\leq \tilde{\beta}_e(|(\tilde{x}_0(0), \tilde{\chi}(0))|, t) + \epsilon\big(|\zeta|_{[0,t)} + |\tilde{\eta}|_{[0,t)}\big) \end{aligned} \qquad (6.31)$$

for some $\beta_{0\chi} \in \mathcal{KL}$ and $a_{03} > 0$. Thus, standard small-gain arguments can be used to claim that, for any $\ell > \ell_3^\star := \max\{\ell_2^\star, a_\zeta a_{03}\}$, the following estimate holds

$$|(\tilde{x}_0, \tilde{\chi}, \tilde{\zeta})| \leq \beta_x(|(\tilde{x}_0(0), \tilde{\chi}(0), \tilde{\zeta}(0)|, t) + a_x|\tilde{\eta}|_{[0,t)} \qquad (6.32)$$

for suitable $\beta_x \in \mathcal{KL}$ and $a_x \in \mathbb{R}_+$. Small-gain arguments can be also used to show that, in view of (6.32), equations (6.30) and (6.31) imply the existence of a $\ell_4^\star \geq \ell_3^\star$ (possibly dependent on ϵ) such that, for all $\ell > \ell_4^\star$, the following holds

$$\begin{aligned} |\tilde{x}(t)| &\leq \beta_x(|\tilde{x}(0)|, t) + a_x|\tilde{\eta}|_{[0,t)} \\ |\tilde{e}(t)| &\leq \beta_e(|\tilde{x}(0)|, t) + a_e(\ell, \epsilon)|\tilde{\eta}|_{[0,t)} \end{aligned}$$

as far as $x \in X$, with $\beta_e \in \mathcal{KL}$ and $a_e(\ell, \epsilon) > 0$ which can be made arbitrarily small by opportunely reducing ϵ and consequently increasing ℓ. In particular, it can be shown that for any $\rho_i > 0$ there exists $\epsilon^\star$ and $\ell^\star(\epsilon) \geq \ell_4^\star$ such that, for any $\epsilon < \epsilon^\star$ and $\ell > \ell^\star(\epsilon)$, $a_e(\epsilon, \ell)\rho_i < 1$, and this concludes the proof. ■

6.3 Examples of Designs

In this section we present two examples showing how in the context of the framework proposed Section 6.1 we can deal with additional outputs not vanishing at the steady state, thus complementing the more general approach developed in Section 2.3 for the non-adaptive case. The first example is academic and it is aimed at showing how systems that cannot be stabilized by error feedback (and thus cannot be dealt with by using a pre-processing approach) can be regulated if additional measurements are present. The second example is an application to the control of the lateral dynamics of the VTOL (Vertical Take-Off and Landing aircraft). The purpose is to show how we can regulate the system by using state feedback and, mainly, to give an example on how the identification problem can be approached.

6.3.1 Dealing with Additional Non-Vanishing Outputs

In this section we present an example showing how the high-gain strategy presented in the previous sections can be easily extended to deal with additional outputs that need not to vanish at the steady state. We consider the system

$$\begin{aligned} \dot{x}_1 &= f_1(x_1) + \gamma_1(w, x_2) + x_3 \\ \dot{x}_2 &= \gamma_2(w, x) + u_1 + u_2 \\ \dot{x}_3 &= \gamma_3(w, x) - b(w)u_1 + (1 - b(w))u_2 \end{aligned} \tag{6.33}$$

with regulation error

$$e := x_2,$$

with γ_i locally Lipschitz functions and with b differentiable. We observe that if we choose the functions γ_i so that $\partial\gamma_2(0)/\partial x_i = 0$, $i = 1, 3$, then we lose detectability from e of the linear approximation of (6.33) at 0. Thus, e is not enough to stabilize (6.33), and additional outputs are needed. We specifically assume to have available for feedback the other two variables, i.e. $y_{\rm m} := {\rm col}(x_1, x_3)$, that, however, do not to vanish at the ideal steady state in which $e = 0$. We also observe that a control strategy based on a preliminary inner-loop that uses $y_{\rm m}$ to reduce to the case of Section 6.2 is hard to imagine, as u_1 and u_2 affect both the equations of $\dot{x}_2$ and $\dot{x}_3$, and they both must be used in case x_3 is pre-stabilized. We observe, thus, that this case does not fit into any of the previous pre-processing frameworks.

In the rest of the section we build a regulator based on Proposition 6.1. For, we suppose to know a function κ such that

$$(s_1 - s_2)(f_1(s_1) - f_1(s_2) + \kappa(s_2) - \kappa(s_1)) \leq 0. \tag{6.34}$$

We change coordinates as

$$\zeta_2 := x_3 + x_1 + \kappa(x_1), \qquad \zeta_1 := x_2, \qquad p := x_1,$$

transforming (6.33) to

$$\dot{p} = g(w, p, \zeta_1) + \zeta_2, \quad \dot{\zeta} = \rho(w, p, \zeta) + B(w)u$$

with ρ, g and B reading as

$$\begin{aligned} g(w, p, \zeta_1) &:= -p + f_1(p) - \kappa(p) + \gamma_1(w, \zeta_1) \\ \rho(w, p, \zeta) &:= \begin{pmatrix} \gamma_2(w, (p, \zeta_1, \zeta_2 - p - \kappa(p))) \\ \gamma_3(w, (p, \zeta_1, \zeta_2 - p - \kappa(p))) + (1 + \kappa'(p))(g(w, p, \zeta_1) + \zeta_2) \end{pmatrix} \\ B(w) &:= \begin{pmatrix} 1 & 1 \\ -b(w) & 1 - b(w) \end{pmatrix}. \end{aligned}$$

With $\alpha > 0$ fulfilling

$$4\alpha > 1 + \sup_{w \in W} (1 + b(w))^2$$

, let

$$\mathcal{L} := \begin{pmatrix} \alpha^2 & 0 \\ 0 & \alpha \end{pmatrix}, \qquad \mathcal{S}(w) := \begin{pmatrix} 1 + b(w)^2 & b(w) \\ b(w) & 1 \end{pmatrix}.$$

Then the high-frequency matrix $B(w)$ fulfills

$$\mathcal{L}^T B(w, x)^T \mathcal{S}(w) + \mathcal{S}(w) B(w, x) \mathcal{L} = \begin{pmatrix} 2\alpha^2 & \alpha(1 + b(w)) \\ \alpha(1 + b(w)) & 2\alpha \end{pmatrix} =: M(w). \tag{6.35}$$

As $2\alpha^2 > 0$ and

$$\begin{aligned}\det M(w) &= 4\alpha^3 - \alpha^2(1+b(w))^2 = \alpha^2(4\alpha - (1+b(w))^2)\\ &> 4\alpha^2 + \alpha^2\left(4\sup_{w\in W}(1+b(w))^2 - (1+b(w))^2\right) > 4,\end{aligned}$$

then by definition of α, $M(w)$ is positive definite and there exists $m > 0$ such that, for all $x \in \mathbb{R}^2$,

$$x^T M(x) x \geq m|x|^2. \tag{6.36}$$

With y_η a virtual input and k a control parameter, consider the control law

$$u = -k\mathcal{L}\begin{pmatrix}\zeta_1 + y_\eta\\ \zeta_2\end{pmatrix} = \begin{pmatrix}-k\alpha^2(\zeta_1+y_\eta)\\ -k\alpha\zeta_2\end{pmatrix}.$$

It can be shown that this control law ensures bounded solutions semi-globally, with the domain of validity that increases with k. In these variables, the regulator equations give

$$\begin{aligned}\zeta_1^\star &= 0\\ \dot{p}^\star &= g(w,p^\star,0) + \zeta_2^\star\\ \dot{\zeta}_2^\star &= \rho_2(w,p^\star,\zeta^\star) - b(w)u_1^\star + (1-b(w))u_2^\star\\ 0 &= \rho_1(w,p^\star,\zeta^\star) + u_1^\star + u_2^\star.\end{aligned}$$

from which we deduce

$$y_\eta^\star = -u_1^\star/(\alpha^2 k), \qquad u_2^\star = -k\alpha\zeta_2^\star,$$

i.e.

$$\begin{aligned}\dot{\zeta}_2^\star &= \rho_2(w,p^\star,\zeta^\star) + b(w)\rho_1(w,p^\star,\zeta^\star) - k\alpha\zeta_2^\star\\ y_\eta^\star &= \big(\rho_1(w,p^\star,\zeta^\star) - k\alpha\zeta_2^\star\big)/(\alpha^2 k).\end{aligned}$$

From the latter equations, and by letting, as in Section 6.2, $y_\eta = \Gamma(\eta) := \eta_1$, we can thus infer a possible class $\mathcal{C}_\eta^\star$ of ideal steady states for η defined by $\eta_1^\star = y_\eta^\star$ and $\eta_i^\star = \dot{\eta}_{i-1}^\star$. We then fix the identifier degrees of freedom and the matrices G_i by following the guidelines of Section 6.1.3 and Lemma 6.1, thus obtaining the ideal steady state $z^\star$ as detailed in Section 6.1.3. Consider the change of variables

$$\begin{aligned} \tilde{p} &:= p - p^\star, & \tilde{\eta} &:= \eta - \mathrm{col}\left(\eta^\star_{[1,d-1]},\ \phi(\eta^\star_{[1,d-1]}, \theta^\star)\right) \\ \tilde{\zeta}_2 &:= \zeta_2 - \zeta_2^\star, & \tilde{z} &:= z - z^\star, \quad \tilde{\zeta}_1 := \zeta_1 + \tilde{\eta}_1 \end{aligned}$$

that yields

$$\begin{aligned} \dot{\tilde{p}} &= \tilde{g}(w, \tilde{p}, \tilde{\zeta}, p^\star) + \tilde{\zeta}_2 \\ \dot{\tilde{\eta}} &= \tilde{\Phi}(\tilde{\eta}, \tilde{z}, \eta^\star, z^\star, \varepsilon^\star) + G\tilde{\zeta}_1 \\ \dot{\tilde{z}} &= \tilde{\mu}(\tilde{z}, \tilde{\eta}, \eta^\star, z^\star) \\ \dot{\tilde{\zeta}} &= \tilde{\rho}(w, \tilde{p}, \tilde{\zeta}, p^\star, \zeta^\star) - kB(w)\mathcal{L}\tilde{\zeta} \end{aligned}$$

with $\tilde{\rho}$ and $\tilde{g}$ that, in view of (6.34), fulfill

$$\begin{aligned} |\tilde{\rho}(w, \tilde{p}, \tilde{\zeta}, p^\star, \zeta^\star)| &\leq r_1(|\tilde{\zeta}| + |\tilde{p}| + |\tilde{\eta}_{[1,2]}|) \\ \tilde{p}\tilde{g}(w, \tilde{p}, \tilde{\zeta}, p^\star) &\leq -\frac{1}{2}|\tilde{p}| + \frac{1}{2}r_2(|\tilde{\zeta}_1| + |\tilde{\eta}_1|) + \frac{1}{2}|\tilde{\zeta}_2| \end{aligned} \tag{6.37}$$

for some $r_1, r_2 \in \mathcal{K}$ that are locally Lipschitz. Thus, as $\mathcal{S}(w) > 0$, quite standard high-gain arguments can be used to show that, considering the function

$$V := |\tilde{p}| + \sqrt{\tilde{\zeta}^T \mathcal{S}(w) \tilde{\zeta}},$$

and noting that (6.35), (6.36) imply $-k2\tilde{\zeta}^T\mathcal{S}(w)B(w)\mathcal{L}\tilde{\zeta} \leq -km|\zeta|^2$, then, in view of (6.37), for each compact set of initial conditions $X \subset \mathbb{R}^3$, we can find $k^\star > 0$ such that, for all $k > k^\star$, the hypotheses of Proposition 6.1 hold with $\nu = 0$, which in turn yields the result of Theorem 6.1 with $\nu_{\mathrm{s}} = 0$.

6.3.2 Application to the Control of the VTOL

In this section we present an application to the regulation of the lateral position of the VTOL aircraft. The aim of this example is to show how the design of the identifier can be approached in the high-gain setting developed in the previous sections (again in presence of additional output variables not necessarily vanishing at the steady state). For compactness we disregard the equations of the vertical dynamics, as it can be controlled in a separate control loop. The dynamics of the lateral (x_1, x_2) and angular (x_3, x_4) positions of the VTOL aircraft can be described by the equations (Isidori et al., 2003)

$$\begin{array}{rclcrcl} \dot{x}_1 & = & x_2 & \qquad & \dot{x}_3 & = & x_4 \\ \dot{x}_2 & = & q(w) - g\tan x_3 + v & & \dot{x}_4 & = & Bu\,. \end{array} \tag{6.38}$$

with $M > 0$ the VTOL mass, $g > 0$ the gravitational constant and $B = 2LJ^{-1} > 0$, with $L > 0$ the length of the wings and J the moment of inertia (typically uncertain). The input u is the force on the wingtips, v is a vanishing input taking into account the (controlled) vertical dynamics and $q(w) := M^{-1}q_0(w)$, with $q_0(w)$ that is the lateral force produced by the wind. The control goal is to eliminate the wind action from the lateral position dynamics, i.e. the regulation error is defined as $e(t) = x_1(t)$. We also suppose to have available for feedback the entire state, namely $y = x$. We stress that, although it is usually the case in practice to have the whole state available for feedback, previous output regulation solutions (see e.g. (Marconi and Praly, 2008a)) allow to use only e as a control variable. Here instead we take advantage from the additional information. Let $w(t)$ be generated by an exosystem of the form (6.2). The corresponding solution $(x^\star, u^\star)$ to the regulator equations fulfil $x_1^\star = x_2^\star = 0$, $x_3^\star = \tan^{-1}(q(w)/g)$, $x_4^\star = gL_s q(w)/(g^2 + q(w)^2)$ and

$$u^\star = \frac{g}{B}\left(\frac{L_s^2 q(w)}{g^2 + q(w)^2} - \frac{2(L_s q(w))^2 q(w)}{(g^2 + q(w)^2)^2}\right).$$

We consider the change of coordinates $x \mapsto \chi$, where

$$\begin{array}{ll} \chi_1 := x_1, & \chi_2 := x_2, \\ \chi_3 := -g\tan x_3 + q(w), & \chi_4 = L_s q(w) - g x_4/(\cos x_3)^2 \end{array}$$

that yields

$$\begin{array}{rclcrcl} \dot{\chi}_1 & = & \chi_2 & \quad & \dot{\chi}_3 & = & \chi_4 \\ \dot{\chi}_2 & = & \chi_3 & & \dot{\chi}_4 & = & b(w,\chi) - \Omega(w,\chi)u \end{array}$$

with $b(w,\chi)$ and $\Omega(w,\chi)$ given by

$$\begin{aligned} b(w,\chi) &:= L_s^2 q(w) - \frac{1}{g}(\chi_4 - L_s q(w))^2 \sin\left(2\tan^{-1}\left(\frac{q(w)-\chi_3}{g}\right)\right) \\ \Omega(w,\chi) &:= gB\left(\cos\left(\tan^{-1}\left(\frac{q(w)-\chi_3}{g}\right)\right)\right)^{-2}. \end{aligned}$$

With y_η an auxiliary input, $c \in \mathbb{R}^4$ and $k, \ell > 0$ design parameters we think of a control law of the kind

$$u = \ell\Big(c_1 k^4(x_1 + y_\eta) + c_2 k^3 x_2 + c_3 k^2(-g \tan x_3) + c_4 k(-g x_4/\cos^2 x_3)\Big).$$

that in the χ coordinates reads as

$$u = \ell\Big(c_1 k^4(\chi_1 + y_\eta) + c_2 k^3 \chi_2 + c_3 k^2 \chi_3 + c_4 k \chi_4 - c_3 k^2 q(w) - c_4 k L_s q(w)\Big)\,.$$

Therefore the ideal steady-state value of $y_\eta^\star$ is given by

$$y_\eta^\star := \frac{c_3}{c_1 k^2} q(w) + \frac{c_4}{c_1 k^3} L_s q(w) + \frac{1}{c_1 \ell k^4} \Omega(w,0)^{-1} b(w,0)\,. \tag{6.39}$$

We approach the design of the internal model unit by letting as before $y_\eta = \eta_1$, with η_1 the first component of an adaptive internal model unit of the form (6.5), with the order d, the function $\phi(\eta,\theta)$ and the identifier subsystem z that are chosen on the basis of the class $\mathcal{C}_\eta^\star$ of functions that, in view of (6.39), are linear combinations of $\dot q(w)$, $q(w)$ and $\Omega(w,0)^{-1}b(w,0)$. For clarity of exposition, details on the choice of ϕ are postponed to the end of the section. Once fixed ϕ, we fix the matrices G_i and ψ according to Lemma 6.1 and, in the following, we approach the design of k and ℓ so as to fulfill the hypotheses of Proposition 6.2. For, we define $\tilde\eta$ according to (6.17) and, with $c \in \mathbb{R}^4$ chosen so that $c_4 = 1$ and $p(s) := s^3 + c_3 s^2 + c_2 s + c_1$ is an Hurwitz polynomial, we further change variables as $\chi \mapsto (\tilde\chi, \tilde\zeta)$, with

$$\begin{aligned} \tilde\chi_1 &:= \chi_1 + \tilde\eta_1, & \tilde\chi_2 &:= k^{-1}\chi_2 \\ \tilde\chi_3 &:= k^{-2}\chi_3, & \tilde\zeta &:= k^{-3}\chi_4 + c_1\tilde\chi_1 + c_2\tilde\chi_2 + c_3\tilde\chi_3\,. \end{aligned}$$

In the new coordinates we obtain

$$\begin{aligned} \dot{\tilde\chi} &= kM\tilde\chi + f(\tilde\chi,\tilde\zeta,\tilde\eta) \\ \dot{\tilde\zeta} &= -\ell k \Pi(w,\tilde\chi,\tilde\zeta,\tilde\eta)\tilde\zeta + \Delta(w,\tilde\chi,\tilde\zeta,\tilde\eta) \end{aligned}$$

where M is Hurwitz,

$$f(\tilde\chi,\tilde\zeta,\tilde\eta) := \mathrm{col}(\tilde\eta_2 - G_1\tilde\eta_1 + G_1\tilde\chi_1, 0, k\tilde\zeta),$$

being G_1 is the same matrix of the internal model unit (6.5) and

$$\begin{aligned}
\Pi(w,\tilde{\chi},\tilde{\zeta},\tilde{\eta}) &:= \Omega\left(w,(\tilde{\chi}_1-\tilde{\eta}_1,k\tilde{\chi}_2,k^2\tilde{\chi}_3,k^3(\tilde{\zeta}-c_{[1,3]}^T\tilde{\chi}))\right)\\
\Delta(w,\tilde{\chi},\tilde{\zeta},\tilde{\eta}) &:= c_{[1,3]}^T\big(kM\tilde{\chi}+f(\tilde{\chi},\tilde{\zeta},\tilde{\eta})\\
&\qquad + k^{-3}b\big(w,(\tilde{\chi}_1-\tilde{\eta}_1,k\tilde{\chi}_2,k^2\tilde{\chi}_3,k^3(\tilde{\zeta}-c_{[1,3]}^T\tilde{\chi}))\big)\\
&\qquad - k^{-3}\Pi(w,\tilde{\chi},\tilde{\zeta},\tilde{\eta})\Omega(w,0)^{-1}b(w,0)\,.
\end{aligned}$$

For any choice of G_1, there exists $L_f>0$ such that

$$|f(\tilde{\chi},\tilde{\zeta},\tilde{\eta})|\le L_f(|\tilde{\chi}|+|\tilde{\eta}|+k|\tilde{\zeta}|).$$

Moreover, $\Pi(w,\tilde{\chi},\tilde{\zeta},\tilde{\eta})$ depends on $(\tilde{\chi},\tilde{\zeta},\tilde{\eta})$ only throughout $\tilde{\chi}_3$ and it is bounded by above and below in each compact subset of $\mathbb{R}^{n_w}\times\mathbb{R}^4\times\mathbb{R}^d$ and $\Delta(w,\tilde{\chi},\tilde{\zeta},\tilde{\eta})$ is locally Lipschitz and vanishes when $(\tilde{\chi},\tilde{\zeta},\tilde{\eta})=0$, for any $w\in\mathbb{R}^{n_w}$. By standard high-gain arguments, it is thus possible to conclude that, for any compact subset $X\subset\mathbb{R}^4$, there exist $k^\star,\ell^\star(k)>0$ such that for all $k>k^\star$ and $\ell>\ell^\star(k)$ the assumptions of Proposition 6.1 hold with $\nu=0$.

We propose now a design example for the internal model unit and the identifier in the case in which $q(w(t))$ is a quasi-periodic signal characterized by a stronger dominant frequency component and weaker higher harmonics, and the goal is to learn and compensate the dominant harmonic. We first observe that the constants that multiply the terms $L_s q(w)$ and $\Omega(w,0)^{-1}b(w,0)$ in the expression (6.39) of $y_\eta^\star$ are much smaller than those multiplying $q(w)$. In order to simplify the problem, and to have better insight on $y_\eta^\star$, we thus approximate the class $\mathcal{C}_\eta^\star$ as

$$\mathcal{C}_\eta^\star\approx\widehat{\mathcal{C}}_\eta^\star:=\{\eta^\star\;:\;y_\eta^\star=\eta_1^\star=\alpha q(w)\,,\dot{\eta}_i^\star=\eta_{i+1}^\star\,,\alpha\in(0,1)\}$$

(we assumed without loss of generality that $k^2>c_3/c_1$). The approximate class $\widehat{\mathcal{C}}_\eta^\star$ contains thus signals with the same frequency content of $q(w)$. In order to fix the identifier, we infer a prediction error model of the kind

$$\ddot{y}_\eta^\star=-\theta y_\eta^\star\,,\qquad\theta\in\mathbb{R}\,,$$

which captures the dynamical model of a single harmonic. In view of the discussion developed in Section 6.1, we choose the order of the internal model as

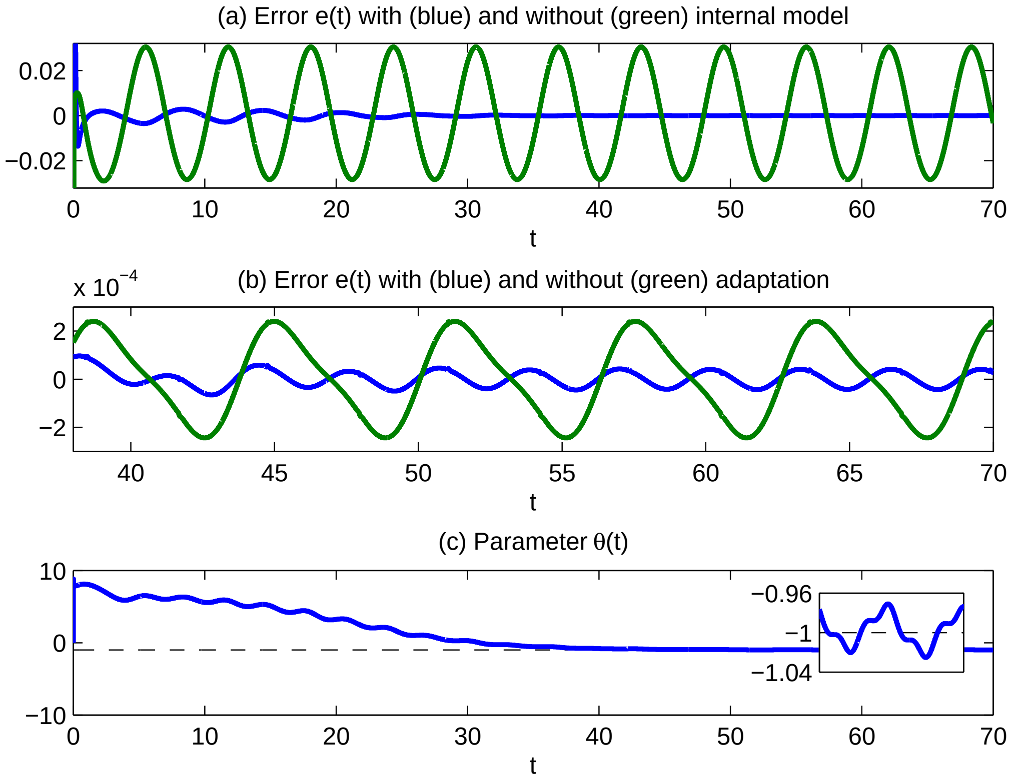

Figure 6.2: Simulation results: figure (a) shows the error trajectory with and without (i.e. with $u_\eta = 0$) adaptive internal model unit. Figure (b) instead shows the comparison between the steady state regulation error when adaptation is used and when instead $\psi(\eta, \theta) = 0$. Figure (c) shows the trajectory of the parameter $\theta(t)$.

$d = 2 + 1 = 3$ and we use a continuous-time least-squares identifier of the kind introduced in Section 4.3, with $d = 1$ and $\sigma(\eta) := \eta_1$. This yields a function ϕ given by

$$\phi(\eta, \theta) = \theta \eta_1 .$$

We conclude the design by defining ψ by any opportunely saturated version of (6.15) and, as mentioned before, by designing the matrices G_i according to Lemma 6.1. Figure 6.2 shows the result of a simulation obtained with $M = 5 \cdot 10^4$Kg, $L = 2$m, $J = 1.25 \cdot 10^4$Kg/m^2 and where we let $q_0(w) = 2 \cdot 10^7 w_1 + 10^6 w_3$, with $w \in \mathbb{R}^4$ that is generated by the system

$$\begin{aligned} \dot{w}_1 &= w_2 & \qquad \dot{w}_3 &= w_4 \\ \dot{w}_2 &= -w_1 & \qquad \dot{w}_4 &= -4 w_3 \,. \end{aligned}$$

with initial condition $w(0) = \mathrm{col}(1, -1, 0, -1)$. Thus $q_0(w)$ is a periodic signal with a dominant harmonic at frequency 1 rad/s and such that $q(w) = q_0(w)/M$ has the same order of magnitude of the weight of the VTOL. The control parameters have been chosen as: $G_1 = 15$, $G_2 = 75$, $G_3 = 125$, $k = 170$, $\ell = 250$, $\lambda = 0.2$ and $\Gamma = 10^{-8}$. Figure (b) shows how the dominant component of $q(w)$ is "learned", thus leading to a considerable compensation of the corresponding harmonic in the regulation error, in which only high frequency components can be observed.

6.4 Adaptive Regulation of Linear Systems via Slow Identifiers

In this section we consider the problem of adaptive output regulation for general multivariable linear systems. Unlike the design examples proposed so far, in this case we do not rely on a "high-gain" strategy to fit into the hypotheses of Theorem 6.1 and, rather, we leverage on the separation of the time-scales obtained by letting the identifier to be slow enough compared to the controlled plant. We postpone a literature overview of adaptive output regulator designs for linear systems to Section 7.1, where the problem is taken on in a more general envelope. Here we consider systems of the form

$$\begin{aligned} \dot{w} &= Sw \\ \dot{x} &= Ax + Bu + Pw \\ e &= C_e x + Q_e w \end{aligned} \tag{6.40}$$

with $w \in \mathbb{R}^{n_w}$, $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, $e \in \mathbb{R}^p$, being $n, n_w, m, p, q \in \mathbb{N}$ such that $m \geq p$, and with S that is an *unknown* matrix. For simplicity, we address the state feedback case, as output feedback can be obtained by means of the same arguments, and we assume that S is simply stable and (A, B) is stabilizable. By following the procedure of Section 6.1, we exploit linearity to fix the structure of the internal model unit and the stabilizer. In particular, in view of Section 1.1.3, we let in (6.5), (6.6) $d = n_w + 1$, $\Gamma(\eta) = \eta$, $G_d = I_p$ and $G_i = 0$ for $i = 1, \dots, d-1$. We then fix the class of stabilizers (6.8) as the class of static state-feedback control laws of the kind

$$u = K_x x + K_\eta \eta \tag{6.41}$$

with $K := \begin{pmatrix} K_x & K_\eta \end{pmatrix}$ to be fixed. To set up the identification problem, we let the model set $\mathcal{M}$ be the set of functions $\phi : \mathbb{R}^{p(d-1)} \to \mathbb{R}^p$ of the form

$$\phi(\eta_{[1,d-1]}, \theta) = (\theta^T \otimes I_p)\eta_{[1,d-1]},$$

and we fix a multivariable version of the continuous-time least-squares identifier introduced in Section 4.3, with $n_\theta := \dim(\theta) = d-1$, $\sigma_i(\eta) = \eta_i$ for $i = 1, \ldots, d-1$, and with

$$\sigma(\eta) := \mathrm{col}\left(\sigma_1(\eta)^T, \ldots, \sigma_{d-1}(\eta)^T\right)$$

and $\lambda > 0$ that is a *small* number to be tuned. By following Section 6.1.2, we consider an expression of ψ obtained according to (6.15), obtaining a function of the form

$$\psi(\eta, z) := \lambda\rho_0(z, \eta) + (\theta^T \otimes I_p)\eta_{[2,d]}, \tag{6.42}$$

for some ρ_0. Here, however, instead of (6.42) we implement the following modified function:

$$\psi(\eta, z) := \lambda\rho(z, \eta) + (\mathrm{p}_{\mathcal{E}}(\theta)^T \otimes I_p)\eta_{[2,d]},$$

where ρ is a bounded function obtained by saturating ρ_0, $\mathcal{E}$ is a compact convex set to be fixed and $\mathrm{p}_{\mathcal{E}}$ denotes any Lipschitz selection of the projection map from $\mathbb{R}^{n_\theta}$ onto $\mathcal{E}$. We can write the internal model unit in the compact form

$$\dot{\eta} = \Psi(\theta)\eta + Ge + \lambda G\rho(z, \eta), \tag{6.43}$$

for some $\Psi : \mathbb{R}^{n_\theta} \to \mathbb{R}^{pd \times pd}$ and, by letting $\xi := \mathrm{col}(x, \eta)$, for appropriate P_ξ, ℓ and ρ_ξ we can write the closed-loop system (4.21), (4.22), (6.40), (6.41), (6.43) as

$$\begin{array}{rcl} \dot{z} &=& \lambda\ell(z, \eta), \\ \dot{w} &=& Sw \\ \dot{\xi} &=& (A_\xi(\theta) + B_\xi K)\xi + P_\xi w + \lambda\rho_\xi(z, \xi) \end{array} \tag{6.44}$$

where

$$A_\xi(\theta) := \begin{pmatrix} A & 0 \\ GC_e & \Psi(\theta) \end{pmatrix}, \quad B_\xi := \begin{pmatrix} B \\ 0 \end{pmatrix}.$$

Let $\Theta_{\rm H}$ be the set of θ for which $A_\xi(\theta) + B_\xi K$ can be made Hurwitz (i.e. for which the non-resonance conditions hold). It can be shown that $\mathbb{R}^{n_\theta} \setminus \Theta_{\rm H}$ is of null Lebesgue measure, so that we can find arbitrarily large compact sets $\mathcal{E}$ inside $\Theta_{\rm H}$. In the following we denote by $\bar{A}_\xi$ the unconstrained map A_ξ obtained with $\mathrm{p}_{\mathcal{E}}(\theta) = \theta$. Fix $\epsilon, r > 0$ such that $r > \epsilon$ and $\bar{\theta} \in \Theta_{\rm H}$ arbitrarily, and let $K \in \mathbb{R}^{m\times(n+pd)}$ be such that $\bar{A}_\xi(\bar{\theta}) + rI + B_\xi K$ is Hurwitz[2]. Then the eigenvalues of $\bar{A}_\xi(\bar{\theta}) + B_\xi K$ have real part smaller than $-r$. Moreover, with $\bar{\mu}$ the eigenvalue of $\bar{A}_\xi(\bar{\theta}) + B_\xi K$ with largest real part, we observe that the map

$$\theta \mapsto \Lambda(\theta) := \max_{\mu\in\sigma(A_\xi(\theta)+B_\xi K)} |\Re[\bar{\mu}] - \Re[\mu]|,$$

where $\Re[\mu]$ denotes the real part of μ, is continuous. Therefore, since the difference $(\bar{A}_\xi(\bar{\theta}) + B_\xi K) - (A_\xi(\theta) + B_\xi K) = \bar{A}_\xi(\bar{\theta}) - A_\xi(\theta)$ is a function only of $\theta - \bar{\theta}$, there exists a non-empty compact convex set $\mathcal{E} \subset \Theta_{\rm H}$ such that $\bar{\theta} \in \mathcal{E}$ and, for all $\theta \in \mathcal{E}$ and all $\mu \in \sigma(A_\xi(\theta) + B_\xi K)$, we have $|\Re[\mu] - \Re[\bar{\mu}]| \le r - \epsilon$, and hence

$$\Re[\mu] = \Re[\bar{\mu}] + (\Re[\mu] - \Re[\bar{\mu}]) \le -r + (r - \epsilon) \le -\epsilon.$$

We remark that the procedure described above leads to a *local* existence result of $\mathcal{E}$, once fixed K, ϵ and r. Nevertheless, r can be taken arbitrarily large, thus potentially allowing $\mathcal{E}$ to be taken arbitrarily large. We also observe that the boundedness of ρ implies that the trajectories of the closed-loop system (6.44) are uniformly ultimately bounded.

The only parameter that remains to fix is λ. We approach its design by noting that low values of λ induce a small gain condition in the interconnection of the systems z and (w, ξ), and in a consequent separation of the time-scales. Let $\Pi(\theta)$ be the unique (smooth in θ) solution to the Sylvester equation $\Pi(\theta)S - (A_\xi(\theta) + B_\xi K)\Pi(\theta) = P_\xi$. Then, when $\lambda = 0$, the subsystem (w, ξ) of (6.44) has a globally exponentially stable attractor given by the graph of $w \mapsto \Pi(\theta)w$. Let $\Pi_\eta(\theta) \in \mathbb{R}^{pd\times n_w}$ be such that Π can be partitioned as $\Pi = \mathrm{col}(\Pi_x, \Pi_\eta)$ for some $\Pi_x(\theta) \in \mathbb{R}^{n\times n_w}$. Then, when $\xi = \Pi(\theta)w$, the input to the identifier is $\eta = \Pi_\eta(\theta)w$. Let

[2]This is possible as stabilizability of $(\bar{A}_\xi(\theta), B_\xi)$ implies those of $(\bar{A}_\xi(\theta) + rI, B_\xi)$.

$\Pi_{\eta_i}(\theta) \in \mathbb{R}^{p\times n_w}$, $i = 1,\dots,d$, be such that we can write $\Pi_\eta = \mathrm{col}(\Pi_{\eta_1},\dots,\Pi_{\eta_d})$. The structure of $\Psi(\theta)$ gives

$$\begin{aligned} \Pi_{\eta_i} S &= \Pi_{\eta_{i+1}}, \qquad i = 1,\dots,d-1 \\ \Pi_{\eta_i} &= \Pi_{\eta_1} S^{i-1} \quad i = 1,\dots,d. \end{aligned} \tag{6.45}$$

As $d-1 = n_w$, by letting c_i, $i = 1,\dots,n_w$ be such that

$$s^{n_w} + c_{n_w-1}s^{n_w-1} + \ldots + c_1 s + c_0 \tag{6.46}$$

coincides with the characteristic polynomial of S, from (6.45) and by the Cayley-Hamilton Theorem we also have

$$\begin{aligned} \Pi_{\eta_d} = \Pi_{\eta_1} S^{d-1} = \Pi_{\eta_1} S^{n_w} &= -\Pi_{\eta_1}\sum_{i=1}^{n_w} c_{i-1} S^{i-1} \\ = -\sum_{i=1}^{n_w} c_{i-1}\Pi_{\eta_1} S^{i-1} &= -\sum_{i=1}^{n_w} c_{i-1}\Pi_{\eta_i}. \end{aligned} \tag{6.47}$$

Since the input to the least squares identifier (4.21)-(4.22) in the reduced system reads as

$$\eta_{[1,d-1]} = \mathrm{col}(\Pi_{\eta_1}(\theta),\dots,\Pi_{\eta_{d-1}}(\theta))w, \quad \eta_d = \Pi_{\eta_d}(\theta)w, \tag{6.48}$$

then (6.47) implies that the least squares problem (4.20) with $\Omega = 0$ has a global solution given by

$$\theta^\circ := -\,\mathrm{col}(c_0,\dots,c_{n_w-1}).$$

The quantity θ° is also the *unique* minimum along the solutions that satisfy the following *strong persistence of excitation property*:

Definition 6.1. *With $\epsilon_R > 0$ the input η is said to have the ϵ_R-*strong persistence of excitation property *if there exists $T > 0$ such that, along the solutions to* (4.21), (4.22) *with input* (6.48), *it holds that $\min\sigma(R(t)) \geq \epsilon_R$ for all $t \geq T$.*

Clearly, if η has the ϵ_R-strong persistence of excitation property it also has the persistence of excitation property of Definition 4.3 for $t \geq T$. Thus, Proposition 4.2 and the continuity of R as a function of η can be invoked to claim that the identifier (4.21) with input (6.48) is such that $\theta \to \theta^\circ$. Suppose that $\theta^\circ \in \mathcal{E}$, then by definition of Ψ, G and Π_η, and by using again (6.45) and the Cayley-Hamilton

Theorem, we obtain that, if $\theta = \theta^\circ$, then the quantity $\Pi_e := C_e \Pi_x + Q_e$ fulfills

$$\begin{aligned}\Pi_e(\theta^\circ) &= \Pi_{\eta_d}(\theta^\circ) S - \sum_{i=1}^{d-1} \theta_i^\circ \Pi_{\eta_{i+1}}(\theta^\circ) \\ &= -\sum_{i=1}^{d-1} (c_{i-1} + \theta_i^\circ) \Pi_{\eta_i} S = 0.\end{aligned}$$

Moreover, we observe that Π is continuous in θ and that the function ρ_ξ in (6.44) is vanishing in $z = z^\star$ (where $\theta^\star = \theta^\circ$ is constant), locally Lipschitz with a Lipschitz constant possibly depending on the particular ϵ_R for which Definition 6.1 holds, and it is multiplied by λ in the equation of $\dot{\xi}$. Therefore, standard small-gain arguments and Proposition 4.2 can be used to show that, if λ is taken sufficiently small, then the stability requirement holds with $\varepsilon^\star = 0$ and with $\nu = 0$ along the solutions for which η has the strong persistence of excitation property with a fixed ϵ_R. We summarize the result in he following proposition.

Proposition 6.3. *The closed-loop system* (4.21), (4.22), (6.40), (6.41), (6.43) *has bounded trajectories. If in addition there exists $c \in \mathcal{E}$ such that* (6.46) *holds, then for any $\epsilon_R > 0$ there exists $\lambda^\star > 0$ such that, for all $\lambda \in (0, \lambda^\star)$, any solution of the closed-loop system for which η has the ϵ_R-strong persistence of excitation property also fulfills*

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

7

Adaptive Output Regulation for Linear Systems via Discrete-Time Identifiers

This chapter presents a different approach to the problem of output regulation for linear multivariable systems that relies on *discrete-time identifiers*. Other than the usage of discrete-time identifiers, which turns the closed-loop system into a hybrid system, the proposed approach differs from the adaptive framework of the previous sections by the fact that also the stabilizer depends on the identifier's state. In the general nonlinear terms of Section 6.1, having a stabilizer that depends on the identifier would result in a problematic invertibility condition taking place instead of (6.9), with a chicken-egg dilemma that would have to be extended to include also the identifier. Nevertheless, the assumption of linearity provides a powerful way to break the chicken-egg dilemma and allows us to consider more flexible stabilization procedures that also involve the identifier's state.

The focus to linear systems is further motivated by the fact that, up to our knowledge, no adaptive design is known that can overcome uncertainties in the exosystem in a general multivariable and not necessarily minimum phase case (except for the design of Section 6.4). We present here a simple solution to this problem, with a regulator that is based on the classical design presented in Section 1.1.3, in which and identifier adapts the internal model frequencies in a data-driven fashion. We approached the problem in the same philosophy of the previous chapters, strongly inspired by the intuition of dual control and iterative identification: the closed-loop system alternates continuous-time flows to jump times in which the adaptation of the internal model parameters takes place; the stabilizer is designed to ensure that, for each "guess" of the internal model parameters, the closed-loop system has a well-defined (quasi) steady-state that attracts the solutions during the successive flow time. The temporal distance between two successive jumps is taken large enough to let the closed-loop trajectories to get close to the steady-state. The identifier is designed independently from the underlying problem and, interestingly enough, it is built to fit in the same framework of Section 4.1 and to fulfill the same conditions of the identifier requirement (see requirement 4.1). Linearity then plays a crucial role in making sure that the identification problem solved by the identifier makes sense for each sequence of quasi steady-states and, in particular, that an unique solution exists despite the fact that each quasi steady state depends on the (wrong) previous guess of the parameters.

This section contains original results submitted for publication in (Bin et al., 2018c).

7.1 Previous Designs of Adaptive Linear Regulators

The general problem of designing a regulator that ensures asymptotic regulation for linear systems in presence of uncertainties in the exosystem is still open, and the existing results only cover limited classes of plants or exosystems. Single-input single-output (SISO) linear systems have been considered in (Marino and Tomei, 2003; Marino and Santosuosso, 2007) using adaptive observers. In the first work the order of the exosystem is known, in the second the knowledge of its upper bound is sufficient. In both the papers, though, the plant's matrices are assumed to be perfectly known, so as robustness relative to exosystems perturbations is traded for those relative to the plant. For what concerns adaptive designs for multivariable linear systems, strongly minimum-phase normal forms have been considered in (Muzimoto and Iwai, 2007), while state-feedback tracking for more general linear systems has been studied in (Bado and Ichikawa, 2006). Other approaches have been developed in the context of nonlinear systems. Nonlinear systems in output feedback form driven by uncertain linear exosystems have been considered for instance in (Nikiforov, 1998; Ding, 2003), where adaptive backstepping techniques are used. Nonlinear minimum-phase SISO normal forms have been considered in many papers. For example, in (Serrani et al., 2001) an ad hoc adaptation algorithm is constructed based on Lyapunov arguments, in (Delli Priscoli et al., 2006) adaptation is carried out by using the theory of adaptive observers, in (Isidori et al., 2012) unknown linear exosystems are immersed into larger parameterless nonlinear exosystem whose dynamic is known and that can be dealt with in a nonlinear regulation setting. The same idea was applied to a class of uncertain nonlinear exosystems in (Forte et al., 2013; Bin et al., 2016). Finally, in (Forte et al., 2017; Bin and Marconi, 2017a, 2018a) adaptation is cast as a system identification problem, and parameter estimation is performed by any continuous- or discrete-time algorithm satisfying some strong stability properties.

In this chapter we consider the output regulation problem for general multivariable linear systems, with the reference signals and the disturbances that are generated by an unknown exosystem. We endow the linear regulator of Davison (1976) with a *discrete-time* identification unit which adapts the internal model on the basis of the closed-loop measurable states. The identification algorithm implements a recursive least-squares scheme of the kind presented in Section 4.2. The regulator is designed to ensure the existence of a “temporary” steady state between two successive updates of the identifier, despite the possible wrong value of the estimated parameters. Even if the regulator errors do not vanish in this temporary steady state, the controlled plant still oscillates with the same modes of the exosystem, thus unveiling to the identifier the unknown frequency content of the external excitation. This, in turn, allows the identifier to eventually estimate the “right” parameters, no matter how “wrong” are the temporary steady states, as long as the dimension of the internal model is sufficiently large and a persistency of excitation condition is fulfilled. Moreover, the regulator corresponding to the "right" parameter has the internal model property, and it thus guarantees asymptotic regulation. In this respect, we observe that the proposed approach can be framed in a *"dual control"* perspective (see e.g. Feldbaum (1960); Gevers (2005)), where the regulator plays the double role of inducing the right dynamics making the identification of the unknown parameters possible, and then stabilizing the "right" steady state when the parameter is correctly estimated.

7.2 Problem Formulation

We consider linear systems of the form

$$\begin{aligned} \dot{x} &= Ax + Bu + Pw + M\nu \\ y &= Cx + Qw + R\nu \end{aligned} \tag{7.1}$$

where $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^m$ the control input, $y \in \mathbb{R}^{n_y}$ the measured outputs and $(w, \nu) \in \mathbb{R}^{n_w} \times \mathbb{R}^{n_\nu}$ represent exogenous signals acting on the system, such as references to be tracked and disturbances to be rejected. In particular, $w(t)$ represents the "modeled' part of the exogenous signals, and we suppose that it belongs to the family of solutions of an *exosystem* of the form

$$\dot{w} = Sw, \tag{7.2}$$

with *unknown* state matrix S and with a dimension $n_w \in \mathbb{N}$ that is upperbounded by a *known* integer $d \in \mathbb{N}$. The signal $\nu(t)$, instead, is a bounded hybrid input representing unknown unmodeled disturbances, i.e. disturbances acting on the plant whose nature is not known and that are not supposed to be generated by any external process. We associate to (7.1) a further set of outputs $e \in \mathbb{R}^{n_e}$, $n_e \in \mathbb{N}$, defined as

$$e = C_e x + Q_e w + N_e \nu. \tag{7.3}$$

We refer to the quantity e as the *regulation errors*. They represent those outputs on which the effect of the exosystem must be ideally removed such as, for instance, tracking errors. More precisely, we seek an output feedback regulator of the form

$$\begin{array}{rcll} \dot{x}_c & \in & F_c(x_c, y) & (x_c, y) \in C_c \\ x_c^+ & \in & G_c(x_c, y) & (x_c, y) \in D_c \\ u & = & \gamma(x_c, y)\,, & \end{array} \tag{7.4}$$

with state x_c taking values in an Euclidean space $\mathcal{X}_c$, ad where C_c, $D_c \subset \mathcal{X}$, such that all the solutions to the closed-loop system (7.1), (7.4) are bounded and it has the *ε-approximate regulation property* for some $\varepsilon > 0$, i.e. any solution also satisfies

$$\limsup |e| \leq \varepsilon.$$

We say that (7.1), (7.4) has the *asymptotic regulation property* if it has the 0-approximate regulation property. On the system (7.1), (7.3) we make the following assumptions:

Assumption 7.1. *The regulation errors are included in the measured output, i.e.* $y = \mathrm{col}(e, y_m)$ *for some* $y_m \in \mathbb{R}^{n_m}$*,* $n_m := n_y - n_e$*.*

Assumption 7.2. (A, B) *is stabilizable,* (C, A) *is detectable,* $\mathrm{rank}\, B = m \geq \mathrm{rank}\, C_e = n_e$*.*

Assumption 7.3. *The solutions of* (7.2) *range in a compact set* $W \subset \mathbb{R}^{n_w}$ *and* $|W|$ *is known.*

Assumptions 7.1 and 7.2 are close to being necessary if a *robust* asymptotic regulation result is sought. As a matter of fact, *readability* in the sense of (Francis and Wonham, 1975) of e from y is proved in (Francis and Wonham, 1975) to be necessary to obtain a structurally stable solution in the case in which (7.2) is known. If readability holds, on the other hand, we can always change coordinates to have Assumption 7.1 fulfilled. Furthermore, as in the classical solution of (Davison, 1976), we will augment the plant (7.1) with an *internal model unit* that is driven by the regulation errors. Then Assumption 7.2 turns out to be necessary to have stabilizability and detectability of the resulting cascade. Regarding Assumption 7.3, this assumption limits the size of the initial conditions of w and requires S to be stable though not typically Hurwitz. The set W can be arbitrarily large as soon as $|W|$ is known. This latter quantity represents a constant that must be dominated by some control parameters. Thus, in this sense, the forthcoming result could also be rephrased by fixing the control parameters and adjusting the “admissible” W accordingly.

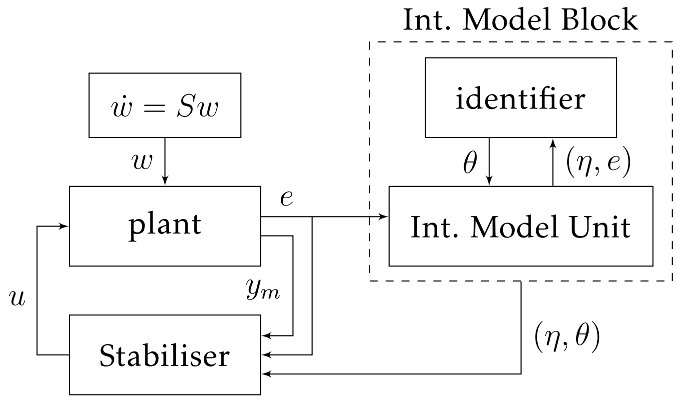

Figure 7.1: Block-diagram of the closed-loop system.

7.3 The Regulator Structure

In this section we construct a regulator of the form (7.4) that guarantees closed-loop stability and, under suitable persistence of excitation, has the ε-approximate regulation property with ε that is proportional to $\limsup |\nu|$. The regulator consists of two main blocks: the *internal model block* and the *stabilizer* (see Figure 7.1). The internal model block is itself composed of two subsystems: the *internal model* and the *identifier*. The internal model is a system driven by the regulation errors and it replicates the structure proposed by Davison in (Davison, 1976). The spectrum of the internal model's state matrix is adapted by the identifier to match the modes of the unknown exosystem (7.2). The identifier is a discrete-time system built to solve asymptotically an optimization problem defined on the time evolution of the state of the internal model and the regulation errors. Under suitable persistence of excitation and if $\nu = 0$ it turns out that the identifier optimal trajectory is uniquely determined and the corresponding internal model includes all the exosystem's modes, so as asymptotic regulation is achievable. The stabilizer is a subsystem that processes all the measured signals and robustly stabilizes the cascade of the plant and the internal model block. In the rest of the section we detail all these three components.

7.3.1 The Internal Model Unit

With d any known upper bound on $n_w = \dim w$, the internal model unit is a system with state $\eta \in \mathbb{R}^{n_e d}$ satisfying the following equations

$$\begin{aligned} \dot{\eta} &= \Phi(\theta)\eta + Ge \\ \eta^+ &= \eta \end{aligned} \tag{7.5}$$

where

$$\Phi(\theta) := \begin{pmatrix} 0_{n_e(d-1)\times n_e} & I_{n_e(d-1)} \\ \multicolumn{2}{c}{\theta^T \otimes I_{n_e}} \end{pmatrix} \quad G := \begin{pmatrix} 0_{n_e(d-1)\times n_e} \\ I_{n_e} \end{pmatrix}$$

and with $\theta \in \mathbb{R}^d$ that is a parameter adapted by the identifier. The characteristic polynomial of $\Phi(\theta)$ reads as

$$\varphi_{\Phi(\theta)}(\lambda) = \left(\lambda^d - \theta_d \lambda^{d-1} - \cdots - \theta_2 \lambda - \theta_1\right)^{n_e}, \tag{7.6}$$

so as, if S were known, the internal model of (Davison, 1976) could be obtained by letting the components of θ to be the coefficients of any polynomial that has the eigenvalues of S as roots. We then let $\mathcal{Q}$ be the set

$$\mathcal{Q} := \left\{ \theta \in \mathbb{R}^d \ : \text{rank} \begin{pmatrix} A - \lambda I & B \\ C_e & 0 \end{pmatrix} < n + n_e,\ \lambda \in \sigma(\Phi(\theta)) \right\},$$

and we pick a compact set satisfying

$$\mathcal{E} \subset \mathbb{R}^d \setminus \mathcal{Q}.$$

The existence of a non-empty set $\mathcal{E}$ satisfying the above property, which in general has to be assumed, is in turn necessary for the solvability of the problem. As a matter of fact, as shown for instance in (Isidori, 2017, Lem. 4.1), for the regulation problem to have a solution for a given exosystem (7.2), $\mathbb{R}^n \setminus \mathcal{Q}$ must contain at least one $\theta \in \mathbb{R}^d$ such that $\sigma(S) = \sigma(\Phi(\theta))$. Moreover, we have the following sufficient condition.

Lemma 7.1. *Suppose that Assumptions 7.1-A7.2 hold and assume that the set of $\lambda \in \mathbb{C}$ for which the transfer function $C_e(A - \lambda I)^{-1}B$ loses rank is finite. Then for each $r \geq 0$, $\mathbb{R}^d \setminus (\mathcal{Q} + \mathbb{B}r)$ is closed and not empty.*

Proof. Let

$$F(\lambda) := \begin{pmatrix} A - \lambda I & B \\ C_e & 0 \end{pmatrix}.$$

We first prove that the set $\mathcal{U} := \{\lambda \in \mathbb{C} \, : \, \operatorname{rank} F(\lambda) < n + n_e\}$ is finite. Pick $\lambda \in \mathbb{C}/\sigma(A)$. Then $A - \lambda I$ is invertible and the matrix

$$M(\lambda) := \begin{pmatrix} I_n & 0_{n \times n_e} \\ C_e(A - \lambda I)^{-1} & -I_{n_e} \end{pmatrix}$$

is well-defined and full rank. Thus

$$\begin{aligned} \operatorname{rank} F(\lambda) &= \operatorname{rank}(M(\lambda)F(\lambda)) \\ &= \operatorname{rank} \begin{pmatrix} A - \lambda I & B \\ 0_{n_e} & C_e(A - \lambda I)^{-1}B \end{pmatrix} \\ &\geq n + \operatorname{rank}(C_e(A - \lambda I)^{-1}B). \end{aligned}$$

By assumption, as $\operatorname{rank} B \geq n_e$, the number of $\lambda \in \mathbb{C}$ for which $\operatorname{rank}(C_e(A - \lambda I)^{-1}B) < n_e$ is finite. Since $\sigma(A)$ has at most n elements, then we conclude that $\mathcal{U}$ is finite. Let $\mathcal{P}^d$ denote the set of monic polynomials of degree d with coefficients in $\mathbb{R}$. Each element of $\mathcal{P}^d$ can be written as $p_a(s) = s^d + a_d s^{d-1} + \cdots + a_2 s + a_1$, for some $a \in \mathbb{R}^d$, so as there is a natural isomorphism $\iota : \mathbb{R}^d \to \mathcal{P}^d$, $a = \operatorname{col}(a_1, \ldots, a_d) \mapsto p_a(s)$. Let $P(\lambda) \subset \mathcal{P}^d$ denote the set of polynomials in $\mathcal{P}^d$ that have λ as a root. Suppose $\lambda \in \mathbb{R}$, each element in $P(\lambda)$ can be univocally written as

$$\begin{aligned} p_a(s) &= (s - \lambda)(s^{d-1} + a_{d-1}s^{d-2} + \cdots + a_2 s + a_1) \\ &= s^d + (a_{d-1} - \lambda)s^{d-1} + \cdots + (a_1 - a_2\lambda)s - a_1\lambda, \end{aligned}$$

so as $\iota^{-1}(P(\lambda))$ is the $(d-1)$-dimensional affine subspace of $\mathbb{R}^d$ given by $\iota^{-1}(P(\lambda)) = \{(-\lambda a_1, a_1 - \lambda a_2, \ldots, a_{d-2} - \lambda a_{d-1}, a_{d-1} - \lambda) \in \mathbb{R}^d \, : \, (a_1, \ldots, a_{d-1}) \in \mathbb{R}^{d-1}\}$. If instead $\lambda \in \mathbb{C}/\mathbb{R}$, in the same way as before it can be seen that $\iota^{-1}(P(\lambda))$ is a $(d-2)$-dimensional affine subspace of $\mathbb{R}^d$. As $\mathcal{Q}$ can be written as $\mathcal{Q} = \{\theta \in \mathbb{R}^d \, : \, \iota(\theta) \in P(\lambda), \, \lambda \in \mathcal{U}\} = \cup_{\lambda \in \mathcal{U}} \iota^{-1}(P(\lambda))$, then $\mathcal{Q}$ is the union of a finite number of affine subspaces of $\mathbb{R}^d$ of dimension $d-1$ or $d-2$. Pick $r > 0$ arbitrarily and notice that the set $\mathcal{Q} + \mathbb{B}r$ is open as it is a union of open sets. Hence $\mathbb{R}^d \setminus (\mathcal{Q} + \mathbb{B}r)$ is closed and it remains to show that it is not empty. We thus notice that for each $\lambda \in \mathcal{U}$ we can write $\iota^{-1}(P(\lambda)) = x_\lambda + \operatorname{Im} A_\lambda$, for some $x_\lambda \in \mathbb{R}^d$ and $A_\lambda \in \mathbb{R}^{d\times(d-1)}$. Pick any $\lambda \in \mathcal{U}$ and define the set $S := \{x_\lambda + y + v \in \mathbb{R}^{d-1} : y \in \operatorname{Im} A_\lambda,\ v \in (\operatorname{Im} A_\lambda)^\perp\}$. We now show that $S \cap (\mathbb{R}^d \setminus (\mathcal{Q} + \mathbb{B}r)) \neq \emptyset$. For, let $\mathcal{U}_\lambda \subset \mathcal{U}$ be the set of $\mu \in \mathcal{U}$ such that $\operatorname{Im} A_\lambda \subseteq \operatorname{Im} A_\mu$. Then for each $\mu \in \mathcal{U}_\lambda$, each $y \in \operatorname{Im} A_\lambda$, and with $p = x_\lambda + y + v \in S$, in view of (Deutsch, 2001, Thm. 4.9), we have

$$|p|_{\iota^{-1}(P(\mu))} = \inf_{z \in \operatorname{Im} A_\mu} |x_\mu + z - x_\lambda - y - v| = |(x_\mu - x_\lambda)' - v'|$$

where $(x_\mu - x_\lambda)'$ and v' denote the projection of $x_\mu - x_\lambda$ and v onto $(\operatorname{Im} A_\mu)^\perp$. Hence, for each $v \in (\operatorname{Im} A_\lambda)^\perp$ fulfilling $|v'| > r + \max_{\mu \in \mathcal{U}_\lambda} |x_\mu - x_\lambda|$ we obtain $|p|_{\iota^{-1}(P(\mu))} \geq |v'| - |(x_\mu - x_\lambda)'| > r$, and this in turn shows that, for every $y \in \operatorname{Im} A_\lambda$ and for sufficiently large v, $p = x_\lambda + y + v \in \mathbb{R}^d \setminus (\cup_{\mu \in \mathcal{U}_\lambda} \iota^{-1}(P(\mu)) + \mathbb{B}r)$. Pick now $\mu \in \mathcal{U} \setminus \mathcal{U}_\lambda$ and fix a $v \in (\operatorname{Im} A_\lambda)^\perp$ satisfying the above bound, then $\operatorname{Im} A_\lambda \cap (\operatorname{Im} A_\mu)^\perp \neq \emptyset$ and we get

$$|p|_{\iota^{-1}(P(\mu))} = \inf_{z \in \operatorname{Im} A_\mu} |x_\mu + z - x_\lambda - y - v| = |(x_\mu - x_\lambda)'' - y'' - v''|$$

with $(x_\mu - x_\lambda)''$, y'' and v'' the projections of $x_\mu - x_\lambda$, y and v onto $(\operatorname{Im} A_\mu)^\perp$. Hence choosing y so that $|y''| > r + \max_{\mu \in \mathcal{U}\setminus\mathcal{U}_\lambda} |x_\mu - x_\lambda| + |v''|$ yields $|p|_{\iota^{-1}(P(\mu))} > r$, and this in turn proves that there exists a $p \in S$ satisfying $p \in \mathbb{R}^d \setminus (\mathcal{Q} + \mathbb{B}r)$. Hence the claim. ■

Lemma 7.1 implies that for any $r > 0$ and any sufficiently large compact set $\Theta \subset \mathbb{R}^d$, the set $\{\theta \in \mathbb{R}^d : |\theta|_{\mathcal{Q}} \geq r\} \cap \Theta$ is compact and not empty, thus qualifying as a possible choice of $\mathcal{E}$. The definition of $\mathcal{E}$ is justified by the following fact:

Lemma 7.2. *Let* $\operatorname{rank} B \geq \operatorname{rank} C_e = n_e$. *Then the pair*

$$\left(\begin{pmatrix} A & 0 \\ GC_e & \Phi(\theta) \end{pmatrix}, \begin{pmatrix} B \\ 0 \end{pmatrix} \right) \tag{7.7}$$

is stabilizable/controllable for all $\theta \in \mathcal{E}$ *if and only if* (A, B) *is stabilizable/controllable. Moreover, the pair*

$$\left(\begin{pmatrix} C & 0 \\ 0 & I \end{pmatrix}, \begin{pmatrix} A & 0 \\ GC_e & \Phi(\theta) \end{pmatrix} \right)$$

is detectable/observable for all $\theta \in \mathbb{R}^d$ if and only if (C, A) is detectable/observable.

Lemma 7.2 is based on the PBH test (Isidori, 2017), and its proof is a straightforward consequence of the definition of $\mathcal{Q}$ and it is thus omitted. In the forthcoming sections we will force θ to range in the set $\mathcal{E}$. Although this limits the number of internal models that can be eventually implemented, it also guarantees that the cascade of the plant and the internal model is stabilizable independently of the identifier trajectories.

7.3.2 The Identifier

The identifier subsystem is a multivariable version of the discrete-time least-squares identifier introduced in Section 4.2. We will recall its construction hereafter. The identifier measures two inputs, α and β, at the jump times and it tries to infer a model relating the two inputs on the basis of the observed samples. The model is a linear regression of order d and the parameter is chosen to minimize a cost function that weights a sum of historical prediction errors produced by the candidate model. At the end of the section we will interconnect the identifier with the internal model, obtaining a hybrid system that acts as a feedforward generator during flows and optimizes its internal model during jumps.

In the following we identify $(\mathbb{R}^{n_e})^d$ with $\mathbb{R}^{n_e d}$ and, for each of its elements p, we let $p_1, \dots, p_d$ denote the elements of $\mathbb{R}^{n_e}$ such that $p = \operatorname{col}(p_1, \dots, p_d)$. Then we define the matrix

$$\gamma(p) := \begin{pmatrix} p_1 & p_2 & \cdots & p_d \end{pmatrix}^T \in \mathbb{R}^{d \times n_e}.$$

Let α and β be hybrid inputs taking values in $\mathbb{R}^{n_e d}$ and $\mathbb{R}^{n_e}$ respectively. We define the identifier subsystem as a discrete-time system defined on the state space $\mathcal{Z} \times \mathbb{R}^d$, where $\mathcal{Z} := \mathbb{R}^{d\times d} \times \mathbb{R}^d$, and with state $z := (R, v) \in \mathcal{Z}$ and $\theta \in \mathbb{R}^d$ satisfying the following equations

$$\left\{\begin{array}{rcl} \dot{R} & = & 0 \\ \dot{v} & = & 0 \\ \dot{\theta} & = & 0 \end{array}\right. \quad \left\{\begin{array}{rcl} R^+ & = & \mu R + \gamma(\alpha)\gamma(\alpha)^T \\ v^+ & = & \mu v + \gamma(\alpha)\beta \\ \theta^+ & \in & p_{\mathcal{E}}(R^\dagger v), \end{array}\right. \tag{7.8}$$

with output θ, where $\cdot^\dagger$ denotes the Moore-Penrose pseudoinverse, $\mu \in (0, 1)$ is a design parameter and $p_{\mathcal{E}}(\cdot)$ is the projection map $\theta \mapsto p_{\mathcal{E}}(\theta) := \arg\inf_{\theta_{\mathcal{E}} \in \mathcal{E}} |\theta - \theta_{\mathcal{E}}|$. We endow $\mathcal{Z}$ with the norm $|(R, v)| := \sqrt{|R|^2 + |v|^2}$.

We recall that the identifier (7.8) is constructed to asymptotically find the "best" linear model relating the *regressor* input α and the input β. More precisely, we associate to (7.8) the *prediction model* $\hat{\beta} : \mathcal{E} \times \mathbb{R}^{n_e d} \to \mathbb{R}^{n_e}$ given by

$$\hat{\beta}(\theta, \alpha) := (\theta^T \otimes I_{n_e})\alpha = \sum_{i=1}^{d} \theta_i \alpha^i \tag{7.9}$$

and the corresponding *prediction error* $\varepsilon : \mathcal{E} \times \mathbb{R}^{n_e d} \times \mathbb{R}^{n_e} \to \mathbb{R}^{n_e}$ given by

$$\varepsilon(\theta, \alpha, \beta) := \beta - \hat{\beta}(\theta, \alpha) = \beta - (\theta^T \otimes I_{n_e})\alpha.$$

For fixed θ, the prediction model $\hat{\beta}(\theta, \alpha)$ represents the identifier's guess of β given α. The identifier is constructed to choose θ so that the guess $\hat{\beta}(\theta, \alpha)$ is the best possible among all those producible by d-dimensional linear models of the form (7.9). As in Section 4.1, "best" is defined relatively to a cost functional that weights the prediction performance of the candidate models on all the historical data. More precisely, with $\mathcal{A}(\mathbb{R}^d, \mathbb{R}_+)$ the set of functions $\mathbb{R}^d \to \mathbb{R}_+$, we associate to each input (α, β) a function $\mathcal{J}_{\alpha,\beta} : \mathrm{dom}(\alpha, \beta) \to \mathcal{A}(\mathbb{R}^d, \mathbb{R}_+)$ defined by

$$\mathcal{J}_{\alpha,\beta}(\theta)(t, j) := \sum_{i=0}^{j-1} \mu^{j-1-i} \left| \varepsilon\big(\theta, \alpha(t^i, i), \beta(t^i, i)\big) \right|^2 . \tag{7.10}$$

At a given $(t, j) \in \mathrm{dom}(\alpha, \beta)$, the best linear model is the one given by (7.9) with θ minimizing $\mathcal{J}_{\alpha,\beta}(t, j)$. We associate to (7.10) the following (set-valued) map

$$\theta^{\circ}_{\alpha,\beta}(t, j) := \arg\min_{\theta \in \mathbb{R}^d} \mathcal{J}_{\alpha,\beta}(\theta)(t, j) , \tag{7.11}$$

whose value at each (t, j), contains the "optimal" parameters θ that minimize (7.10).

We recall that the intuition behind the definition of the identifier (7.8), in relation to the minimization problem (7.10), resides in the fact that the optimal trajectory (7.11) can be proved to satisfy

$$\theta^{\circ}_{\alpha,\beta}(t, j) = \big\{ \theta \in \mathbb{R}^d \,:\, R^\star(t, j)\theta = v^\star(t, j) \big\} , \tag{7.12}$$

being

$$\begin{aligned} R^\star(t,j) &:= \sum_{i=0}^{j-1} \mu^{j-i-1}\gamma(\alpha(t^i,i))\gamma(\alpha(t^i,i))^T \\ v^\star(t,j) &:= \sum_{i=0}^{j-1} \mu^{j-i-1}\gamma(\alpha(t^i,i))\beta(t^i,i)^T\,. \end{aligned} \tag{7.13}$$

As stated in more general terms in the forthcoming proposition, it can be shown that $z^\star := (R^\star, v^\star)$ is a solution to the subsystem z of (7.8) which is also (robustly) asymptotically stable. It is worth noting, moreover, that while the definition of (7.10) requires in principle the knowledge of an unbounded number of samples, in view of (7.13), the information that is necessary to define (7.12) can be encoded in the *finite* dimensional quantities $R^\star$ and $v^\star$, and this permits to track the optimal trajectory (7.11) with a finite-dimensional system (as it is (7.8)).

When we will interconnect the identifier and the internal model, the inputs (α, β) will be set to some functions of the state η and the regulation error e. These signals, in turn, carry some "ideal" information about $w(t)$, that is useful to infer the right model, corrupted by some additional disturbances given by transitory artifacts and residual noise dependent on $\nu(t)$. To take into account this situation in the characterization of the properties of the identifier (7.8), we consider inputs (α, β) given by $\alpha = \alpha^* + \delta_\alpha$, $\beta = \beta^* + \delta_\beta$, with (α^*, β^*) an *ideal* input and $\delta = (\delta_\alpha, \delta_\beta)$ an additive disturbance, thus fitting exactly in the framework of Section 4.1. The following proposition expresses the optimality of the identifier (7.8) with respect to the cost functional $\mathcal{J}_{\alpha^\star,\beta^\star}(\theta)(t,j)$ when $\delta = 0$ and the robust stability properties of the optimal trajectory in case of $\delta \neq 0$.

Proposition 7.1. *The following hold:*

1. *For each* $z_1 := (R_1, v_1), z_2 := (R_2, v_2) \in \mathcal{Z}$ *and each* $(\alpha_1, \beta_1), (\alpha_2, \beta_2) \in \mathbb{R}^{dn_e} \times \mathbb{R}^{n_e}$ *there exists* $\rho > 0$ *such that, with* $(\delta_\alpha, \delta_\beta) := (\alpha_1, \beta_1) - (\alpha_2, \beta_2)$ *and, for* $i = 1, 2$, $z_i^+ := (\mu R_i + \gamma(\alpha_i)\gamma(\alpha_i)^T, \mu v_i + \gamma(\alpha_i)\beta_i)$, *it holds that*

$$|z_1^+ - z_2^+|^2 \le \mu^2 |z_1 - z_2|^2 + \rho |(\delta_\alpha, \delta_\beta)|^2,$$

where, for $i = 1, 2$, *we let* $z_i^+ := (\mu R_i + \gamma(\alpha_i)\gamma(\alpha_i)^T, \mu v_i + \gamma(\alpha_i)\beta_i)$.

2. *For each hybrid input* $(\alpha^\star, \beta^\star) : \mathrm{dom}(\alpha^\star, \beta^\star) \to \mathbb{R}^{dn_e} \times \mathbb{R}^{n_e}$ *there exists* $z^\star := (R^\star, v^\star) : \mathrm{dom}(\alpha^\star, \beta^\star) \to \mathcal{Z}$ *such that* $(z^\star, (\alpha^\star, \beta^\star))$ *is a solution pair to the subsystem* z *to* (7.8) *and the corresponding "unconstrained output"* $\theta^\star_{un} := (R^\star)^\dagger v^\star$

satisfies

$$\theta^\star_{un}(t,j) \in \theta^\circ_{\alpha^\star,\beta^\star}(t,j).$$

The proof of Proposition 7.1 follows directly from the arguments used to prove Proposition 4.1. Although Proposition 7.1 guarantees that for every input $(\alpha,\beta) = (\alpha^\star,\beta^\star) + (\delta_\alpha,\delta_\beta)$ there exists a "optimal trajectory" $(z^\star,\theta^\star_{un})$ such that, for any selection $\theta^\star$ of $p_{\mathcal{E}}(\theta^\star_{un})$, $((z^\star,\theta^\star),(\alpha^\star,\beta^\star))$ solves (7.8) and is asymptotic stable (with $\delta = 0$), it is not in general true that $z \to z^\star$ implies $\theta \to \theta^\star$, due to the pseudoinverse operator that is not, in general, continuous. To recover this "detectability" property, as well as to ensure single-valuedness of (7.11), we associate to the input α the following *persistence of excitation* condition:

Definition 7.1. *With $J \in N$ and $\epsilon > 0$, a complete hybrid input $\alpha : \operatorname{dom}\alpha \to \mathbb{R}^{n_e d}$ is said to be (J,ϵ)*-persistently exciting *if, for all integers $j \geq J$*

$$\min\sigma\left(\sum_{i=0}^{j-1}\mu^{j-1-i}\gamma(\alpha(t^i,i))\gamma(\alpha(t^i,i))^T\right) \geq \epsilon\,. \tag{7.14}$$

Remark 7.1. This definition of persistence of excitation is stronger than those given in Section 4.2. As a matter of fact, in Definition 4.2, the matrix involved in the PE condition was allowed not to be full rank. Instead (7.14) implies that, for sufficiently large times, the regressor matrix is positive definite. This tighter condition is motivated by Lemma 7.3, in which (7.14) is proved to imply that the solution map (7.12) is single valued. △

In the following we will often abbreviate "(J,ϵ)-persistently exciting" with "(J,ϵ)-PE". Lemma 7.3 relates persistence of excitation of α with those of $\alpha^\star$ and, thus, with single-valuedness of the map $\theta^\circ_{\alpha^\star,\beta^\star}(t,j)$, when the disturbance δ_α is small enough at the jump times. Lemma 7.4, instead, links persistence of excitation and "detectability" from the output θ.

Lemma 7.3. *Let α, $\alpha^\star$ and δ_α be bounded hybrid inputs defined over the same time domain and such that $\alpha = \alpha^\star + \delta_\alpha$. Then for any $\epsilon > 0$ there exists $\bar\delta > 0$ such that, if α is (J,ϵ)-PE for some $J \in \mathbb{N}$ and $|\delta_\alpha(t^j,j)| \leq \bar\delta$ for all $j \geq J$, then there exists $(J',\epsilon') \in \mathbb{N}\times\mathbb{R}^*_+$ such that $\alpha^\star$ is (J',ϵ')-PE. Moreover, $\theta^\circ_{\alpha^\star,\beta^\star}(t,j)$ is a singleton for all $(t,j) \in \operatorname{dom}\alpha^\star$ such that $j \geq J'$.*

Proof. Let ℓ_∞ be the space of bounded sequences $s = (s_n)_{n\in\mathbb{N}}$ and, for $s \in \ell_\infty$, let $|s|_{n_1,n_2} := \sup_{n_1\leq n\leq n_2}|s_n|$. Let $\phi(\alpha) := \operatorname{col}(\alpha(t^0,0),\alpha(t^1,1),\dots) \in \ell_\infty$ and, for

$k, j \in \mathbb{N}$, let $\Sigma_k^j : \ell_\infty \to \mathbb{R}^{d\times d}$ be the function

$$s \mapsto \Sigma_k^j(s) := \sum_{i=k}^{j-1} \mu^{j-i-1}\gamma(s_i)\gamma(s_i)^T .$$

For any two $q_1, q_2 \in \mathbb{R}^{n_e d}$ there exists $c_0 > 0$ such that $|\gamma(q_1)\gamma(q_1)^T - \gamma(q_2)\gamma(q_2)^T| \leq c_0|q_1 - q_2|$. As a consequence, for each two $s^1, s^2 \in \ell_\infty$ there exists $c_0 > 0$ such that, for each $k, j \in \mathbb{N}$, we have

$$\begin{aligned}|\Sigma_k^j(s_1) - \Sigma_k^j(s_2)| &= \sum_{i=k}^{j-1} \mu^{j-i-1}(\gamma(s_i^1)\gamma(s_i^1)^T - \gamma(s_i^2)\gamma(s_i^2)^T) \\ &\leq \left(\sum_{i=k}^{j-1}\mu^{j-i-1}\right) c_0|s^1 - s^2|_{k,\infty}.\end{aligned}$$

Noting that

$$\sum_{i=k}^{j-1}\mu^{j-i-1} \leq \sum_{i=0}^{j-1}\mu^{j-i-1} = \sum_{\ell=0}^{j-1}\mu^\ell \leq \sum_{\ell=0}^{\infty}\mu^\ell = \frac{1}{1-\mu},$$

we thus obtain

$$|\Sigma_k^j(s_1) - \Sigma_k^j(s_2)| \leq c_1|s^1 - s^2|_{k,\infty}, \tag{7.15}$$

for each $k, j \in \mathbb{N}$ and with $c_1 := c_0/(1-\mu)$. As the map $R \in \mathbb{R}^{n_e d\times n_e d} \mapsto \min\sigma(R) \in \mathbb{R}_+$ is continuous, for every $\upsilon > 0$ there exists $r_\upsilon > 0$ such that

$$\begin{aligned}&|\Sigma_0^j(\phi(\alpha)) - \Sigma_0^j(\phi(\alpha^\star))| \leq r_\upsilon \\ &\quad\implies |\min\sigma(\Sigma_0^j(\phi(\alpha))) - \min\sigma(\Sigma_0^j(\phi(\alpha^\star)))| \leq \upsilon.\end{aligned} \tag{7.16}$$

As Σ_0^j has symmetric positive semi-definite values then

$$\begin{aligned}&\min\sigma(\Sigma_0^j(\phi(\alpha^\star))) = |\min\sigma(\Sigma_0^j(\phi(\alpha^\star)))| \\ &= \big|\min\sigma(\Sigma_0^j(\phi(\alpha))) - \big(\min\sigma(\Sigma_0^j(\phi(\alpha))) - \min\sigma(\Sigma_0^j(\phi(\alpha^\star)))\big)\big| \\ &\geq \big||\min\sigma(\Sigma_0^j(\phi(\alpha)))| - |\min\sigma(\Sigma_0^j(\phi(\alpha))) - \min\sigma(\Sigma_0^j(\phi(\alpha^\star)))|\big|.\end{aligned} \tag{7.17}$$

Noting that $\Sigma_0^j(\phi(\alpha))$ is exactly the matrix appearing in (7.14), and since α is (J,ϵ)-PE, then $\min\sigma(\Sigma_0^j(\phi(\alpha))) > \epsilon$. Pick $\epsilon' < \epsilon$ and $\upsilon \in (0, \epsilon - \epsilon')$ arbitrarily. Thus if for some $J' \geq J$ we have $|\Sigma_0^j(\phi(\alpha)) - \Sigma_0^j(\phi(\alpha^\star))| \leq r_\upsilon$ for all $j \geq J'$, then (7.16) and (7.17) give

$$\min \sigma(\Sigma_0^j(\phi(\alpha^\star))) \geq \epsilon - \upsilon \geq \epsilon',$$

for all $j \geq J'$, that is the first claim. We thus have to show that (7.16) holds for sufficiently small $\bar{\delta}$. Note that, for each $j > J$

$$|\Sigma_0^j(\phi(\alpha)) - \Sigma_0^j(\phi(\alpha^\star))| \leq \mu^{j-J}|\Sigma_0^J(\phi(\alpha)) - \Sigma_0^J(\phi(\alpha^\star))| + |\Sigma_J^j(\phi(\alpha)) - \Sigma_J^j(\phi(\alpha^\star))|. \tag{7.18}$$

As the first term of (7.18) is a constant multiplied by μ^j, and $\mu < 1$, we claim the existence of $J' \geq J$ such that in (7.18) we have that $\mu^{j-J}|\Sigma_0^J(\phi(\alpha)) - \Sigma_0^J(\phi(\alpha^\star))| \leq r_\upsilon/2$ for all $j \geq J'$. Moreover, in view of (7.15), $|\phi(\delta_\alpha)|_{J,\infty} = |\phi(\alpha) - \phi(\alpha^\star)|_{J,\infty} \leq r_\upsilon/(2c_1)$ implies $|\Sigma_J^j(\phi(\alpha)) - \Sigma_J^j(\phi(\alpha^\star))| \leq r_\upsilon/2$ for all $j \geq J$. Thus (7.18) yields (7.16) for all $j \geq J'$. This in turn proves the first claim, with $\bar{\delta} := r_\upsilon/(2c_1)$. To see that $\theta^\circ_{\alpha^\star,\beta^\star}(t,j)$ is single valued, notice that it is the set of θ in which the gradient of (7.10) vanishes, which is given by (7.12)-(7.13) with $(\alpha^\star, \beta^\star)$ in place of (α, β). In view of (7.13), $v^\star(t,j) \in \operatorname{Im} R^\star(t,j)$. Thus, noting that for $j \geq J'$, $R^\star(t,j) = \Sigma_0^j(\phi(\alpha^\star))$ is nonsingular, then (7.12) is a singleton and the second claim follows. ■

Lemma 7.4. *Let $(z_1, (\alpha_1, \beta_1))$ and $(z_2, (\alpha_2, \beta_2))$ be solution pairs to (7.8) with the same time domain. Suppose that, for $i = 1, 2$, α_i is (J_i, ϵ_i)-PE, for some $(J_i, \epsilon_i) \in \mathbb{N} \times \mathbb{R}^*_+$ and let $\theta_i := p_\mathcal{E}(R_i(t,j)^\dagger v_i(t,j))$. Then there exists $J \in \mathbb{N}$ and $a \geq 0$ such that*

$$|\theta_1(t,j) - \theta_2(t,j)|^2 \leq a|z_1(t,j) - z_2(t,j)|^2,$$

for all $(t,j) \in \operatorname{dom} z$ satisfying $j \geq J$.

Proof. For sake of readability, we will omit the time dependency. By letting $\theta_i^u(t,j) := R_i(t,j)^\dagger v_i(t,j)$, then for suitable selections $s_\mathcal{E}^1$ and $s_\mathcal{E}^2$ of $p_\mathcal{E}$, we obtain

$$\begin{aligned}
|\theta_1 - \theta_2|^2 &= |s_\mathcal{E}^1(\theta_1^u) - s_\mathcal{E}^2(\theta_2^u)|^2 = |s_\mathcal{E}^1(\theta_1^u) - \theta_1^u + \theta_1^u - \theta_2^u + \theta_2^u - s_\mathcal{E}^2(\theta_2^u)|^2 \\
&\leq \inf_{\theta_\mathcal{E} \in \mathcal{E}} |\theta_1^u - \theta_\mathcal{E}|^2 + \inf_{\theta_\mathcal{E} \in \mathcal{E}} |\theta_2^u - \theta_\mathcal{E}|^2 + |\theta_1^u - \theta_2^u|^2 \\
&\leq 3|\theta_1^u - \theta_2^u|^2 = 3|R_1^\dagger v_1 - R_2^\dagger v_2|^2 \\
&\leq 3|R_1^\dagger - R_2^\dagger|^2 |v_1|^2 + 3|R_2^\dagger|^2 |v_1 - v_2|^2.
\end{aligned}$$

In view of (Campbell and Meyer, 2009, Thm. 10.4.5), we have $|R_1^\dagger - R_2^\dagger|^2 \leq$

$9\max\{|R_1^\dagger|^4, |R_2^\dagger|^4\}|R_1 - R_2|^2$, which yields

$$|\theta_1 - \theta_2|^2 \\ \leq 3\max\{9\max\{|R_1^\dagger|, |R_2^\dagger|\}^4|v_1|^2, |R_2^\dagger|^2\}|z_1 - z_2|^2 \tag{7.19}$$

By direct solution we obtain, for $i = 1, 2$

$$R_i(t,j) = \mu^j R_i(0,0) + \sum_{k=0}^{j-1} \mu^{j-1-k}\gamma(\alpha_i(t^k,k))\gamma(\alpha_i(t^k,k))^T. \tag{7.20}$$

Since $\mu < 1$ the first term of (7.20) vanishes exponentially with j. Thus using the fact that α_i is (J_i, ϵ_i)-PE, the same arguments of Lemma 7.3 can be used to show that (7.20) implies that for any $\epsilon_i' \in (0, \epsilon_i)$, there exists $J_i' \geq J_i$ such that, for all $(t,j) \in \operatorname{dom} R_i$ such that $j \geq J_i'$, $\min \sigma(R_i(t,j)) \geq \epsilon_i'$. As a consequence, by letting $J := \max\{J_1', J_2'\}$ and $\epsilon := \min\{\epsilon_1', \epsilon_2'\}$, we obtain $|R_i(t,j)^\dagger| \leq 1/\epsilon$ for all $(t,j) \in \operatorname{dom} R$ such that $j \geq J$. Thus the result follows from (7.19) by noting that boundedness of (α_i, β_i) for $i = 1, 2$ implies those of v_i. ■

With the above definitions in mind, we interconnect the identifier (7.8) and the internal model (7.5) by letting θ in (7.5) be the same state variable of (7.8), and by letting in (7.8) $\alpha = \eta$ and $\beta = G^T\dot{\eta} = (\theta^T \otimes I_{n_e})\eta + e$, i.e.

$$\begin{cases} \dot{R} &= 0 \\ \dot{v} &= 0 \\ \dot{\theta} &= 0 \end{cases} \quad \begin{cases} R^+ &= \mu R + \gamma(\eta)\gamma(\eta)^T \\ v^+ &= \mu v + \gamma(\eta)\big((\theta^T \otimes I_{n_e})\eta + e\big) \\ \theta^+ &\in p_{\mathcal{E}}(R^\dagger v). \end{cases} \tag{7.21}$$

7.3.3 The Stabilizer

The stabilizer is defined as the composition of a continuous-time output feedback controller for the cascade (x, η, z) and a clock subsystem that activates the update law of the identifier. It reads as follows:

$$
\begin{array}{l}
\left\{\begin{array}{rcl}\dot{\tau} & = & 1\\ \dot{\xi} & = & H_\xi(\theta)\xi + H_y(\theta)y + H_\eta(\theta)\eta\end{array}\right.\\
\qquad\qquad (\tau,\xi,y,\eta)\in[0,\overline{\mathrm{T}}]\times\mathbb{R}^{n_\xi+n_y+n_e d}\\
\left\{\begin{array}{rcl}\tau^+ & = & 0\\ \xi^+ & = & \xi\end{array}\right.\\
\qquad\qquad (\tau,\xi,y,\eta,\theta)\in[\underline{\mathrm{T}},\overline{\mathrm{T}}]\times\mathbb{R}^{n_\xi+n_y+(n_e+1)d}
\end{array}
\tag{7.22}
$$

with state $(\tau,\xi)\in\mathbb{R}\times\mathbb{R}^{n_\xi}$, $n_\xi\in\mathbb{N}$, and output

$$
u = D_\xi(\theta)\xi + D_y(\theta)y + D_\eta(\theta)\eta\,. \tag{7.23}
$$

The parameters $\underline{\mathrm{T}},\overline{\mathrm{T}} > 0$ define the jump times, while $\underline{\mathrm{T}}$ must be taken sufficiently large to achieve closed-loop stability (see Proposition 7.2), $\overline{\mathrm{T}}$ is only constrained to be larger or equal to $\underline{\mathrm{T}}$. In this way the update law can be triggered with non-periodic timing strategies in the limits of the stability constraints. The subsystem ξ is instead a continuous-time system that is designed to stabilize the closed-loop system during flows and with $w = 0$. More precisely, (7.22)-(7.23) is designed so that:

P1. H_ξ, H_y, H_η, D_ξ, D_y and D_η are locally Lipschitz functions of θ.

P2. The matrix

$$
F(\theta) = \begin{pmatrix} A + BD_y(\theta)C & BD_\eta(\theta) & BD_\xi(\theta)\\ GC_e & \Phi(\theta) & 0\\ H_y(\theta)C & H_\eta(\theta) & H_\xi(\theta)\end{pmatrix} \tag{7.24}
$$

is Hurwitz for all $\theta\in\mathcal{E}$.

Remark 7.2. We remark that, in view of Lemma 7.2, P1 and P2 can be always achieved. As a matter of fact, θ is available for feedback and the only matrix in the equations of (x,η) that depends on θ is $\Phi(\theta)$, whose dependency is smooth. $\triangle$

7.3.4 The Overall Regulator

The overall regulator, obtained by interconnecting the subsystems (7.5), (7.21), (7.22), (7.23), is thus a hybrid system described by the following equations

$$\begin{aligned}
&\begin{cases}
\dot{\tau} &= 1\\
\dot{\eta} &= \Phi(\theta)\eta + Ge\\
\dot{R} &= 0,\ \dot{v} = 0,\ \dot{\theta} = 0\\
\dot{\xi} &= H_\xi(\theta)\xi + H_y(\theta)y + H_\eta(\theta)\eta
\end{cases}\\
&\qquad (\tau, \eta, R, v, \theta, \xi, y) \in [0, \overline{\mathrm{T}}] \times \mathbb{R}^{n_e d} \times \mathcal{Z} \times \mathbb{R}^{d+n_\xi+n_y}\\
&\begin{cases}
\tau^+ &= 0\\
\eta^+ &= \eta\\
R^+ &= \mu R + \gamma(\eta)\gamma(\eta)^T\\
v^+ &= \mu v + \gamma(\eta)\big((\theta^T \otimes I_{n_e})\eta + e\big)\\
\theta^+ &\in p_{\mathcal{E}}(R^\dagger v)\\
\xi^+ &= \xi
\end{cases}\\
&\qquad (\tau, \eta, R, v, \theta, \xi, y) \in [\underline{\mathrm{T}}, \overline{\mathrm{T}}] \times \mathbb{R}^{n_e d} \times \mathcal{Z} \times \mathbb{R}^{d+n_\xi+n_y}
\end{aligned} \tag{7.25}$$

with input y and with output

$$u = D_\xi(\theta)\xi + D_y(\theta)y + D_\eta(\theta)\eta\,. \tag{7.26}$$

7.4 Asymptotic Properties of the Regulator

We let for convenience $\mathbf{w} := (w, \tau)$, $\chi := (x, \eta, \xi)$ and

$$\begin{aligned}
s_{\mathbf{w}}(\mathbf{w}) &:= \mathrm{col}(Sw, 1)\\
g_{\mathbf{w}}(\mathbf{w}) &:= \mathrm{col}(w, 0)\\
G_z(z, \chi) &:= \{\mu R + \gamma(\eta)\gamma(\eta)^T\} \times \{\mu v + \gamma(\eta)((\theta^T \otimes I_{n_e})\eta + e)\}\\
E(\theta) &:= \mathrm{col}(P + BD_y(\theta)Q, GQ_e, H_y(\theta)Q)\\
L(\theta) &:= \mathrm{col}(M + BD_y(\theta)R, GN_e, H_y(\theta)N).
\end{aligned}$$

Then, the closed loop system given by (7.1), (7.5), (7.21), (7.22), (7.23) reads as follows

$$\left\{\begin{array}{rcl}\dot{\mathbf{w}} & = & s_{\mathbf{w}}(\mathbf{w})\\ \dot{\chi} & = & F(\theta)\chi + E(\theta)w + L(\theta)\nu\\ \dot{z} & = & 0\\ \dot{\theta} & = & 0\end{array}\right. \qquad \left\{\begin{array}{rcl}\mathbf{w}^{+} & = & g_{\mathbf{w}}(\mathbf{w})\\ \chi^{+} & = & \chi\\ z^{+} & = & G_z(z,\chi)\\ \theta^{+} & \in & p_{\mathcal{E}}(R^{\dagger}v)\end{array}\right. \tag{7.27}$$

with flow and jump sets given by $\mathcal{C} := W \times [0, \overline{\mathrm{T}}] \times \mathbb{R}^{n_\chi} \times \mathcal{Z} \times \mathbb{R}^d \times \mathcal{N}$ and $\mathcal{D} := W \times [\underline{\mathrm{T}}, \overline{\mathrm{T}}] \times \mathbb{R}^{n_\chi} \times \mathcal{Z} \times \mathbb{R}^d \times \mathcal{N}$, being $n_\chi := n_x + n_e d + n_\xi$ and $\mathcal{N} \subset \mathbb{R}^{n_\nu}$ an arbitrarily large compact set. In the definition of $\mathcal{C}$ and $\mathcal{D}$ we restricted the flow and jump sets of (w, ν) to the compact set $W \times \mathcal{N}$. In this way we consider only solutions for which $w(t) \in W$ and $\nu(t) \in \mathcal{N}$. Since W is assumed to be forward invariant for the exosystem (7.2), we maintain completeness of the solutions for all inputs ν satisfying $\nu(t,j) \in \mathcal{N}$.

As long as $n_w \leq d$, the Cayley-Hamilton theorem guarantees the existence of $\omega \in \mathbb{R}^d$ such that the exosystem's state matrix S satisfies

$$S^d - \omega_d S^{d-1} - \cdots - \omega_2 S - \omega_1 I = 0\,. \tag{7.28}$$

As mentioned in the previous sections, if the internal model unit (7.5) is implemented with $\theta = \omega$, for any ω for which (7.28) holds, then asymptotic regulation is achieved. As we constrained θ to range in $\mathcal{E}$, we will eventually rely on the following assumption:

Assumption 7.4. *There exists $\omega \in \mathcal{E}$ such that* (7.28) *holds.*

The following Proposition, which is the main result of the chapter, states the main asymptotic properties of the proposed regulator.

Proposition 7.2. *Suppose that Assumptions 7.1 and 7.3 hold and let* (7.22) *be chosen such that P1 and P2 hold. Then there exists $\underline{\mathrm{T}}_1^\star$, such that if $\underline{\mathrm{T}} \geq \underline{\mathrm{T}}_1^\star$, all the solutions of* (7.27) *are bounded. If in addition Assumption 7.4 holds, for any $\epsilon > 0$ there exist $\underline{\mathrm{T}}_2^\star \geq \underline{\mathrm{T}}_1^\star$ and $\overline{\nu}, c \geq 0$ such that, if $\underline{\mathrm{T}} \geq \underline{\mathrm{T}}_2^\star$, for each complete solution pair to* (7.27) *for which η is (J,ϵ)-PE, for some $J \in \mathbb{N}$, and $|\nu|_\infty \leq \bar{\nu}$ the following holds*

$$\limsup |e| \leq c \limsup |\nu|\,. \tag{7.29}$$

With reference to the proof of Proposition 7.2 (that is below this paragraph), we note that boundedness of the trajectories is obtained if $\underline{\mathrm{T}}$ is larger than a quantity that only depends on the closed-loop system's data and that can be fixed after the regulator is designed. The bound (7.29) is instead more complex. It is in fact a property guaranteed just along the trajectories for which η is (J, ϵ)-PE, for some $(J, \epsilon) \in \mathbb{N} \times \mathbb{R}_+^*$, and only if T is larger and $\limsup |\nu|$ is smaller than constants that, in general, depend on ϵ. Therefore the bound (7.29) is local in $\nu(t)$, with the same constants, though, that work for any trajectory for which η is (J', ϵ')-PE with $J' \in \mathbb{N}$ and $\epsilon' \geq \epsilon$. We also remark that there is no uniformity in the convergence (7.29), as the convergence rate strongly depends on the particular J for which the (J, ϵ)-PE condition holds. This, however, matches with the intuition that the correct adaptation can take place only after the input signal to the identifier becomes sufficiently informative. Therefore, uniformity in the choice of $\underline{\mathrm{T}}_2^\star$ and in the convergence (7.29) is possible only inside the class of solutions to the closed-loop system that are (J, ϵ)-PE with the same J and ϵ. Finally we note that if $\nu = 0$, i.e. if no unmodeled disturbances are present, then (7.29) implies asymptotic regulation, i.e. $e(t, j) \to 0$.

Proof of Proposition 7.2. The existence of $\underline{\mathrm{T}}_1^\star$ such that for $\underline{\mathrm{T}} \geq \underline{\mathrm{T}}_1^\star$ the maximal trajectories of (7.27) are complete and bounded follows from standard "slow-switching" arguments (see for instance (Hespanha and Morse, 1999)) once noted that $F(\theta)$ and $E(\theta)$ are bounded uniformly in θ and boundedness of η implies those of z. In proving the second claim, we articulate the discussion in the following 4 points.

1) Quasi steady state of the stabilized cascade χ

We prove now that during the flow intervals the system χ evolves towards a "quasi" steady state determined by w and parametrized by θ. As $F(\theta)$ is Hurwitz for each $\theta \in \mathcal{E}$, there exist Lipschitz maps $P : \mathcal{E} \to \mathbb{R}^{n_\chi \times n_\chi}$ and $\Pi : \mathcal{E} \to \mathbb{R}^{n_\chi \times n_w}$, with $P(\cdot)$ having symmetric and positive definite values, that are point-wise solutions to

$$
\begin{aligned}
&F(\theta)^T P(\theta) + P(\theta) F(\theta) = -I_{n_\chi} \\
&\Pi(\theta) S = F(\theta) \Pi(\theta) + E(\theta) \,.
\end{aligned}
$$

Define the function

$$V(w,\chi,z,\theta) := \big(\chi - \Pi(\theta)w\big)^T P(\theta)\big(\chi - \Pi(\theta)w\big)\,.$$

Then, by letting $\underline{\sigma} := \min\{\lambda \in \mathbb{R} \,:\, \lambda \in \sigma(P(\theta)),\, \theta \in \mathcal{E}\}$ and $\bar{\sigma} := \max\{\lambda \in \mathbb{R} \,:\, \lambda \in \sigma(P(\theta)),\, \theta \in \mathcal{E}\}$, simple computations show that V fulfills $\underline{\sigma}|\chi - \Pi(\theta)w|^2 \le V(w,\chi,z) \le \bar{\sigma}|\chi - \Pi(\theta)w|^2$ in the whole state space and

$$L_{\mathcal{F}}V(w,\chi,z,\theta) \le -\lambda V(w,\chi,z,\theta) + r_0|\nu|^2 \tag{7.30}$$

for all (w,χ,z,θ) such that $(\mathbf{w},\chi,z,\theta) \in \mathcal{C}$, with $\mathcal{F} := (Sw, F(\theta)\chi + E(\theta)w + L(\theta)\nu, 0)$ and with $\lambda > 0$. On the other hand, for all $(\mathbf{w},\chi,z,\theta) \in \mathcal{D}$ and with $(w^+,\chi^+,z^+,\theta^+) = (w,\chi,G_z(z,\chi),(\mu R + \gamma(\eta)\gamma(\eta)^T)^\dagger(\mu v + \gamma(\eta)(\theta^T \otimes I)\eta + e))$ we obtain

$$\begin{aligned} V(w^+,\chi^+,z^+,\theta^+) &\le \bar{\sigma}|\chi - \Pi(\theta^+)|^2 \\ &\le \bar{\sigma}\big(|\chi - \Pi(\theta)w|^2 + |\Pi(\theta)w - \Pi(\theta^+)w|^2\big) \\ &\le r_1 V(w,\chi,z,\theta) + r_2|\theta - \theta^+|^2 \end{aligned} \tag{7.31}$$

with $r_1 = \bar{\sigma}/\underline{\sigma}$ and with $r_2 > 0$ properly chosen by using the fact that $w \in W$, $\theta^+ \in \mathcal{E}$ and $\Pi(\cdot)$ is Lipschitz on $\mathcal{E}$.

2) Properties of the identifier

We show now that if the flow is long enough the distance of χ to its quasi steady state at jump times, where adaptation takes place, is sufficiently small to conclude that if η is persistently exciting, then the identification problem associated to the steady-state signals has an unique optimum. Let $\Pi_x(\theta) \in \mathbb{R}^{n_x \times n_w}$, $\Pi_\eta(\theta) \in \mathbb{R}^{n_\eta \times n_w}$ and $\Pi_\xi(\theta) \in \mathbb{R}^{n_\xi \times n_w}$ be such that $\Pi(\theta) = \mathrm{col}(\Pi_x(\theta), \Pi_\eta(\theta), \Pi_\xi(\theta))$ and, for $i = 1,\ldots,d$, let $\Pi_{\eta_i}(\theta)$ be such that $\Pi_\eta(\theta) = \mathrm{col}(\Pi_{\eta_1}(\theta),\ldots,\Pi_{\eta_d}(\theta))$. As a consequence of the structure of the matrix $\Phi(\theta)$, we have

$$\begin{cases} \Pi_{\eta_i}(\theta)S = \Pi_{\eta_{i+1}}(\theta), \qquad i = 1,\ldots,d-1 \\ \Pi_{\eta_d}(\theta)S = (\theta^T \otimes I_{n_e})\Pi_\eta(\theta) + \Pi_e(\theta) \end{cases} \tag{7.32}$$

with $\Pi_e(\theta) := C_e\Pi_x(\theta) + Q_e$. Equation (7.32) also yields

$$\eta = \Pi_\eta(\theta)w + \delta_\alpha$$
$$(\theta^T \otimes I_{n_e})\eta + e = \Pi_{\eta_d}(\theta)Sw + \delta_\beta$$

where $\delta_\alpha := \eta - \Pi_\eta(\theta)w$ and $\delta_\beta := (\theta^T \otimes I_{n_e})(\eta - \Pi_\eta(\theta)w) + (e - \Pi_e(\theta)w)$ that satisfy

$$|(\delta_\alpha, \delta_\beta)|^2 \leq r_3 V(w, \chi, z, \theta) + r_4|\nu|^2, \tag{7.33}$$

with $r_3 := (1 + |C_e|^2 + |\mathcal{E}|^{2n_e})/\underline{\sigma}$ and $r_4 := |N_e|^2$. The identifier subsystem can be thus seen as a system with input $(\alpha, \beta) = (\alpha^\star + \delta_\alpha, \beta^\star + \delta_\beta)$, where $(\alpha^\star, \beta^\star) = (\Pi_\eta(\theta)w, \Pi_{\eta_d}(\theta)Sw)$ and $(\delta_\alpha, \delta_\beta)$ defined as before. This yields two consequences:

1) Single-valued Optimum: As $\nu(t,j) \in \mathcal{N}$ for each solution pair to (7.27), and λ and r_1 in (7.30)-(7.31) are constants, we can assume without loss of generality that $\underline{\mathrm{T}}_1^\star$ is chosen large enough so that (7.30)-(7.31) can be turned, by using standard average dwell-time conditions (Cai et al., 2008; Liberzon et al., 2014), to a ISS-Lyapunov function (Cai and Teel, 2013). This in turn implies the existence of a $\Delta_0 > 0$, depending on $\mathcal{E}$ and $\mathcal{N}$, such that, for each solution pair $((\mathbf{w}, \chi, z, \theta), \nu)$ to (7.27) there exists $\bar{s}_1 > 0$ such that $|\chi(t,j)| \leq \Delta_0$ for all $(t,j) \in \mathrm{dom}(\mathbf{w}, \chi, z, \theta)|_{\geq \bar{s}_1}$. Thus, in particular there exists $\Delta_1 > 0$ such that $|(\alpha(t,j), \beta(t,j))| = |(\eta(t,j), (\theta(t,j)^T \otimes I_{n_e})\eta(t,j) + e(t,j))| \leq \Delta_1$ for all $(t,j) \in \mathrm{dom}(\mathbf{w}, \chi, z, \theta)|_{\geq \bar{s}_1}$. Also, as W and $\mathcal{E}$ are compact, there exists $\Delta_2 > 0$ such that $|(\alpha^\star(t,j), \beta^\star(t,j))| = |(\Pi_\eta(\theta(t,j))w(t,j), \Pi_{\eta_d}(\theta(t,j))Sw(t,j))| \leq \Delta_2$ for all $(t,j) \in \mathrm{dom}(\mathbf{w}, \chi, z, \theta)$. Suppose that $\alpha = \eta$ is (J, ϵ)-PE, for some $(J, \epsilon) \in \mathbb{N} \times \mathbb{R}_+$. Then, in view of Lemma 7.3, there exist $(J', \epsilon') \in \mathbb{N} \times \mathbb{R}_{>0}$ and $\bar{\delta} > 0$, depending on $\mathcal{E}, \mathcal{N}$ and ϵ, such that $|(\delta_\alpha(t_j, j), \delta_\beta(t_j, j))| \leq \bar{\delta}$ for all $j \geq J$ implies that $(\alpha^\star, \beta^\star)$ is (J', ϵ')-PE. Equation (7.30) gives

$$V(w(t^j, j), \chi(t^j, j), z(t^j, j), \theta(t^j, j)) \leq \bar{V} e^{-\lambda \underline{\mathrm{T}}} + r_0 |\nu|_\infty^2 / \lambda$$

for all $j \in \mathbb{N}$ such that $t_j + j \geq \bar{s}_1$ and with $\bar{V} := \bar{\sigma}(\Delta_0^2 + |\Pi(\mathcal{E})|^2 |W|^2)$. Therefore, as long as

$$|\nu|_\infty \leq \bar{\nu} := \bar{\delta} \sqrt{\frac{1}{3} \max\left\{\frac{\lambda}{r_0 r_3}, \frac{1}{r_4}\right\}}$$

$$\underline{\mathrm{T}} \geq \underline{\mathrm{T}}_{\epsilon}^{\star} := \max\left\{\underline{\mathrm{T}}_1^{\star}, (1/\lambda)\log\left(3\bar{V}r_3/\bar{\delta}^2\right)\right\}$$

then (7.33) implies that every solution pair $((\mathbf{w}, \chi, z, \theta), \nu)$ to (7.27) with $\underline{\mathrm{T}} \geq \underline{\mathrm{T}}_{\epsilon}^{\star}$ for which η is (J, ϵ)-PE fulfils $|(\delta_{\alpha}(t_j, j), \delta_{\beta}(t_j, j))| \leq \bar{\delta}$ for all $j \geq J$ such that $t_j + j \geq \bar{s}_1$. Noting that (J, ϵ)-PE implies (j, ϵ)-PE for all $j \geq J$, we thus conclude that there exists a $J_2 \geq \max\{J, J', \inf_{j\in\mathbb{N}} t_j + j \geq \bar{s}_1\}$ such that $\alpha^{\star} = \Pi_{\eta}(\theta)w$ is (J_2, ϵ')-PE, and the map $\theta^{\circ}_{\alpha^{\star},\beta^{\star}}(t, j)$ is a singleton for all $(t, j) \in \mathrm{dom}(\mathbf{w}, \chi, z, \theta)|_{\geq \bar{s}_2}$, having denoted $\bar{s}_2 := t_{J_2} + J_2$.

2) Stability: In view of the aforementioned bounds Δ_1 and Δ_2 on (α, β) and $(\alpha^{\star}, \beta^{\star})$, Proposition 7.1 implies that, with the same $\rho > 0$, for each solution pair $((\mathbf{w}, \chi, z, \theta), \nu)$ to (7.27) there exist $z^{\star} : \mathrm{dom}(\mathbf{w}, \chi, z, \theta) \to \mathcal{Z}$ such that, for all $(t, j) \in \Gamma(\mathrm{dom}(\mathbf{w}, \chi, z, \theta)|_{\geq \bar{s}_2})$ and with $\tilde{z} := z - z^{\star}$, we have

$$|\tilde{z}^{+}|^2 \leq \mu^2|\tilde{z}|^2 + \rho r_3 V(w, \chi, z, \theta) + \rho r_4|\nu|^2, \tag{7.34}$$

where we omitted the argument (t, j) and we let $\tilde{z}^{+} := \tilde{z}(t, j{+}1)$. In the following, for an arbitrary $\epsilon > 0$, we let $\mathcal{S}_{\epsilon}$ be the class of the solution pairs $((\mathbf{w}, \chi, z, \theta), \nu)$ to the closed-loop system with $\underline{\mathrm{T}} \geq \underline{\mathrm{T}}_{\epsilon}^{\star}$ and such that α is (J, ϵ)-PE for some $J \in \mathbb{N}$. We stress that the above discussion, and thus in particular, that $\alpha^{\star} = \Pi_{\eta}(\theta)w$ is (J_2, ϵ')-PE and $\theta^{\circ}_{\alpha^{\star},\beta^{\star}}(t, j)$ is a singleton for $t + j \geq \bar{s}_2$, holds for all such solutions, with only J_2 and $\bar{s}_2$ that possibly depend on the particular solution. In the following we will make reference to the solution-dependent quantities introduced above (such as J_2 and $\bar{s}_2$) with the remark that they are meant to be defined in the same way as before and they refer to the particular solution considered.

3) (J, ϵ)-PE and Assumption 7.4 yield the internal model property

We show now that the single-valued solution to the identification problem associated to the quasi steady-state inputs is independent on θ and coincides with ω of (7.28). In view of Assumption 7.4, there exists $\omega \in \mathcal{E}$ such that (7.28) holds. As a consequence, (7.32) yields

$$\Pi_{\eta_d}(\theta)S = \Pi_{\eta_1}(\theta)S^d = \Pi_{\eta_1}(\theta)\Big(\omega_d S^{d-1} + \cdots + \omega_2 S + \omega_1 I\Big) \tag{7.35}$$

Then, for any solution pair $((\mathbf{w}, \chi, z, \theta), \nu)$ to the closed-loop system, (7.32) and

(7.35) yield

$$
\begin{aligned}
&|\beta^\star(t^j,j) - (\omega^T \otimes I_{n_e})\alpha^\star(t^j,j)| \\
&= \left|\Big(\Pi_{\eta_d}(\theta(t^j,j))S - \Pi_{\eta_1}(\theta(t^j,j))(\omega_d S^{d-1} + \cdots + \omega_1 I)\Big) w(t^j,j)\right| \\
&= 0
\end{aligned}
$$

for all $(t^j,j) \in \operatorname{dom}(\mathbf{w},\chi,z,\theta)$. By definition of $\mathcal{J}_{\alpha^\star,\beta^\star}$ in (7.10), this also implies that $\mathcal{J}_{\alpha^\star,\beta^\star}(\omega)(t,j) = 0$, and thus that $\omega \in \theta^\circ_{\alpha^\star,\beta^\star}(t,j)$, for all $(t,j) \in \operatorname{dom}(\mathbf{w},\chi,z,\theta)$. Pick $((\mathbf{w},\chi,z,\theta),\nu) \in \mathcal{S}_\epsilon$, as $\alpha^\star$ is (J_2,ϵ')-PE and $\theta^\circ_{\alpha^\star,\beta^\star}(t,j)$ is single valued for all $(t,j) \in \operatorname{dom}(\mathbf{w},\chi,z,\theta)|_{\geq \bar{s}_2}$, then necessarily

$$
\theta^\circ_{\alpha^\star,\beta^\star}(t,j) = \{\omega\} \subset \mathcal{E}, \quad \forall (t,j) \in \operatorname{dom}(\mathbf{w},\chi,z,\theta)|_{\geq \bar{s}_2}. \tag{7.36}
$$

A further consequence of (7.32) is that

$$
\begin{aligned}
\Pi_e(\omega) &= \Pi_{\eta_d}(\omega)S - (\omega^T \otimes I_{n_e})\Pi_\eta(\omega) \\
&= \Pi_{\eta_1}(\omega)\Big(S^d - \omega_d S^{d-1} - \cdots - \omega_1 I\Big) = 0.
\end{aligned} \tag{7.37}
$$

so as if $\limsup |\chi - \Pi(\omega)w|$ is proportional to $\limsup |\nu|$, then (7.29) is proved. In the next paragraph we show that this is the case whenever T is sufficiently large.

4) Large T yields small gain

Finally, we show here that if the jump times are distant enough, a small-gain like condition holds, and "modulo ν", θ tends to the optimum ω and χ to the error-zeroing steady state $\Pi(\omega)w$. Thus the claim of the proposition follows. Pick $((\mathbf{w},\chi,z,\theta),\nu) \in \mathcal{S}_\epsilon$. As $\omega \in \mathcal{E}$, in view of (7.36) and of Proposition 7.1, for all $(t,j) \in \operatorname{dom}(\mathbf{w},\chi,z,\theta)|_{\geq \bar{s}_2}$ we have $\omega = \theta^\star(t,j)$, with $\theta^\star(t,j)$ the unique element of $p_{\mathcal{E}}(R^\star(t,j)^\dagger v^\star(t,j))$. Then Lemma 7.4 can be invoked to claim the existence of $J_3 \geq J_2$ and $r_5 \geq 0$, depending on ϵ, such that, with $\bar{s}_3 := t_{J_3} + J_3$, the following holds

$$
|\theta(t,j) - \omega|^2 \leq r_5 |\tilde{z}(t,j)|^2 \tag{7.38}
$$

for all $(t,j) \in \operatorname{dom}(\mathbf{w},\chi,z,\theta)|_{\geq \bar{s}_3}$. As a consequence of (7.38) and (7.34), for all $(t,j) \in \Gamma(\operatorname{dom}(\mathbf{w},\chi,z,\theta)|_{\geq \bar{s}_3})$, we obtain

$$
|\theta - \theta^+|^2 = |\theta - \omega + \omega - \theta^+|^2 \leq |\theta - \omega|^2 + |\theta^+ - \omega|^2 \leq r_5\big(|\tilde{z}|^2 + |\tilde{z}^+|^2\big)
$$

$$\leq r_6|\tilde{z}|^2 + r_7 V(w,\chi,z,\theta) + r_8|\nu|^2,$$

with $r_6 := r_5(1+\mu^2)$, $r_7 := r_5\rho r_3$, $r_8 := r_5\rho r_4$ and where again we omitted the argument (t,j) and we let $\theta^+ := \theta(t,j+1)$ and $\tilde{z}^+ := \tilde{z}(t,j+1)$. We further develop (7.31) to obtain, for all $(t,j) \in \Gamma(\mathrm{dom}(\mathbf{w},\chi,z,\theta)|_{\geq \bar{s}_3})$

$$V(w^+,\chi^+,z^+,\theta^+) \leq r_9 V(w,\chi,z,\theta) + r_{10}|\tilde{z}|^2 + r_{11}|\nu|^2,$$

being $r_9 := r_1 + r_2 r_7$, $r_{10} := r_2 r_6$ and $r_{11} := r_2 r_8$. In summary, for all $(t,j) \in \mathrm{dom}(\mathbf{w},\chi,z,\theta)|_{\geq \bar{s}_3}$ such that $t \in (t_j, t_{j+1})$ we have:

$$\dot{V}(w,\chi,z,\theta) \leq -aV(w,\chi,z,\theta) + r_0|\nu|^2, \quad D^+|\tilde{z}|^2 = 0 \tag{7.39}$$

and for all $(t,j) \in \Gamma(\mathrm{dom}(\mathbf{w},\chi,z,\theta)|_{\geq \bar{s}_3})$

$$\begin{aligned} V(w^+,\chi^+,z^+,\theta^+) &\leq r_9 V(w,\chi,z,\theta) + r_{10}|\tilde{z}|^2 + r_{11}|\nu|^2 \\ |\tilde{z}^+|^2 &\leq \mu^2|\tilde{z}|^2 + \rho r_3 V(w,\chi,z,\theta) + \rho r_4|\nu|^2, \end{aligned} \tag{7.40}$$

where we omitted the argument (t,j) and we let $(w^+,\chi^+,z^+,\theta^+,\tilde{z}^+) := (w(t,j+1),\chi(t,j+1),z(t,j+1),\theta(t,j+1),\tilde{z}(t,j+1))$. We then have the following:

Lemma 7.5. *There exist $c_0 \geq 0$ and $\underline{\mathrm{T}}_2^\star \geq \underline{\mathrm{T}}_\epsilon^\star$, independent on J_3, such that for any solution pair in $\mathcal{S}_\epsilon$ with $\underline{\mathrm{T}} \geq \underline{\mathrm{T}}_2^\star$*

$$\limsup(V(w,\chi,z,\theta) + |\tilde{z}|^2) \leq c_0 \limsup|\nu|^2.$$

Lemma 7.5 (proved at the end of this proof) implies in particular

$$\begin{aligned} &\limsup|\chi - \Pi(\theta)w| \leq c_1 \limsup|\nu| \\ &\limsup|\theta - \omega| \leq c_2 \limsup|\nu| \end{aligned}$$

for $c_1 := \sqrt{c_0/\underline{\sigma}}$ and $c_2 := \sqrt{r_5 c_0}$. In view of (7.37), $\Pi_e(\omega) = 0$, so that

$$\begin{aligned} |e| &= |e - \Pi_e(\omega)w| \\ &\leq |C_e||\chi - \Pi(\omega)| + |C_e||w||\Pi(\theta) - \Pi(\omega)| + |N_e||\nu|. \end{aligned}$$

As $\Pi(\cdot)$ is Lipschitz on $\mathcal{E}$, this suffices to conclude (7.29), with $c := |C_e|c_1 +$

$|C_e||W|L_\Pi c_2 + |N_e|$, where L_Π is the Lipschitz constant of Π on $\mathcal{E}$. The result of the Proposition follows then from the arbitrariness of ϵ. ■

Proof of Lemma 7.5. Pick any $k \in (0, a)$ and $q \in (\mu^2, 1)$, and let

$$\ell_1 \in (0, q - \mu^2)\,, \qquad\qquad \ell_2 \in (1, q/(\mu^2 + \ell_1))\,.$$

then $0 < \ell_1 < 1$, $\ell_2 > 1$ and $(\ell_1 + \mu^2)\ell_2 < q < 1$. Let

$$\psi \geq r_{10}/\ell_1, \quad \underline{\mathrm{T}}_2^\star := \max\left\{\underline{\mathrm{T}}_\epsilon^\star, \frac{1}{k}\log\left(\frac{r_9 + \rho r_3 \psi}{q}\right)\right\} \tag{7.41}$$

pick in the jump set of (7.27) $\overline{\mathrm{T}} \geq \underline{\mathrm{T}} \geq \underline{\mathrm{T}}_2^\star$ and let

$$h \in (0, \log(\ell_2)/\overline{\mathrm{T}}). \tag{7.42}$$

Define the function

$$W(\mathbf{w}, \chi, z, \theta, \tilde{z}) := e^{k\tau} V(w, \chi, z, \theta) + \psi e^{-h\tau} |\tilde{z}|^2.$$

Then, clearly,

$$V(w, \chi, z, \theta) + |\tilde{z}|^2 \leq \max\{1, e^{h\overline{\mathrm{T}}}/\psi\} W(\mathbf{w}, \chi, z, \theta, \tilde{z}). \tag{7.43}$$

Pick a solution $(\mathbf{w}, \chi, z, \theta) \in \mathcal{S}_\epsilon$ with $\overline{\mathrm{T}} \geq \underline{\mathrm{T}} \geq \underline{\mathrm{T}}_2^\star$. Then, using (7.22), for all $(t, j) \in \mathrm{dom}(\mathbf{w}, \chi, z, \theta)|_{\geq \bar{s}_3}$ such that $t \in (t_j, t_{j+1})$, (7.39) yields (we omit the time dependency)

$$\dot{W}(\mathbf{w}, \chi, z, \theta, \tilde{z}) \leq -a_W W(\mathbf{w}, \chi, z, \theta, \tilde{z}) + r_{12}|\nu|^2, \tag{7.44}$$

with $a_W := \min\{a - k, h\}$ and $r_{12} := e^{k\overline{\mathrm{T}}} r_0$. As $\tau^+ = 0$, for all $(t, j) \in \Gamma(\mathrm{dom}(\mathbf{w}, \chi, z, \theta)|_{\geq \bar{s}_3})$, instead, (7.40) yields

$$W(\mathbf{w}^+, \chi^+, z^+, \theta^+, \tilde{z}^+) = (r_9 + \rho r_3) V(w, \chi, z, \theta) + (r_{10} + \psi\mu^2)|\tilde{z}|^2 + r_{13}|\nu|^2,$$

with $r_{13} := r_{11} + \rho r_4 \psi$. As for each $(t, j) \in \Gamma(\mathrm{dom}(\mathbf{w}, \chi, z, \theta))$, necessarily, $\underline{\mathrm{T}} \leq$

$\tau(t, j) \le \overline{\mathrm{T}}$, we get

$$W(\mathbf{w}^+, \chi^+, z^+, \theta^+, \tilde{z}^+) \le (r_9 + \rho r_3 \psi) e^{-k\underline{\mathrm{T}}} e^{k\tau} V(w, \chi, z, \theta) + (r_{10} + \psi\mu^2) e^{h\overline{\mathrm{T}}} e^{-h\tau} |\tilde{z}|^2 + r_{13}|\nu|^2.$$

(7.41)-(7.42) and $\underline{\mathrm{T}} \ge \underline{\mathrm{T}}_2^\star$ gives:

$$\begin{aligned}(r_9 + \rho r_3 \psi) e^{-k\underline{\mathrm{T}}} &\le (r_9 + \rho r_3 \psi) e^{-k\underline{\mathrm{T}}_2^\star} \le q \\ (r_{10} + \psi\mu^2) e^{h\overline{\mathrm{T}}} &\le \psi(r_{10}/\psi + \mu^2)\ell_2 c \le \psi q,\end{aligned}$$

so that we obtain

$$W(\mathbf{w}^+, \chi^+, z^+, \theta^+, \tilde{z}^+) \le q W(\mathbf{w}, \chi, z, \theta, \tilde{z}) + r_{13}|\nu|^2, \tag{7.45}$$

and the claim follows from (7.43), (7.44) and (7.45). ■

7.5 An Example

We consider here a plant of the form (7.1) with

$$A = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 2 \end{pmatrix}.$$

The control goal is to drive $y_1 := x_2$ to a desired set point $y_1^\star$ chosen by the user and to make $y_2 := x_3$ follow a sinusoid $y_2^\star(t)$ at any desired frequency, despite the disturbances $Pw(t)$ acting on the system. We suppose that $w(t)$ is a combination of a constant term, a harmonic at the same frequency of $y_2^\star(t)$ and a third unknown harmonic. The disturbance $Pw(t)$ and the references $y_1^\star$ and $y_2^\star(t)$ can be thus modeled as outputs of an exosystem of the form (7.2) with $S := \mathrm{blkdiag}(S_1, S_2, S_3)$, where

$$S_1 = \gamma_1 \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad S_2 = \gamma_2 \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad S_3 = 0$$

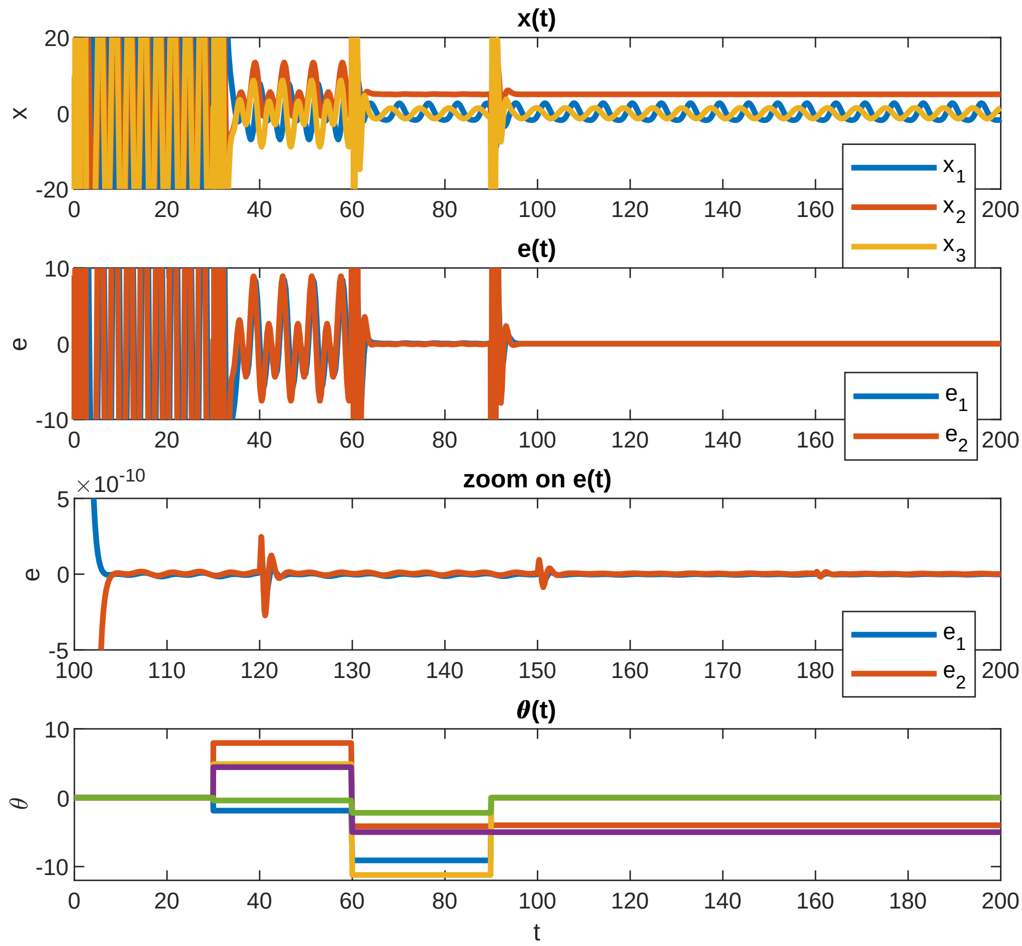

Figure 7.2: Simulation results: The first plot shows the time evolution of the state x, the second plot the regulation error $e(t)$, the third plot is a zoom on the regulation error at a scale of the order of 10^{-10} and the last plot shows the evolution of the parameters θ.

and with $\gamma_1, \gamma_2 \in \mathbb{R}_+^*$ and $e_2 := y_2 - y_2^\star$, that are obtained as in (7.3) with the choice

$$C_e = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad\qquad Q_e := \begin{pmatrix} 0 & 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

and noting that different set-points $y_1^\star$, as well as different amplitudes and phases for $y_2^\star(t)$ can be obtained by changing the initial conditions of w and need not to be known at the design stage. The amplitude and the phase of the disturbance $Pw(t)$ depends, other on $w(0)$, on the matrix P which is not known by the designer. It is worth noting that if we let $x^* := \Pi w$ and $u^* := \Gamma w$ be the corresponding steady state functions such that $e^* := C_e x^* + Q_e w = 0$, then, by letting $\tilde{x} := x - x^*$, we obtain that $e = 0$ and $u = u^\star$ imply[1]

$$\dot{\tilde{x}}_1 = \tilde{x}_1 \,,$$

so that the plant considered is not minimum phase relatively to the graph of Π and, as a consequence, this example does not fit in the frameworks addressed in the existing literature. In the simulation, for simplicity, we used a state-feedback stabilizer, i.e. we assumed $y = \mathrm{col}(e, x_1)$ and $u = K(\theta)y$ with $K(\theta)$ properly designed. Figure 7.2 shows the result of a simulation of the proposed control system implemented with $d = 5$, $\mu = 0.9$ and with $T = 30$s. In the simulation we let $y_2^* = 5$, $\gamma_1 = 1$, $\gamma_2 = 2$, $w(0) = \mathrm{col}(1, -1, 0, 1, y_2^*)$, $x(0) = \mathrm{col}(5, -10, 10)$ and

$$P = \begin{pmatrix} 1 & 1 & 0 & 1 & -1 \\ -1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix} .$$

[1] This can be verified by noting that Π and Γ fulfill the *regulator equations* $\Pi S = A\Pi + B\Gamma + P$ and $0 = C_e\Pi + Q_e$, and that $e = 0$ implies $\tilde{x}_2 = \tilde{x}_3 = 0$.

Conclusion

This last part of the thesis was dedicated to the design of adaptive regulators for multivariable systems, by leveraging on the post-processing regulator of Chapter 2 and on the identification framework of Chapter 4. In particular, Chapter 6 presented a general post-processing design procedure hinging on a "non-equilibrium" framework (Byrnes and Isidori, 2003), in which the regulator equations (6.4) are allowed to admit solutions $(x^{\star}, u^{\star})$ that are not necessarily dependent only on the exosystem variables. The uncertainties typically characterizing $(x^{\star}, u^{\star})$ and the need to face the chicken-egg dilemma motivated the adoption of an adaptive internal model, in which adaptation is cast as an identification problem. The chicken-egg dilemma is taken on by moving the overall uncertainty on $(x^{\star}, u^{\star})$ (coming from the uncertainties in the plant and exosystem and form the fact that the stabilizer is still floating when the structure of the internal model is fixed) to the identification level, where algorithms can be developed to deal with it. In line with the identification viewpoint, we considered a more suitable approximate, rather than asymptotic, regulation objective, and the prediction error ε of the identified model was shown in Theorem 6.1 to be directly related to the bound on the asymptotic regulation error, with asymptotic regulation that is obtained only in the idealistic case in which a "true model" exists in the model set. General requirements are introduced to guide the design of the identifier (Requirements 6.1 and 6.3) and the other degrees of freedom related to stabilization (Requirement 6.2), with the performances of the final regulator that result to be dependent on the "quality" of the internal model, influencing $\varepsilon^\star$, and that of the stabilizer, influencing the asymptotic distance to the ideal steady state (Theorem 6.1).

We presented some representative design examples to illustrate how the proposed framework can embrace different regulation problems. A high-gain strategy was proposed to systematically deal with the class of systems possessing a (partial) normal form, with the dimension of the input not necessarily equal to those of the regulation errors. Two examples have been given to show how additional measured outputs can be easily included in the stabilization loop, thus showing how in this post-processing approach we can solve problems that do not fit in the previous pre-processing framework (even in the non-adaptive case). Finally, we presented a possible (continuous-time) least-squares approach to the problem of adaptive regulation for general multivariable linear systems, based on the separation of the time-scales of the learning dynamics and the system evolution.

Chapter 7 proposed a different adaptive solution for linear systems, by using discrete-time identifiers and a regulator that is also identifier-dependent. Interestingly enough, even if in different terms, the main requirement on the identifier was to be *slow enough*, as it happened for the approach of Section 6.4. As a matter of fact, we have shown that if the identifier is slow enough the regulator ensures boundedness of the closed-loop trajectories and, if an upper bound on the order of the exosystem is known, then along the solutions that satisfy a persistence of excitation condition asymptotic regulation is achieved. We also showed robustness of the proposed scheme to sufficiently small unmodeled disturbances.

The material presented is far from being a complete answer to the problem of multivariable nonlinear output regulation, which is definitely an open and challenging research field. The strength of the framework of Section 6.1 is that the solution to the regulator equations are just used in a "qualitative" way in order to select the most appropriate internal model and identifier, this being in sharp contrast with existing design principles that have definitely a "friend-centric" nature. Moreover, many research directions are open by the proposed vision. Large emphasis has to be put in better supporting the identifier and stability requirements so as to enlarge the class of systems that can be dealt with by using the general approach of Section 6.1. For what concerns the identification problem, a road that is definitely worth to investigate is the adoption of *universal approximators* (such as Wavelets and Neural Networks), which permit to further weaken the chicken-egg dilemma, dealing to practical regulation without virtually any a-priori knowledge on the system.

Conclusions and Future Directions

In this thesis we dealt with the problem of output regulation for nonlinear systems under different points of view, and the presented results touched various fundamental pillars of the theory. The first part was dedicated to the contributions concerning the robustness issue and the structural properties of the nonlinear regulators. We presented a broad introduction to output regulation for linear and nonlinear systems (Chapter 1), highlighting the main limits and drawbacks of the state-of-art solutions and proposing a new high-level vision on the problem (chapters 2 and 3). The re-formalization of the concept of robustness given in Section 3.4 provides a formal playground in which the robustness issues relative to arbitrary asymptotic properties can be analyzed in a generalized framework. This opens new questions about robustness, a topic that constituted the most celebrated property of the linear regulator and that was almost forgotten in most of the recent nonlinear literature. At a conceptual level, though, the most important contribution of the three chapters is perhaps the general vision presented about the problem, which leaded us to formulate the chicken-egg dilemma and to become aware of the fact that the community was essentially avoiding to face the problem. We thus presented sufficient conditions for the existence of regulators of the post-processing type (Section 2.3), by showing how with very few additional effort we could deal with additional measured outputs and control inputs, thus overcoming one of the most annoying conceptual limitations of the existing approaches. Even if it is true that the regulator presented in Section 2.3 does not provide conditions easy enough to qualify as "constructive", the overall insight on the problem permitted us to realize that the chicken-egg dilemma could be dealt with using adaptation.

In this direction, Chapter 4 presented an original framework where to cast adaptive control problems in system identification terms, maintaining a formal control theoretical playground. For clarity of exposition,in Chapter 5 we chose to focus on an adaptive observation problem, as anyway the problem considered owns the same essential features of the adaptive output regulation problem. In these chapters the main conceptual contribution was the way we looked at system identification schemes in terms of (hybrid) systems, by re-framing the estimation phase in terms of a stability property with respect to an ideal optimal steady state defined by the inputs. Thus, we have shown how the identifiers fitting in the proposed framework can be co-designed with high-gain observers with the resulting state estimation error that is directly related to the prediction capabilities of the identified model. This latter result is in perfect line with the system identification viewpoint, in which no such thing as a "true model" usually exists, and where, by analogy, the property of *asymptotic state estimation* (or later regulation) is idealistic and somewhat pointless. The "identifier requirement", namely the stability and regularity properties asked to an identifier to work in the proposed framework, was supported by different examples, thus showing how that requirement is not so restrictive and, on the contrary, fits nicely on the usual properties (such as the persistence of excitation) observed in the identification algorithms. Moreover, interestingly enough, the same identifier requirement comes out as a sufficient condition in all the approaches considered in the chapters 5, 6 and 7, although coming from different thoughts.

Chapter 6 is where the vision of Chapter 1 and the theory of Chapter 4 merge in a framework for adaptive nonlinear output regulation problems. The conceptual contribution of this chapter was the intuition that the chicken-egg dilemma could be faced by means of adaptation, as anyway from the identification viewpoint there is no much difference between plant's or exosystem's uncertainties and the indeterminateness coming from the chicken-egg dilemma. The general framework proposed at the beginning of the chapter consists of a generic guideline and a meta-result that, however, is quite tautological (on the other hand the more one the discussion is general the less one can conclude). Nevertheless, the guidelines were applied to state-of art classes of problems, such as minimum-phase (partial) normal forms, and to problems not solved yet, such as general multivariable linear systems and non-square multivariable (partial) normal forms. Finally, Chapter 7 presented an alternative and more powerful approach to the problem of adaptive output regulation for linear systems that uses discrete-time identifiers. Interestingly enough, the identifier was designed by following the same guidelines presented in Chapter 4, and the key to stability was, as in the other approaches, a separation of the time-scales of the learning dynamics and the controlled system's evolution.

Overall, the thesis started from the simple and plain problem and solution of the linear case and ended in the quite complex and intricate problem of the general nonlinear case, showing how the linear problem was just the tip of the iceberg and how the interplay of identification and control presented as a quite natural solution to deal with the increasing complexity that goes with the increasing generality.

Where are we going?

The work presented in this thesis leaves perhaps more open questions than the answered ones, and *"what will be the future of output regulation?"* is a challenging point that, however, deserves a little of introspection. If we look at the problem in a *single-experiment* perspective, i.e. we seek for designs that care about the results one might have by performing "isolated experiments" (essentially all the existing literature so far), the research challenges are the old ones: we need more advanced observers and stabilization techniques to deal with larger classes of systems. In recent years the regulation theory has been extended on networks (see e.g. Wieland et al., 2011; Isidori et al., 2014), and we are at the dawn of output regulation of hybrid system (Marconi and Teel, 2013; Cox et al., 2013; Carnevale et al., 2016), which represent, perhaps, the most probable subject of the closest future. Hybrid systems are interesting per se, and also the linear case is still an open, very challenging, problem, sharing many crucial features of nonlinear continuous-time systems as, for instance, the fact that the exosystem is in general far to be sufficient for the design of the internal model (Carnevale et al., 2016). Moreover, hybrid systems also candidate as a nice mathematical playground in which *multiple-experiment* problems (see below) can be described. Another hot point concerning the single-experiment class of problem is the "co-design" of stabilizers and internal model units so as to solve the chicken-egg dilemma. We used adaptation in this thesis, though that is not the unique way to proceed, and a systematic procedure for the synergistic design of the two units is a definitely open research field, still completely unexplored. For what concerns the design of adaptive regulators, we have just hit the surface of the problem, and a lot of work is needed (just think that before the regulators proposed in the last part of the thesis also the general linear case was open). Nevertheless, the main criticism that people usually moves to the adaptive approaches is that the heavy additional complexity that they introduce is not worth with respect to the relatively-small additional performance that it guarantees. As a matter of fact, phrases of the kind: *"it is sufficient to take the gains large enough to obtain the same asymptotic bound"* are very often stated when debating about the effective (dis)advantage of adaptation. Those critics are probably true, but they hold true because of the conceptual limits of the a single-experiment perspective, in which a single "MATLAB simulation" exhausts the whole spectrum of the applications of interest. Thus an interesting milestone in the future research on output regulation will be to switch from a single to a *multi-experiment* perspective, where complex real world problems can be faced by "concatenating" different regulation problems and by developing a "theory of concatenation" to do that in autonomy. Multi-experiment means that we can endow a control system with the ability to store different internal models and a "logic" (that is an internal model itself) allowing to switch from one internal model to another to accomplish an uncountable number of complex tasks. In this perspective, adaptation is key, as new internal models need to be grown, tuned and deleted in autonomy, and the "logic" that allows to build complex behavior out of their composition has to be learned from observations. If a distinction between stabilizer and internal model unit is done, this also means that the stabilizer must evolve with the internal model, for instance the "control gains" can be taken ever smaller as the knowledge on the task to execute and of the environment grows. This in turn is consistent with our everyday experience: the first time we perform a motor task we move very rigidly (the "high-gain" prevails), as soon as we get familiar with the right movements to perform, we go open-loop and a high-gain feedback is not needed anymore (a guy that, by following the single-experiment perspective, insists on taking the gains high, instead of using adaptation, will perform worse). In this perspective the integration of hybrid systems in the regulation framework seems to be necessary, as it directly enters in what would be a more suitable *representation theory* of complex tasks, for which a single ODE, such as an exosystem, is reductive. The quest of extending the theory of representation of the outside world also motivates looking towards a *stochastic framework*, so as we can represent “partial”, “approximate” or “high-level” information about the outside world without capturing all the determinism. In other words,we can go beyond the concept of “periodicity” for a more suitable notion of “recurrence” of tasks.

In conclusion, we can say that output regulation for nonlinear systems is very far to be a closed problem, and the new mathematical tools represented by (stochastic) hybrid systems project us towards a new representation theory and an underlying multi-experiment perspective that is by now essentially unexplored. This, in turn, motivates the need of a copious amount of research, with a ever more high impact in both theoretical an practical terms.

Appendices

A

Hybrid Systems

In this appendix we very briefly collect the main notions and notations about hybrid systems borrowing the formalism and the framework from (Goebel et al., 2012). For basic concepts about hybrid systems the reader is referred to (Goebel et al., 2012; Goebel and Teel, 2006; Cai et al., 2007, 2008). We also report the main notions of stability theory used in the text.

A.1 Hybrid Systems

Let $\mathcal{X}$ be a normed vector space. We represent *hybrid systems* on $\mathcal{X}$ by means of the following differential and difference inclusions (Goebel et al., 2012):

$$\mathcal{H} : \begin{cases} \dot{x} & \in & F(x) & \quad x \in \mathcal{C} \\ x^+ & \in & G(x) & \quad x \in \mathcal{D} \end{cases} \tag{A.1}$$

where $x \in \mathcal{X}$ is the state, $F, G : \mathcal{X} \rightrightarrows \mathcal{X}$ denote respectively the *flow* and *jump* maps and $\mathcal{C}, \mathcal{D} \subset \mathcal{X}$ are the flow and jump sets. The set $\mathcal{C}$ defines the region of the state space where the trajectories may evolve according to

$$\dot{x} \in F(x).$$

The set $\mathcal{D}$ is instead the region in which the trajectories can jump according to

$$x^+ \in G(x).$$

A.1.1 Hybrid Time Domains, Hybrid Arcs and Solutions

The solutions of (A.1) are defined on *hybrid time domains*:

Definition A.1. *A subset $\mathcal{E} \subset \mathbb{R}_+ \times \mathbb{N}$ is called a* compact hybrid time domain *if $\mathcal{E}$ can be written as*

$$\mathcal{E} := \bigcup_{j=0}^{J-1} \big([t_j, t_{j+1}] \times \{j\}\big),$$

for some finite sequence $0 = t_0 \leq t_1 \leq \cdots \leq t_J$. It is called a hybrid time domain *if for all $(T, J) \in \mathcal{E}$, the set $\mathcal{E} \cap ([0, T] \times \{0, 1, \ldots, J\})$ is a compact hybrid time domain.*

If $\mathcal{E} \subset \mathbb{R}\times\mathbb{N}$ is a hybrid time domain and $(t, j), (s, i) \in \mathcal{E}$, we write $(t, j) \prec (s, i)$ if $t + j < s + i$. The symbols $\prec$, $=$, $\succ$ and $\succeq$ are defined in a similar way. We will also frequently use the shortcut $0 = (0, 0)$. With $T \in \mathbb{R}_+$ we let $\mathcal{E}|_{\geq T} := \{(t, j) \in \mathcal{E} : t + j \geq T\}$, $\sup_{\mathrm{t}} \mathcal{E} := \sup\{t \in \mathbb{R}_+ : (t, j) \in \mathcal{E}\}$ and $\sup_{\mathrm{j}} \mathcal{E} := \sup\{j \in \mathbb{N} : (t, j) \in \mathcal{E}\}$.

Hybrid arcs are functions defined on hybrid time domains. Given a hybrid arc $\varphi : \operatorname{dom}\varphi \to \mathcal{X}$ and a $(t, j) \in \operatorname{dom}\varphi$, we let

$$\begin{aligned}
t_j &:= \inf_{t\in\mathbb{R}} (t, j) \in \operatorname{dom}\varphi \\
t^j &:= \sup_{t\in\mathbb{R}} (t, j) \in \operatorname{dom}\varphi \\
j_t &:= \inf_{j\in\mathbb{N}} (t, j) \in \operatorname{dom}\varphi \\
j^t &:= \sup_{j\in\mathbb{N}} (t, j) \in \operatorname{dom}\varphi.
\end{aligned}$$

We denote by $\Gamma(\varphi)$ the set of $(t, j) \in \operatorname{dom}\varphi$ such that $(t, j + 1) \in \operatorname{dom}\varphi$ and by $\mathcal{I}(\varphi)$ the set of $(t, j) \in \operatorname{dom}\varphi$ such that $t \in (t_j, t^j)$ and we let

$$\operatorname{length}(\varphi) := \sup_{\mathrm{t}} \operatorname{dom}\varphi + \sup_{\mathrm{j}} \operatorname{dom}\varphi.$$

If $\varphi : \operatorname{dom}\varphi \to \mathcal{X}$ is a hybrid arc converging to a point $\bar{\varphi} \in \mathcal{X}$, we write $\varphi(t,j) \to \bar{\varphi}$ as a shortcut for

$$\lim_{\substack{(t,j)\in\operatorname{dom}\varphi \\ t+j\to\infty}} \varphi(t,j) = \bar{\varphi}.$$

We also write

$$\limsup \varphi$$

as a short for

$$\limsup_{\substack{(t,j)\in\operatorname{dom}\varphi \\ t+j\to\infty}} \varphi(t,j).$$

If $(\varphi^n)_n$ is a sequence of hybrid arcs such that the sequence $(\operatorname{graph}\varphi^n)_n$ converges to the graph $\operatorname{graph}\bar{\varphi}$ of a hybrid arc $\bar{\varphi}$, then we write

$$\bar{\varphi} = \text{gph-lim}_{n\to\infty} \varphi^n.$$

We say that the hybrid arc φ fulfills an *average dwell-time condition* (Hespanha and Morse, 1999) with parameters $(\lambda, N_0) \in \mathbb{R}_+ \times \mathbb{N}$ if

$$\begin{aligned} &\forall (t,j), (s,i) \in \operatorname{dom}\varphi, \\ &(s,i) \prec (t,j) \implies j - i \leq \lambda(t-s) + N_0. \end{aligned} \tag{A.2}$$

he condition (A.2) ensures persistence of flow intervals in φ. We say that a hybrid arc φ satisfies a *reverse average dwell-time condition* (Hespanha et al., 2005) with parameters $(r, N) \in (\mathbb{R}_+)^2$ if

$$\begin{aligned} &\forall (t,j), (s,i) \in \operatorname{dom}\varphi, \\ &(s,i) \prec (t,j) \implies t - s \leq r(j-i) + N. \end{aligned} \tag{A.3}$$

The reverse average dwell-time condition (A.3) ensures persistence of jumps.

We say that a sequence $(\varphi^n)_n$ of hybrid arcs $\varphi^n : \operatorname{dom}\varphi^n \to \mathcal{X}$ is *locally eventually bounded* if for any $m > 0$ there exists a compact set $X \subset \mathcal{X}$ and $n_0 >\in \mathbb{N}$ such that $\varphi^n(t,j) \in K$ for all $n > n_0$ and all $(t,j) \in \operatorname{dom}\varphi^n$ with $t + j < m$. The sequence is said to be *eventually bounded* if there exists a compact set $X \subset \mathcal{X}$ and $n_0 >\in \mathbb{N}$ such that $\varphi^n(t,j) \in K$ for all $n > n_0$ and all $(t,j) \in \operatorname{dom}\varphi^n$.

A solution to a hybrid system of the form (A.1) is a hybrid arc satisfying the following definition:

Definition A.2 (Solution to a hybrid system)**.** *A hybrid arc* $\varphi : \operatorname{dom}\varphi \to \mathcal{X}$ *is a solution to* (A.1) *if* $\varphi(0) \in \overline{\mathcal{C}} \cup \mathcal{D}$ *and*

1. *For all* $j \in \mathbb{N}$ *such that* $I^j := \{t \in \mathbb{R}_+ : (t,j) \in \operatorname{dom}\varphi\}$ *has non-empty interior*

$$\begin{aligned} &\varphi(t,j) \in \mathcal{C} && \forall t \in \operatorname{int} I^j \\ &\dot{\varphi}(t,j) \in F(\varphi(t,j)) && \textit{a.e. in } I^j. \end{aligned}$$

2. *For all* $(t,j) \in \operatorname{dom}\varphi$ *such that* $(t, j+1) \in \operatorname{dom}\varphi$

$$\begin{gathered} \varphi(t,j) \in \mathcal{D} \\ \varphi(t,j+1) \in G(\varphi(t,j)). \end{gathered}$$

We call a solution *maximal* if it cannot be extended further and *complete* if its time domain is unbounded. For a hybrid system $\mathcal{H}$, $\mathcal{S}_{\mathcal{H}}(X)$ denotes the set of all the maximal solutions of $\mathcal{H}$ originating in $X \subset \mathcal{X}$. We say that $\mathcal{H}$ is *forward complete* from $X \subset \mathcal{X}$ if every maximal solution originating in X is complete. Let $x \in \mathcal{S}_{\mathcal{H}}$ and

$$V : \mathcal{X} \to \mathbb{R}.$$

We define the *Dini derivative* of V along the flow of (A.1) at $(t,j) \in \mathcal{I}(x)$ as

$$D^+V(x(t,j)) = \limsup_{h \to 0^+} \frac{1}{h}\big(V(x(t+h,j)) - V(x(t,j))\big)$$

and we often omit (t,j) and write $D^+V(x)$ when obvious. We also let

$$\dot{V}(x) := D^+V(x)$$

whenever $\limsup$ can be substituted with $\lim$. We also use the short notation

$$x(t,j)^+ := x(t,j+1),$$

whenever $(t,j) \in \Gamma(x)$.

A.1.2 Nominal Well-Posedness and the Hybrid Basic Conditions

Nominal well-posedness, as defined in (Goebel et al., 2012), is essentially a sequential compactness requirement on the set of solutions of a hybrid system.

Definition A.3. *A hybrid system $\mathcal{H}$ is called* nominally well-posed *if the following holds: for every graphically convergent sequence*[1] $(\varphi^n)_n$ *of solutions to $\mathcal{H}$ with $\varphi^n(0) \to \xi \in \mathcal{X}$,*

1. *if the sequence $(\varphi^n)_n$ is locally eventually bounded then the sequence $(\mathrm{length}(\varphi^n))_n$ is either convergent or properly divergent to ∞ and*

$$\varphi = \text{gph-lim}_{n\to\infty} \varphi^n$$

   is a solution to $\mathcal{H}$ with $\varphi(0) = \xi$ and $\mathrm{length}(\varphi) = \lim_{n\to\infty} \mathrm{length}(\varphi^n)$.

2. *If the sequence $(\varphi^n)_n$ is not locally eventually bounded then there exists a $m \in \mathbb{R}_+^*$ for which there exists $(t_n, j_n) \in \mathrm{dom}\, \varphi^n$ such that $\lim_{n\to\infty} |\varphi^n(t_n, j_n)| = \infty$ and*

$$\varphi = (\text{gph-lim}_{n\to\infty} \varphi^n)\,|_{t+j<m}$$

   is a maximal solution to $\mathcal{H}$ with $\mathrm{length}(\varphi) = m$ and

$$\lim_{t\to\sup_t \dim \varphi} |\varphi(t, \sup_j \mathrm{dom}\, \varphi)| = \infty.$$

The following conditions (Goebel et al., 2012) are basic regularity conditions on the data that allow to conclude nominal well-posedness of a hybrid system:

Definition A.4. *A hybrid system of the form* (A.1) *is said to satisfy the* hybrid basic conditions *if:*

1. *$\mathcal{C}$ and $\mathcal{D}$ are closed.*

2. *F is outer semicontinuous*[2] *and locally bounded relative to $\mathcal{C}$, $\mathcal{C} \subset \mathrm{dom}\, F$, and $F(x)$ is convex for every $x \in \mathcal{C}$.*

[1] i.e. the sequence of sets obtained by taking the graphs of the elements of the sequence of solutions converges in the usual set-theoretical sense.

[2] A set-valued map $M : \mathcal{X} \rightrightarrows \mathcal{X}$ is outer semicontinuous in $\mathcal{X}$ if, for all $x \in \mathcal{X}$ and all sequences $(x_n)_n$ and $(y_n)_n$ such that $x_n \to x$, $y_n \in M(x_n)$ and $y_n \to y$, we have $y \in M(x)$.

3. *G is outer semicontinuous and locally bounded relative to $\mathcal{D}$, and $\mathcal{D} \subset \operatorname{dom} G$.*

Theorem A.1. *If* (A.1) *satisfies the hybrid basic conditions then it is nominally well-posed.*

A.1.3 Hybrid Systems with Inputs

Let $\mathcal{U}$ be a normed vector space, a *hybrid input* (Cai and Teel, 2009) $u : \operatorname{dom} u \to \mathcal{U}$ is defined as a Lebesgue measurable and locally essentially bounded hybrid arc. We write a hybrid system with input as[3]

$$\mathcal{H}_u : \begin{cases} \dot{x} & \in & F(x,u) & (x,u) \in \mathcal{C} \\ x^+ & \in & J(x,u) & (x,u) \in \mathcal{D} \end{cases} \tag{A.4}$$

Definition A.5. *With $x : \operatorname{dom} x \to \mathcal{X}$ and $u : \operatorname{dom} u \to \mathcal{U}$, a pair (x,u) is called a solution pair to* (A.4) *if $\operatorname{dom} x = \operatorname{dom} u$, $(x(0), u(0)) \in \mathcal{C} \cap \mathcal{D}$ and*

1. *For almost all $(t,j) \in \mathcal{I}(x,u)$, $(x(t,j), u(t,j)) \in \mathcal{C}$ and $\dot{x}(t,j) \in F(x(t,j), u(t,j))$.*

2. *For all $(t,j) \in \Gamma(x,u)$, $(x(t,j), u(t,j)) \in \mathcal{D}$ and $x(t,j+1) \in G(x(t,j), u(t,j))$.*

With slight abuse of notation, we let $\mathcal{S}_{\mathcal{H}_u}(X)$ denote the set of all maximal solution pairs of (A.4) with $x(0) \in X$. For a hybrid input u and a time instant $(t,j) \in \operatorname{dom} u$, we let

$$|u|_{(t,j)} := \max \left\{ \operatorname*{ess.sup}_{\substack{(s,i) \in \operatorname{dom} u / \Gamma(u) \\ (0,0) \preceq (s,i) \preceq (t,j)}} |u(s,i)|, \sup_{\substack{(t,j) \in \Gamma(u), \\ (0,0) \preceq (s,i) \preceq (t,j)}} |u(s,i)| \right\}$$

and we let

$$|u|_\infty = |u|_{(t,j)}$$

whenever $t + j \to \infty$.

A.2 Stability Notions

In this section we will recall the main definition related to stability theory. We will consider both autonomous systems of the form (A.1), referred to as $\mathcal{H}$, and

[3]For basic literature on hybrid systems with input and notions about input-to-state stability we refer to (Cai and Teel, 2009, 2013).

systems with inputs (A.4), referred to as $\mathcal{H}_u$, and we will make reference to the stability properties of a given closed set $\mathcal{A} \subset \mathcal{X}$.

Definition A.6 (Invariance notions). *Given a hybrid system $\mathcal{H}$, the set $\mathcal{A}$ is said to be:*

- **Weakly forward invariant** *if for every $x_0 \in \mathcal{A}$ there exists at least a complete solution $x \in \mathcal{S}_{\mathcal{H}}(x_0)$ such that $x(\tau) \in \mathcal{A}$ for all $\tau \in \operatorname{dom} x$.*
- **Weakly backward invariant** *if for every $x_0 \in \mathcal{A}$ and every $T > 0$, there exists a least one $x \in \mathcal{S}_{\mathcal{H}}(\mathcal{A})$ such that $x(t_0, j_0) = x_0$ for some $(t_0, j_0) \in \operatorname{dom} x$ fulfilling $t_0 + j_0 \geq T$ and such that $x(t, j) \in \mathcal{A}$ for all $(t, j) \preceq (t_0, j_0)$.*
- **Weakly invariant** *if both weakly forward and backward invariant.*
- **Forward invariant** *if for every $x_0 \in \mathcal{S}_{\mathcal{H}}(\mathcal{A})$, $x(\tau) \in \mathcal{A}$ for all $\tau \in \operatorname{dom} x$.*
- **Backward invariant** *if for every $x_0 \in \mathcal{A}$, every solution x of $\mathcal{H}$ that fulfills $x(t_0, j_0) = x_0$ for some $(t_0, j_0) \in \operatorname{dom} x$ also fulfills $x(t, j) \in \mathcal{A}$ for all $(t, j) \preceq (t_0, j_0)$.*
- **Invariant** *if both forward and backward invariant.*

Definition A.7 (Attractiveness notions). *Given a hybrid system $\mathcal{H}$ and a subset $X \subset \mathcal{X}$, the set $\mathcal{A}$ is said to be:*

- **Pre-attractive** *from X if every $x \in \mathcal{S}_{\mathcal{H}}(X)$ is bounded and if complete $|x(t, j)|_{\mathcal{A}} \to 0$.*
- **Attractive** *from X if pre-attractive from X and $\mathcal{H}$ is forward complete from X.*
- **Uniformly pre-attractive** *if pre-attractive from X and for each $\epsilon > 0$ there exists $T > 0$ such that, for all $x \in \mathcal{S}_{\mathcal{H}}(X)$ with $\operatorname{length}(\operatorname{dom} x) \geq T$, it holds that $|x(t, j)|_{\mathcal{A}} \leq \epsilon$ for all $(t, j) \in \operatorname{dom} x|_{\geq T}$.*
- **Uniformly attractive** *from X if uniformly pre-attractive from X and $\mathcal{H}$ is forward complete from X.*

Definition A.8 (Stability notions). *Given a hybrid system $\mathcal{H}$, the set $\mathcal{A}$ is said to be:*

- **Stable** *if for every $\epsilon > 0$ there exists $\delta > 0$ such that $x \in \mathcal{S}_{\mathcal{H}}(\mathcal{A} + \delta\mathbb{B})$ implies $|x(t, j)|_{\mathcal{A}} \leq \epsilon$ for all $(t, j) \in \operatorname{dom} x$.*

- **Pre-asymptotically stable** *from X if stable and pre-attractive from X.*

- **Uniformly pre-asymptotically stable** *from X if stable and uniformly pre-attractive from X.*

- **Asymptotically stable** *from X if stable and attractive from X.*

- **Uniformly asymptotically stable** *from X if stable and uniformly attractive from X.*

As for continuous-time systems (see the notation section), for $\tau \geq 0$, we define the *τ-reachable set* of $\mathcal{H}$ from a subset $X \subset \mathcal{X}$ as

$$\mathcal{R}^{\tau}_{\mathcal{H}}(X) := \Big\{ \bar{x} \in \mathcal{X} \,:\, \bar{x} = x(t,j),\ x \in \mathcal{S}_{\mathcal{H}}(X),\ (t,j) \in \operatorname{dom} x|_{\geq \tau} \Big\},$$

and the Ω-limit set of X as

$$\Omega_{\mathcal{H}}(X) := \lim_{\tau \to \infty} \mathcal{R}^{\tau}_{\mathcal{H}}(X),$$

which also can written as

$$\begin{aligned} \Omega_{\mathcal{H}}(X) &= \bigcap_{\tau > 0} \mathcal{R}^{\tau}_{\mathcal{H}}(X) \\ &= \Big\{ \bar{x} \in \mathcal{X} \,:\, x^n(t_n, j_n) \to \bar{x},\ x^n \in \mathcal{S}_{\mathcal{H}}(X),\ (t_n, j_n) \in \operatorname{dom} x^n,\ t_n + j_n \to \infty \Big\}. \end{aligned}$$

A system $\mathcal{H}$ is said to be *uniformly eventually bounded* from X if there exist a compact set $K \subset \mathcal{X}$ and $\tau \geq 0$ such that

$$\mathcal{R}^{\tau}_{\mathcal{H}}(X) \subset K.$$

We have an analogous of Proposition 3.8 for hybrid systems (which is a direct consequence of (Goebel et al., 2012, Prop. 6.26)):

Proposition A.1. *Let $\mathcal{H}$ be nominally well-posed, then $\Omega_{\mathcal{H}}(X)$ exists and is closed. If $\mathcal{H}$ is uniformly eventually bounded from X, $\Omega_{\mathcal{H}}(X)$ is compact. If there exists at least a complete solution inside $\mathcal{S}_{\mathcal{H}}(X)$, then $\Omega_{\mathcal{H}}(X)$ is non empty, weakly backward invariant and uniformly pre-attractive from X (and it is the smallest (in the sense of inclusion) closed set with this latter property). If in addition $\Omega_{\mathcal{H}}(X) \subset \operatorname{int} X$, then it is stable, and hence pre-asymptotically stable.*

Definition A.9. *A continuous function* $\varpi : \mathcal{X} \to \mathbb{R}_+$ *is said to be a* proper indicator *for* $\mathcal{A}$ *if* $\varpi(\mathcal{A}) = \{0\}$ *and if* $\varpi(x) \to \infty$ *as* $|x|_{\mathcal{A}} \to \infty$.

Definition A.10 (Input-to-state stability)**.** *Given a hybrid system with input* $\mathcal{H}_u$ *of the form* (A.4)*,* $\mathcal{H}_u$ *is said to be* input-to-state stable (ISS) *with respect to the set* $\mathcal{A}$ *and relative to the input* u *if there exists a proper indicator* ϖ *of* $\mathcal{A}$*,* $\beta \in \mathcal{KLL}$ *and* $\rho \in \mathcal{K}$ *such that, for all* $(x, u) \in \mathcal{S}_{\mathcal{H}_u}$*, the following bound holds*

$$\varpi(x(t,j)) \leq \max\left\{\beta(\varpi(x(0)), t, j), \rho(|u|_{(t,j)})\right\},$$

for all $(t, j) \in \mathrm{dom}(x, u)$.

B

Elements of Wavelet Theory

In this appendix we report the main elements behind the wavelet decomposition in Hilbert spaces. The content related to bases and approximation is mainly taken from (Deutsch, 2001; Christensen, 2008). The content about basic wavelet theory, a part from some easy derivations, is taken from (Daubechies, 1988; Daubechies and Lagarias, 1991; Daubechies, 1992; Daubechies and Lagarias, 1992; Cohen et al., 1992; Strang and Nguyen, 1996; Walnut, 2002). In the rest of the section $\mathcal{H}$ will denote a (possibly infinite-dimensional) Hilbert space, $\langle \cdot, \cdot \rangle$ will denote an inner product on $\mathcal{H}$ and $|\cdot|$ a norm induced by the inner product.

B.1 Elements of Hilbert Spaces

B.1.1 Orthonormal Bases in Hilbert Spaces

The usual concept of orthonormal bases in finite-dimensional spaces can be easily defined also for the infinite-dimensional case. A crucial difference though, is that usually bases in infinite-dimensional spaces come with an infinite number of elements.

Definition B.1. *Let* $\{e_k\}_{k=1}^{\infty}$ *be a sequence of elements of* $\mathcal{H}$.

- *The sequence* $\{e_k\}$ *is said to be a* (Schauder) basis *for* $\mathcal{H}$ *if for each* $f \in \mathcal{H}$ *there exists an unique sequence* $\{c_k\}_{k=1}^{\infty}$ *of scalars such that*

$$f = \sum_{k=1}^{\infty} c_k e_k \tag{B.1}$$

- *A basis* $\{e_k\}$ *is said to be an* unconditional basis *if* (B.1) *converges unconditionally for any* $f \in \mathcal{H}$.

- *A basis* $\{e_k\}$ *is an* orthonormal basis *if* $\langle e_i, e_j \rangle = \delta_{i,j}$.

Remark B.1. The "=" in (B.1) is to be intended in the sense that the series $\sum_{k=1}^{\infty} c_k e_k$ is convergent with sum f, i.e.

$$\left| f - \sum_{k=1}^{n} c_k e_k \right| \to 0 \quad \text{as} \quad n \to \infty \ .$$

△

Orthonormal bases admit the following characterization

Theorem B.1. *The following are equivalent:*

1. $\{e_k\}_{k=1}^{\infty}$ *is an orthonormal basis for* $\mathcal{H}$.

2. $\forall f \in \mathcal{H}$, $f = \sum_k \langle f, e_k \rangle e_k$.

3. $\forall f \in \mathcal{H}$, $\sum_k |\langle f, e_k \rangle|^2 = |f|^2$. *(Parseval's Equation)*

4. $\langle f, e_k \rangle = 0 \ \forall k \in \mathbb{N} \implies f = 0$.

5. $\overline{\text{span}}\{e_k\} = \mathcal{H}$.

Therefore if $\{e_k\}$ is an orthonormal basis then each $f \in \mathcal{H}$ has the *unique* expansion

$$f = \sum_{k=1}^{\infty} \langle f, e_k \rangle e_k \ . \tag{B.2}$$

Moreover it can be proved that (B.2) converges unconditionally.

When $\mathcal{H}$ is a functions pace an important class of basis for $\mathcal{H}$ is those obtained by taking functions that are translated (and possibly dilated) copies of a single given function. Translations and dilations of a function are given by the following operators

Definition B.2 (Dilation & Translation). *Let $\mathcal{H}$ be a space of functions $f : I \subseteq \mathbb{R}^d \to \mathbb{R}$, $d > 0$. Let $a \in \mathbb{R}$ and $k \in \mathbb{Z}^d$, then we define the operators*

- Dilation Operator: $D_a f(x) := a^{1/2} f(ax)$.
- Translation Operator: $T_k f(x) := f(x - k)$.

Definition B.3 (Orthonormal basis of translates). *Let $\mathcal{H}$ be a space of functions $\mathbb{R}^d \to \mathbb{R}$, $d > 0$. An* orthonormal basis of translates *for $\mathcal{H}$ is a basis of the form $\{T_k g(x)\}_{k \in \mathbb{Z}^d}$, with $g \in \mathcal{H}$.*

B.1.2 Riesz Bases in Hilbert Spaces

If a Hilbert space $\mathcal{H}$ has an orthonormal basis then, as stated by the following theorem, it has an infinite number of orthonormal bases all linked by unitary operators.

Theorem B.2. *Let $\{e_k\}_{k \in \mathbb{N}}$ be an orthonormal basis for $\mathcal{H}$, then all the orthonormal basis of $\mathcal{H}$ are given by sequences of the form $\{Ue_k\}_{k \in \mathbb{N}}$, with $U : \mathcal{H} \to \mathcal{H}$ a unitary operator.*

Unitary operators are bijective and isometric, and hence the above theorem is quite intuitive. In general we can think to transform a basis by a bijective operator that is not necessarily an isometry. In this way we obtain new objects that are non-orthonormal bases.

Definition B.4 (Riesz Bases). *A Riesz basis for $\mathcal{H}$ is a family of the form $\{Ue_k\}_k$ with $\{e_k\}_k$ an orthonormal basis and $U : \mathcal{H} \to \mathcal{H}$ a bounded bijective operator.*

As for orthonormal bases also for a Riesz basis $\{h_k\}_k$, for each $f \in \mathcal{H}$, there exists a unique family of scalars $\{c_k\}_k$ such that $f = \sum_k c_k h_k$.

Theorem B.3. *Let $\{h_k\}_k$ be a Riesz basis for $\mathcal{H}$ then there exists a* unique *Riesz basis $\{\tilde{h}_k\}_k$ for $\mathcal{H}$ such that for all $f \in \mathcal{H}$*

$$f = \sum_k \langle f, \tilde{h}_k \rangle h_k \tag{B.3}$$

and the series (B.3) *converges unconditionally.*

The unique Riesz basis $\{\tilde{h}_k\}_k$ is called the **dual basis** of $\{h_k\}_k$ and if $\{h_k\}_k$ is given by $\{h_k\}_k = \{Ue_k\}_k$, with $\{e_k\}_k$ an orthonormal basis of $\mathcal{H}$ and $U : \mathcal{H} \to \mathcal{H}$ a bounded bijective operator, then

$$\{\tilde{h}_k\}_k = \{(U^{-1})^* h_k\}_k$$

where $(U^{-1})^*$ denotes the adjoint of U^{-1}. Moreover it can be shown that the dual of $\{\tilde{h}_k\}_k$ is $\{h_k\}_k$, hence they are usually referred to as a *pair of dial Riesz bases*. A pair of dual Riesz bases have a nice property called the **biorthogonality**:

Definition B.5 (Biorthogonal sequences)**.** *Two sequences $\{h_k\}_k$ and $\{g_k\}_k$ in a Hilbert space are said to be* biorthogonal *if*

$$\langle h_k, g_j \rangle = \delta_{k,j} \ .$$

The following result says that dual Riesz bases are biorthogonal

Theorem B.4. *Let $\{h_k\}_k$ and $\{\tilde{h}_k\}_k$ be a pair of dual Riesz bases of $\mathcal{H}$, then*

(a) $\{h_k\}_k$ and $\{\tilde{h}_k\}_k$ are biorthogonal

(b) for any $f \in \mathcal{H}$

$$f = \sum_k \langle f, \tilde{h}_k \rangle h_k = \sum_k \langle f, h_k \rangle \tilde{h}_k$$

(c) There exist $A, B > 0$ such that the following Frame Condition *is satisfied*

$$A|f|^2 \le \sum_k |\langle f, h_k \rangle|^2 \le B|f|^2 \tag{B.4}$$

for all $f \in \mathcal{H}$.

The Frame condition generalizes the Parseval's equation and it can be proved that the maximum A and the minimum B for which (B.4) holds are given by $A =$

$1/(|U^{-1}|^2)$ and $B = |U|^2$. It is easy to se how Riesz bases generalize orthonormal bases, if U is unitary then necessarily $|U| = 1$ and $|U^{-1}| = 1$ and (B.4) reduces to the Parseval's Equation (point 3 of theorem B.1). Thus $\{h_k\}_k$ is orthonormal and $h_k = \tilde{h}_k$. Finally there is the following characterization for Riesz bases of translates in L_2.

Theorem B.5. *Let $\phi(x) \in L_2(\mathbb{R})$ be compactly supported, let $\{T_k\phi(x)\}_{k\in\mathbb{Z}}$ be a Riesz basis for $S \subset L_2(\mathbb{R})$, if there exists $\tilde{\phi}(x) \in L_2(\mathbb{R})$ such that $\{T_k\tilde{\phi}(x)\}_{k\in\mathbb{Z}}$ is biorthogonal to $\{T_k\phi(x)\}_{k\in\mathbb{Z}}$ then*

(a) for every $f \in S$

$$f(x) = \sum_{k\in\mathbb{Z}} \langle f, T_k\tilde{\phi}\rangle T_k\phi(x) \, .$$

(b) for every $f \in S$ there exist $A, B > 0$ such that

$$A|f|^2 \le \sum_{k\in\mathbb{Z}} |\langle f, T_k\tilde{\phi}\rangle|^2 \le B|f|^2 \, .$$

B.2 Best Approximations in Hilbert Spaces

Definition B.6 (Best Approximation)**.** *Let $A \subset \mathcal{X}$ be a subset of a pre-Hilbert Space $\mathcal{X}$ and let $x \in \mathcal{X}$. An element $a \in A$ is called a* best approximation *to x from A if*

$$|x|_A = |x - a|.$$

We denote the (possibly empty) set of all best approximation to x from A as

$$P_A(x) := \{a \in A \, : \, |x|_A = |x - a|\} \, .$$

The map P_A is called the **Metric Projection** onto A. For subspaces of Hilbert Spaces there is the following characterization

Proposition B.1. *Let A be a subspace of a Hilbert space $\mathcal{H}$, then for all $x \in \mathcal{H}$, $P_A(x)$ has exactly one element and*

(a) $x = P_A(x) + P_{A^\perp}(x)$ (or equivalently $P_A + P_{A^\perp}$ equals the identity on $\mathcal{H}$).

(b) $\mathcal{H} = A \oplus A^\perp$.

(c) $a \in P_A(x) \iff x - a \in A^{\perp} \iff \langle x - a, a \rangle = 0.$

Furthermore, if A is a finite-dimensional subspace of $\mathcal{H}$, the following holds

Proposition B.2. *Let A be a n-dimensional subspace of $\mathcal{H}$, let $\{y_1, \ldots, y_n\}$ be a basis for A. Then, for all $x \in \mathcal{H}$*

$$P_A(x) = \sum_{i=1}^{n} \alpha_i y_i$$

where $\alpha_1, \ldots, \alpha_n$ are the unique solution to the ***normal equations***

$$\sum_{i=1}^{n} \alpha_i \langle y_i, y_j \rangle = \langle x, y_j \rangle, \qquad j = 1, \ldots, n \tag{B.5}$$

If in particular $\{y_1, \ldots, y_n\}$ is an orthonormal system, *then $\alpha_i = \langle x, y_i \rangle$, and*

$$P_A(x) = \sum_{i=1}^{n} \langle x, a_i \rangle a_i \ . \tag{B.6}$$

For infinite-dimensional subspaces of $\mathcal{H}$ the following similar result holds.

Theorem B.6. *Let M be a complete subspace of $\mathcal{H}$ and let E be an orthonormal basis for M, then for all $x \in \mathcal{H}$*

$$P_M(x) = \sum_{e \in E} \langle x, e \rangle e \ . \tag{B.7}$$

The following result, known as the **"Reduction Principle"**, says that given a subset M of $\mathcal{H}$ and a subset K of M, then to find the projection of a point $x \in \mathcal{H}$ onto K one can consider the composition of a projection of x to M and a new projection onto K, and these projections commute.

Theorem B.7 (Reduction Principle)**.** *Let K be a convex subset of $\mathcal{H}$ and M be any closed subset of $\mathcal{H}$ containing K, then for every $x \in \mathcal{H}$*

(a) $P_K(x) = (P_K \circ P_M)(x) = (P_M \circ P_K)(x).$

(b) $|x|_K^2 = |x|_M^2 + |P_M(x)|_K^2.$

B.3 Basics of Wavelet Analysis

B.3.1 Generalized Multiresolution Analysis

We denote by $\mathcal{C}_c(\mathbb{R})$ the set of all compactly supported continuous functions $\mathbb{R} \to \mathbb{R}$. A relevant framework in which biorthogonal wavelets can be defined is the so-called *Generalized Multiresolution Analysis (GMRA)* (see for instance (Walnut, 2002)):

Definition B.7 (Generalized Multiresolution Analysis). *A GMRA on $\mathbb{R}$ is a sequence of subspaces $(V_i)_{i\in\mathbb{Z}}$ of $L_2(\mathbb{R})$ satisfying the following properties:*[1]

a) For all $i \in \mathbb{Z}$, $V_i \subset V_{i-1}$.

b) For any $f \in \mathcal{C}_c(\mathbb{R})$ and every $\epsilon > 0$, there exists $i \in \mathbb{Z}$ and a function $g \in V_i$ such that $|f - g| \leq \epsilon$.

c) $\cap_{i\in\mathbb{Z}} V_i = \{0\}$.

d) $f \in V_i$ if and only if $f(2^i \cdot) \in V_0$.

e) There exists a function $\upsilon \in L_2(\mathbb{R})$, called the scaling function *such that $V_0 = \overline{\text{span}}\{\upsilon(\cdot - k)\}_{k\in\mathbb{Z}}$ and $\{\upsilon(\cdot - k)\}_{k\in\mathbb{Z}}$ is a* Riesz basis *for V_0.*

For a function $f \in L_2(\mathbb{R})$ and with $i, k \in \mathbb{Z}$, we define the function $f_{i,k} \in L_2(\mathbb{R})$ as

$$f_{i,k}(s) = 2^{-i/2} f(2^{-i}s - k).$$

If υ is a scaling function of a GMRA $(V_i)_{i\in\mathbb{Z}}$, then we have (Walnut, 2002, Lem. 10.17) that

$$V_i = \overline{\text{span}}\{\upsilon_{i,k}\}_{k\in\mathbb{Z}}$$

and $\{\upsilon_{i,k}\}_{k\in\mathbb{Z}}$ is a Riesz basis for V_i. As the family $\{\upsilon(\cdot - k)\}_{k\in\mathbb{Z}}$ is a Riesz basis of V_0, it admits a *dual basis* $\{\tilde{\upsilon}(\cdot - k)\}_{k\in\mathbb{Z}}$. If such a dual basis is itself the scaling function of a GMRA $(\tilde{V}_i)_{i\in\mathbb{Z}}$ we say that the GMRAs $(V_i)_{i\in\mathbb{Z}}$ and $(\tilde{V}_i)_{i\in\mathbb{Z}}$ are *dual* to each other.

[1] In the literature the nesting order is taken in both the directions, in the sense that some authors (e.g. (Walnut, 2002)) use higher values of i to denote higher resolutions (i.e. $V_i \subset V_{i+1}$), some other (e.g. (Daubechies, 1992)) use the opposite. We chose this latter convention according to (Daubechies, 1992).

With $i \in \mathbb{Z}$ and $f \in L_2(\mathbb{R})$ we define the *projection* operator P_i and the *detail* operator Q_i as

$$\mathrm{P}_i f := \sum_{k\in\mathbb{Z}} \langle f, \tilde{\upsilon}_{i,k}\rangle \upsilon_{i,k}, \qquad \mathrm{Q}_i f := \mathrm{P}_{i-1} f - \mathrm{P}_i f$$

and the same object are defined for the dual basis

$$\tilde{\mathrm{P}}_i f := \sum_{k\in\mathbb{Z}} \langle f, \upsilon_{i,k}\rangle \tilde{\upsilon}_{i,k}, \qquad \tilde{\mathrm{Q}}_i f := \tilde{\mathrm{P}}_{i-1} f - \tilde{\mathrm{P}}_i f.$$

For a given $i \in \mathbb{Z}$, $\mathrm{Q}_{i+1} f$ represents the additional detail that is needed to obtain a *finer* approximation $\mathrm{P}_i f$ of f at scale i starting from a *coarser* approximation $\mathrm{P}_{i+1} f$ at scale $i+1$. Moreover, for every $f \in \mathcal{C}_c(\mathbb{R})$,

$$\lim_{i\to-\infty} |\mathrm{P}_i f - f| = 0,$$

so as for each $f \in \mathcal{C}_c(\mathbb{R})$ and $\epsilon > 0$ there exists $i^\star \in \mathbb{Z}$ such that, for all $i \le i^\star$, the function $\hat{f}_i \in V_i$ given by

$$\hat{f}_i := \mathrm{P}_i f = \sum_{k\in\mathbb{Z}} a_{i,k} \upsilon_{i,k}, \qquad a_{i,k} := \langle f, \tilde{\upsilon}_{i,k}\rangle, \tag{B.8}$$

satisfies

$$|f - \hat{f}_i| \le \epsilon.$$

We also observe that if υ has compact support, then the sum in (B.8) has *finite* terms for each $i \in \mathbb{Z}$.

B.3.2 Biorthogonal Wavelets

We can univocally associate to the scaling functions υ and $\tilde{\upsilon}$ respectively, two functions ψ and $\tilde{\psi}$ in $L_2(\mathbb{R})$, referred to as the *wavelet functions*, that, with $W_i := \overline{\mathrm{span}}\{\psi_{i,k}\}_{k\in\mathbb{Z}}$ and $\tilde{W}_i := \overline{\mathrm{span}}\{\psi_{i,k}\}_{k\in\mathbb{Z}}$, fulfill the following properties (see (Walnut, 2002, Lem. 10.24)):

a) $\psi \in V_{-1}$ and $\tilde{\psi} \in \tilde{V}_{-1}$.

b) $\{\psi_{0,k}\}_{k\in\mathbb{Z}}$ and $\{\tilde{\psi}_{0,k}\}_{k\in\mathbb{Z}}$ are biorthogonal.

c) $\{\psi_{0,k}\}_{k\in\mathbb{Z}}$ is a Riesz basis for W_0 and $\{\tilde{\psi}_{0,k}\}_{k\in\mathbb{Z}}$ for $\tilde{W}_0$.

d) For all $k, \ell \in \mathbb{Z}$, $\langle \psi_{0,k}, \tilde{\upsilon}_{0,\ell} \rangle = 0$ and $\langle \tilde{\psi}_{0,k}, \upsilon_{0,\ell} \rangle = 0$.

e) For all $f \in \mathcal{C}_c(\mathbb{R})$, $\mathrm{Q}_0 f \in W_0$ and $\tilde{\mathrm{Q}}_0 f \in W_0$.

Directly from the definition of P_i and Q_i we obtain that, for all $i, i_0 \in \mathbb{Z}$ such that $i < i_0$, the following holds

$$\mathrm{P}_i f = \mathrm{P}_{i_0} f + \sum_{\ell=i+1}^{i_0} \mathrm{Q}_\ell f, \qquad \tilde{\mathrm{P}}_i f = \tilde{\mathrm{P}}_{i_0} f + \sum_{\ell=i+1}^{i_0} \tilde{\mathrm{Q}}_\ell f$$

Moreover, it can be shown that, for any $f \in \mathcal{C}_c(\mathbb{R})$, the details at scale i can be expressed as combinations of wavelets. Namely

$$\mathrm{Q}_i f = \sum_{k\in\mathbb{Z}} \langle f, \tilde{\psi}_{i,k} \rangle \psi_{i,k}, \qquad \tilde{\mathrm{Q}}_i f = \sum_{k\in\mathbb{Z}} \langle f, \psi_{i,k} \rangle \tilde{\psi}_{i,k}. \tag{B.9}$$

As a consequence, we obtain the following representation[2]

$$\mathrm{P}_i f = \sum_{k\in\mathbb{Z}} a_{i_0,k} \upsilon_{i_0,k} + \sum_{\ell=i+1}^{i_0} \sum_{k\in\mathbb{Z}} b_{\ell,k} \psi_{\ell,k} \tag{B.10}$$

where, for $k \in \mathbb{Z}$ and $\ell = i+1, \ldots, 0$,

$$a_{i_0,k} := \langle f, \tilde{\upsilon}_{i_0,k} \rangle \qquad\qquad b_{\ell,k} := \langle f, \tilde{\psi}_{\ell,k} \rangle.$$

Moreover, the projection at scale $i-1$ is given by

$$P_{i-1} f = P_i f + \sum_{k\in\mathbb{Z}} b_{i,k} \psi_{i,k}.$$

From the definition of Riesz basis, there exists an orthonormal basis $\{e_k\}_{k\in\mathbb{Z}}$ of V_0 and a bounded bijective operator $U : L_2(\mathbb{R}) \to L_2(\mathbb{R})$ such that $\upsilon_{0,k}(\cdot) = U e_k(\cdot)$. Hence, for all $k, \ell \in \mathbb{Z}$, we obtain $\delta_{k,\ell} = \langle e_k, e_\ell \rangle = \langle U^{-1}\upsilon_{0,k}, U^{-1}\upsilon_{0,\ell} \rangle = \langle \upsilon_{0,k}, (U^{-1})^* U^{-1} \upsilon_{0,k} \rangle$, having denoted with $(U^{-1})^*$ the adjoint operator of U^{-1}. Hence, with $C := (U^{-1})^* U^{-1}$, if we define the scalar product $\langle \cdot, \cdot \rangle_C$ on $L_2(\mathbb{R})$ given by $\langle f, g \rangle_C := \langle f, Cg \rangle$, we have $\delta_{k,\ell} = \langle \upsilon_{0,k}, \upsilon_{0,\ell} \rangle_C$. Namely, $\{\upsilon_{0,k}(\cdot)\}_{k\in\mathbb{Z}}$ is an orthonormal basis for V_0 with respect to the scalar product $\langle \cdot, \cdot \rangle_C$. In the same way we obtain that $\{\upsilon_{i,k}(\cdot)\}_{k\in\mathbb{Z}}$ is an orthonormal with respect to $\langle \cdot, \cdot \rangle_C$ basis for

[2]A similar relation clearly also holds for $\tilde{\mathrm{P}}_i f$, however we will omit that for compactness.

V_i. Moreover, as a consequence of property d) of the wavelet function, for each $i, \ell, k \in \mathbb{Z}$ we have

$$\begin{aligned}\langle \psi_{i,k}, \upsilon_{i,\ell} \rangle_C &= \langle \psi_{i,k}, CD_{-i}\upsilon_{0,\ell} \rangle = \langle D_{-i}\psi_{0,k}, D_{-i}\tilde{\upsilon}_{0,\ell} \rangle \\ &= \langle \psi_{0,k}, \tilde{\upsilon}_{0,\ell} \rangle = 0\end{aligned}$$

that is, $\{\psi_{i,k}(\cdot)\}_{k\in\mathbb{Z}}$ and $\{\upsilon_{i,k}(\cdot)\}_{k\in\mathbb{Z}}$ are orthogonal with respect to $\langle \cdot, \cdot \rangle_C$. Therefore, we can write $V_i \perp W_i$ and $V_{i-1} = V_i \oplus W_i$, where orthogonality is with respect to the scalar product $\langle \cdot, \cdot \rangle_C$. We also have $V_{i-1} \perp W_{i-1}$, that implies $W_i \perp W_{i-1}$. As a consequence, in the scalar product $\langle \cdot, \cdot \rangle_C$, we have

$$V_i = V_0 \oplus \left(\bigoplus_{\ell=i+1}^{0} W_\ell \right),$$

which means that once a "coarse" representation of a compactly supported continuous $f \in L_2(\mathbb{R})$ is given, in terms of $\mathrm{P}_0 f(\cdot) \in V_0$, a "finer" representation can be obtained by adding the *details* $\mathrm{Q}_i f(\cdot) \in W_0$ that belong to a subspace which is *orthogonal* to V_0. The same can be iterated to find a finer representation $\mathrm{P}_{i-1} f(\cdot) \in V_i$ starting a coarser representation in V_i and always looking for the additional detail inside an orthogonal subspace. According to (B.9), the details that must be added at each stage can be written as a linear combination of scaled and translated version of the wavelet function. Moreover since $\mathcal{C}_c(\mathbb{R})$ is dense in $L_2(\mathbb{R})$ we also have

$$L_2(\mathbb{R}) = \mathrm{cl}\left(V_0 \oplus \left(\bigoplus_{\ell=\infty}^{0} W_\ell \right) \right).$$

Therefore, given any $f \in L_2$ and any $\epsilon > 0$, there exists $i^\star > 0$, such that, for every $i \leq i^\star$ the function $\mathrm{P}_i f(\cdot)$ given by (B.10) satisfies $|f - \mathrm{P}_i f| \leq \epsilon$.

B.3.3 Wavelets in Higher Dimension

The wavelet theory can be extended to deal with functions in $L_2(\mathbb{R}^m)$ with $m > 1$ by considering the tensor product of m GMRAs (see (Daubechies, 1992)). More precisely, we define the subspaces $\mathbf{V}_i$, $i \in \mathbb{Z}$ as

$$\mathbf{V}_0 := V_0 \otimes V_0 \otimes \cdots \otimes V_0$$

$$= \overline{\text{span}}\{\Upsilon(x) := \upsilon(x_1)\upsilon(x_2)\cdots\upsilon(x_m) \,:\, \upsilon \in V_0\}$$

and

$$f \in \mathbf{V}_i \iff f(2^i\cdot) \in V_0.$$

hen $(\mathbf{V}_i)_{i\in\mathbb{Z}}$ forms a GMRA in $L_2(\mathbb{R}^m)$ (i.e. satisfying analogous properties of Definition B.7), with Υ playing the role of a scaling function. Since $\{\upsilon(\cdot - k)\}_{k\in\mathbb{Z}}$ is a Riesz basis for V_0, the scaling function

$$\Upsilon(x) := \upsilon(x_1 - k_1)\upsilon(x_2 - k_2)\cdots\upsilon(x_m - k_m), \qquad \mathbf{k} := (k_1, \dots, k_m) \in \mathbb{Z}^m,$$

is a Riesz basis for $\mathbf{V}_0$. Moreover, the functions

$$\Upsilon_{i,\mathbf{k}}(x) := \upsilon_{i,k_1}(x_1)\cdots\upsilon_{i,k_m}(x_m), \qquad \mathbf{k} \in \mathbb{Z}^m$$

form a Riesz basis for $\mathbf{V}_i$.

In the same way we construct the dual GMRA $(\tilde{\mathbf{V}}_i)_{i\in\mathbb{Z}}$ and its scaling function $\tilde{\Upsilon}$. Furthermore, we construct a set of 2^m-1 wavelet functions $\Psi^h_{i,\mathbf{k}}$, $i \in \mathbb{Z}$, $\mathbf{k} \in \mathbb{Z}^m$, $h = 1, \dots, 2^m - 1$, by all the possible combinations of products of the form

$$g_1(x_1)g_2(x_2)\cdots g_m(x_m)$$

with $g_\ell(x_\ell)$ taking the value $\upsilon_{i,k_\ell}(x_\ell)$ or $\psi_{i,k_\ell}(x_\ell)$, except for the case in which $g_\ell(x_\ell) = \upsilon_{i,k_\ell}(x_\ell)$ for all $\ell = 1, \dots, m$. Namely we set

$$\begin{aligned}
\Psi^1_{i,\mathbf{k}}(x) &:= \psi_{i,k_1}(x_1)\upsilon_{i,k_2}(x_2)\cdots\upsilon_{i,k_m}(x_m)\\
\Psi^2_{i,\mathbf{k}}(x) &:= \upsilon_{i,k_1}(x_1)\psi_{i,k_2}(x_2)\upsilon_{i,k_3}(x_3)\cdots\upsilon_{i,k_m}(x_m)\\
&\cdots\\
\Psi^{2^m-1}_{i,\mathbf{k}}(x) &:= \psi_{i,k_1}(x_1)\psi_{i,k_2}(x_2)\psi_{i,k_3}(x_3)\cdots\psi_{i,k_m}(x_m)
\end{aligned}$$

Directly from the definition we obtain

$$\begin{aligned}
\mathbf{V}_{i-1} &= V_{i-1}\otimes\cdots\otimes V_{i-1} = (V_i \oplus W_i)\otimes\cdots\otimes(V_i\oplus W_i)\\
&= (V_i\otimes\cdots\otimes V_i)\oplus\left(\bigoplus_{h=1}^{2^m-1}\mathbf{Z}^h_i\right)
\end{aligned}$$

having defined

$$\begin{aligned}
\mathbf{Z}_i^1 &:= W_i \otimes V_i \otimes V_i \otimes \cdots \otimes V_i \\
\mathbf{Z}_i^2 &:= V_i \otimes W_i \otimes V_i \otimes \cdots \otimes V_i \\
&\cdots \\
\mathbf{Z}_i^m &:= V_i \otimes V_i \otimes \cdots \otimes V_i \otimes W_i \\
\mathbf{Z}_i^{m+1} &:= W_i \otimes W_i \otimes V_i \otimes \cdots \otimes V_i \\
&\cdots \\
\mathbf{Z}_i^{2^m-1} &:= W_i \otimes W_i \otimes \cdots \otimes W_i,
\end{aligned}$$

i.e. the subspaces $\mathbf{Z}_i^h$ are obtaining by taking the tensor product of all the possible combinations of m subspaces in $\{V_i, W_i\}$. A similar decomposition works for $\tilde{\mathbf{V}}_{i-1}$ as well, by opportunely defining the subspaces $\tilde{\mathbf{Z}}_i^h$.

We construct a set of basis functions for each of the subspaces $\mathbf{Z}_i^h$ by taking the tensor product of m functions in the corresponding order. Namely, we let

$$\begin{aligned}
\Psi_{i,\mathbf{k}}^1(x) &:= \psi_{i,k_1}(x_1)\upsilon_{i,k_2}(x_2)\cdots\upsilon_{i,k_m}(x_m) \\
\Psi_{i,\mathbf{k}}^2(x) &:= \upsilon_{i,k_1}(x_1)\psi_{i,k_2}(x_2)\upsilon_{i,k_3}(x_3)\cdots\upsilon_{i,k_m}(x_m) \\
&\cdots \\
\Psi_{i,\mathbf{k}}^{2^m-1}(x) &:= \psi_{i,k_1}(x_1)\psi_{i,k_2}(x_2)\psi_{i,k_3}(x_3)\cdots\psi_{i,k_m}(x_m).
\end{aligned}$$

Hence, defined

$$\mathbf{W}_i := \oplus_{h=1}^{2^d-1}\mathbf{Z}_i^h,$$

we obtain

$$\begin{aligned}
\mathbf{V}_i &= \overline{\text{span}}\{\Upsilon_{i,\mathbf{k}}(\cdot)\}_{\mathbf{k}\in\mathbb{Z}^m} \\
\mathbf{W}_i &= \overline{\text{span}}\{\Psi_{i,\mathbf{k}}^h(\cdot)\}_{\mathbf{k}\in\mathbb{Z}^m, h=1,\ldots,2^m-1}
\end{aligned}$$

and, for any $i, i_0 \in \mathbb{Z}$ such that $i < i_0$, we can express the projection of a $F \in \mathcal{C}_c(\mathbb{R}^m)$ onto $\mathbf{V}_i$ with the following expansion, that generalizes (B.10)

$$\mathrm{P}_i f = \sum_{\mathbf{k}\in\mathbb{Z}^m} a_{i_0,\mathbf{k}}\Upsilon_{i_0,\mathbf{k}} + \sum_{h=1}^{2^m-1}\sum_{\ell=i+1}^{i_0}\sum_{\mathbf{k}\in\mathbb{Z}^m} b_{i,\mathbf{k}}^h \Psi_{\ell,\mathbf{k}}^h \tag{B.11}$$

In the general case, in the wavelet expansion (B.11) the sum in $\mathbf{k}$ ranges over the whole set $\mathbb{Z}^m$. Hence, even for fixed scale i, (B.11) might consist of infinite terms. Nevertheless, if *compactly supported* biorthogonal Wavelet and Scaling functions are used, the expansion (B.11) can be reduced to a *finite* sum whenever f has bounded support.

B.4 Smoothness, Compact Support and Approximation

The multiresolution analysis introduced in the previous section is a strong tool to construct biorthogonal and orthonormal scaling functions and wavelet bases. However nothing is said about the smoothness, the support and the approximation capabilities of the basis functions that can be constructed. The importance of a bounded support is clear. In fact, even for fixed i, the sum (B.11) runs over infinite terms, since $\mathbf{k}$ varies in $\mathbb{Z}^m$. In practice, if one wants to approximate a function f having bounded support with a finite number of basis functions, then necessarily the basis functions must have bounded support. In this way, indeed, one can restrict $\mathbf{k}$ to range in a bounded subset of $\mathbb{Z}^m$ since only finitely many basis functions will have a support intersecting those of f.

Moreover, in some applications, some regularity constraints are needed for the basis functions $\Upsilon_{i,\mathbf{k}}$ and $\Psi^h_{i,\mathbf{k}}$. Therefore the question whether there exist wavelet basis having at the same time compact support and the desired smoothness properties is crucial.

B.4.1 Smooth, Compactly Supported, Orthonormal Wavelet Bases

There are several choices of scaling functions that generate a valid orthonormal GMRA (for instance the piece-wise linear GMRA, the band-limited GMRA, which is constructed by using the $sinc$ function, the Meyer GMRA and the Spline GMRA, etc.), however usually they are not smooth and compactly supported at the same time. For instance the Meyer wavelets are C^∞ but they have infinite support, the spline constructions yield wavelet in C^k, with k depending on the polynomial order, but they do not have compact support. Nevertheless, in (Daubechies, 1988), Ingrid Daubechies presented a class of compactly supported wavelet basis constructing a valid orthonormal GMRA and characterized by a arbitrarily high regularity.

The family of Daubechies bases are characterized by the fact that, for each $N \in \mathbb{N}$, there exist mother scaling and wavelet functions (denoted by υ_N and ψ_N) generating an orthonormal GMRA that have N vanishing moments and that are supported in $[0, 2N-1]$. The definition of such functions cannot be obtained analytically, they are constructed by solving the so-called *two-scales equation* (see e.g. Daubechies, 1992; Strang and Nguyen, 1996). The regularity properties are thus given in terms of *vanishing moments* and there are many results providing an estimate of the maximum derivative for each resulting wavelet basis and of its Hölder exponent.

In general, it can be proved that a wavelet $\psi(x)$ generating a valid GMRA necessarily satisfies

$$\int_{\mathbb{R}} \psi(x)dx = 0 . \tag{B.12}$$

Equation (B.12) is called the 0-th moment of $\psi(x)$. In the same way, one defines the *k-th moment* of $\psi(x)$ as

$$\int_{\mathbb{R}} x^k \psi(x)dx . \tag{B.13}$$

Then the family of bases introduced by Daubechies have the property that for each $N \in \mathbb{N}$ there exist a wavelet basis that satisfies (B.13) with $k = N$ and with support equal to $[0, 2N-1]$. The way moments are linked to smoothness is unfortunately a one side implication given by the following theorem (Walnut, 2002)

Theorem B.8. *Let ψ be such that for some $N \in \mathbb{N}$, both $x^N\psi(x)$ and $\omega^{N+1}\hat{\psi}(\omega)$ (with $\hat{\cdot}$ denoting the Fourier Transform of ψ) are $L^1(\mathbb{R})$. If $\{\psi_{j,k}(x)\}_{j,k\in\mathbb{Z}}$ is an orthonormal system then* (B.13) *hold for each $k = 0, \ldots, N$.*

Saying $\omega^{N+1}\hat{\psi}(\omega) \in L^1$ is to be interpreted as a smoothness property, in fact ((Walnut, 2002, Theorem 9.3)) $\omega^{N+1}\hat{\psi}(\omega) \in L^1$ implies that $\psi^{(N+1)}(x)$ is uniformly continuous. Unfortunately the inverse implication does not hold, in the sense that if (B.13) holds for some $k > 0$ then $\psi^{(k)}(x)$ might not even exist. Nevertheless, many different results exist (see e.g. Daubechies, 1992; Daubechies and Lagarias, 1991; Daubechies, 1992; Gripenberg, 1996) about the regularity properties of the Daubechies wavelets. Let $\alpha = n + \beta$, with $n \in \mathbb{N}$ and $\beta \in [0, 1)$ and define the function class C^α to be the set of functions f which are n-times differentiable and such that the n-th derivative $f^{(n)}$ is Hölder continuous with exponent β, namely

$$|f^{(n)}(x_1) - f^{(n)}(x_2)| \leq c|x_1 - x_2|^{\beta}$$

for all $x_1, x_2 \in \mathbb{R}$. Then it turns out that ψ_3 is continuously differentiable (or more precisely is $C^{1.0878}$), and the first Daubechies wavelet which is **twice continuously differentiable** is ψ_6 and in general, asymptotically, we have the relation

$$\upsilon_N, \psi_N \in C^{\mu N},$$

with $\mu \approx 0.2$.

Another important aspect of vanishing moments concerns the approximation capabilities for finite j in the expansion (B.10). The following theorem (Walnut, 2002) says that wavelet basis constructed with wavelet functions having many vanishing moments yield better approximations, in the sense that the coefficients $b_{j,k}$ of the expansion (B.10) decay rapidly as i decreases.

Theorem B.9. *With $N \in \mathbb{N}$, assume that $f \in C^N(\mathbb{R})$ and that $f^{(N)} \in L^{\infty}(\mathbb{R})$. Assume that ψ has compact support and*

$$\int_{\mathbb{R}} x^m \psi(x) dx = 0, \qquad 0 \leq m \leq N-1$$

and that $\int_{\mathbb{R}} |\psi_{i,k}(x)|^2 dx = 1$ for all $i, k \in \mathbb{Z}$, then there exists C, depending on N and f, such that

$$|\langle f, \psi_{i,k} \rangle| \leq C 2^{iN} 2^{i/2} \tag{B.14}$$

for all $i, k \in \mathbb{Z}$.

Finally there's another result relating the number of vanishing moments with the approximation of polynomials.

Theorem B.10. *Let $\upsilon(x)$ be a compactly supported scaling function associated with a GMRA and let $\psi(x)$ be the corresponding mother wavelet. If $\psi(x)$ has N vanishing moments then for every polynomial $p(x)$ of degree $p \leq N-1$ there exist coefficients a_k, $k \in \mathbb{Z}$ such that*

$$\sum_{k \in \mathbb{N}} a_k \upsilon(x-k) = p(x)\ .$$

B.4.2 Smooth, Compactly Supported Biorthogonal Wavelets

Biorthogonal Wavelets address many of the issues of orthonormal bases. Relaxing the tight constraint of orthonormality, indeed, yields basis functions that can easily have compact support and high regularity as well as symmetry properties. Moreover, in general, the basis $\{T_k v(x)\}_k$ and $\{T_k \tilde{v}(x)\}_k$ can also have different smoothness and support. The most famous class of biorthogonal wavelets is those obtained as combinations of B-Spline functions (Cohen et al., 1992). A B-Spline of order n is a piecewise-polynomial function obtained by the (n)-fold convolution of the indicator function

$$s^1(x) = \begin{cases} 1, & x \in [0,1) \\ 0, & \text{otherwise} \end{cases}$$

For instance the 2nd order B-spline is given by the piece-wise linear function

$$s^2(x) = (s^1 \star s^1)(x) = \begin{cases} 1 - (|x| - 1), & x \in [0,2) \\ 0, & \text{otherwise} \end{cases}$$

B-Splines have the following interesting properties

- $\operatorname{supp} s^m(x) = [0, m]$
- $s^m(x) = \dfrac{x}{m-1} s^{m-1}(x) + \dfrac{m-x}{m-1} s^{m-1}(x-1)$
- $s^m(x) \in C^{\max\{0, m-2\}}$, in particular

$$(s^m)'(x) = s^{m-1}(x) - s^{m-1}(x-1)$$

- s^m satisfies he two-scales equation

$$s^m(x) = \sum_{k=0}^{m} 2^{1-m} \binom{m}{k} s^m(2x - k)$$

B-Splines can be used to construct biorthogonal scaling functions and wavelets generating valid dual GMRAs *having compact support and arbitrarily regularity*.

The scaling function $v(x)$ and those generating the dual GMRA $\tilde{v}(x)$ can also be given by different Splines.

Bibliography

Bibliography

C. Afri, V. Andrieu, L. Bako, and P. Dufour. State and parameter estimation: A nonlinear luenberger observer approach. *IEEE Transactions on Automatic Control*, 62(2):973–980, 2017.

V. Andrieu and L. Praly. On the existence of a Kazantzis-Kravaris/Luenberger observer. *SIAM Journal on Control and Optimization*, 45(2):432–456, 2006.

D. Angeli. A lyapunov approach to incremental stability properties. *IEEE Transactions on Automatic Control*, 47(3):410–421, 2002.

D. Astolfi and L. Marconi. A high-gain nonlinear observer with limited gain power. *IEEE Transactions on Automatic Control*, 60(11):3059–3064, 2015.

D. Astolfi and L. Praly. Integral action in output feedback for multi-input multi-output nonlinear systems. *IEEE Transactions on Automatic Control*, 62(4):1559–1574, 2017.

D. Astolfi, A. Isidori, L. Marconi, and L. Praly. Nonlinear output regulation by post-processing internal model for multi-input multi-output systems. In *9th IFAC Symposium on Nonlinear Control Systems*, pages 295–300, Tououse, France, 2013.

D. Astolfi, L. Praly, and L. Marconi. Approximate regulation for nonlinear systems in presence of periodic disturbances. In *IEEE 54th Conference on Decision and Control (CDC 2015)*, pages 7665–7670, 2015.

D. Astolfi, L. Marconi, L. Praly, and A. R. Teel. Sensitivity to high-frequency measurement noise of nonlinear high-gain observers. In *10th IFAC Symposium on Nonlinear Control Systems*, pages 276–278, Monterey, CA, 2016.

D. Astolfi, L. Marconi, and A. R. Teel. Low-power peaking-free high-gain observers. *Automatica*, 98:169–179, 2018.

J. Aubin. *Viability Theory*. Springer, 1991.

J. Aubin and A. Cellina. *Differential Inclusions. Set-valued Maps and Viability Theory*. Springer-Verlag, 1984.

J. Back. An inner-loop controller guaranteeing robust transient performance for uncertain MIMO nonlinear systems. *IEEE Transactions on Automatic Control*, 54(7):1601–1607, 2009.

M. Bado and A. Ichikawa. Adaptive output regulation for linear systems. In *SICE-ICASE International Joint Conference 2006*, Bexco, Busan, Korea, 2006.

G. Basile and G. Marro. *Controlled and Conditioned Invariants in Linear System Theory*. Prentice Hall, Englewood Cliffs, New Jersey, 1992.

G. Bastin and M. Gevers. Stable adaptive observers for nonlinear time varying systems. *IEEE Transactions on Automatic Control*, 33(7):650–658, 1988.

G. Besançon. remarks on nonlinear adaptive observer design. *System & Control Letters*, 41:271–280, 2000.

G. Besançon and A. Ţiclea. An immersion-based observer design for rank-observable nonlinear systems. *IEEE Transactions on Automatic Control*, 52(1): 83–88, 2007.

G. Besançon and A. Ţiclea. On adaptive observers for systems with state and parameter nonlinearities. In *20th IFAC World Congress*, Toulouse, France, 2017.

M. Bin and L. Marconi. Adaptive output regulation via post-processing internal models and hybrid identifiers. In *IEEE 56th Annual Conference on Decision and Control (CDC 2017)*, Melbourne, Australia, 2017a.

M. Bin and L. Marconi. About a post-processing design of regression-like non-linear internal models. In *IFAC 2017 World Congress*, Toulouse, France, 2017b.

M. Bin and L. Marconi. “class-type” identification-based internal models in multivariable nonlinear output regulation, 2018a. Submitted to IEEE Transactions on Automatic Control. Under review. Preprint: arXiv:1805.05629.

M. Bin and L. Marconi. The chicken-egg dilemma and the robustness issue in nonlinear output regulation with a look towards adaptation and universal approximators. In *IEEE 57th Conference on Decision and Control (CDC)*, Miami Beach, FL, USA, 2018b.

M. Bin and L. Marconi. Adaptive observers design for nonlinear systems via hybrid identification techniques. submitted to IEEE Transactions on Automatic Control, 2018c.

M. Bin, D. Astolfi, and L. Marconi. Robust internal model design by nonlinear regression via low-power high-gain observers. In *2016 IEEE 56th Conference on Decision and Control (CDC)*, Las Vegas, USA, 2016.

M. Bin, D. Astolfi, L. Marconi, and L. Praly. About robustness of internal model-based control for linear and nonlinear systems. In *IEEE 57th Conference on Decision and Control (CDC)*, Miami Beach, FL, USA, 2018a.

M. Bin, D. Astolfi, L. Marconi, and L. Praly. About robustness of internal model-based control for linear and nonlinear systems. In *57th IEEE Conference on Decision and Control*, Miami Beach, FL, USA, 2018b.

M. Bin, L. Marconi, and A. R. Teel. Adaptive output regulation for linear systems via discrete-time identifiers. submitted to Automatica, 2018c.

C. Byrnes, F. Delli Priscoli, and A. Isidori. Structurally stable output regulation for nonlinear systems. *Automatica*, 33(3):369–385, 1997a.

C. Byrnes, F. Delli Priscoli, and A. Isidori. *Output Regulation of Uncertain Nonlinear Systems*. Birkhäuser, 1997b.

C. I. Byrnes and A. Isidori. Bifurcation analysis of the zero dynamics and the practical stabilization of nonlinear minimum-phase systems. *Asian Journal of Control*, 4(2):171–185, 2002.

C. I. Byrnes and A. Isidori. Limit sets, zero dynamics and internal models in the problem of nonlinear output regulation. *IEEE Transactions on Automatic Control*, 48:1712–1723, Oct. 2003.

C. I. Byrnes and A. Isidori. Nonlinear internal models for output regulation. *IEEE Transactions on Automatic Control*, 49:2244–2247, Dec. 2004.

C. I. Byrnes, A. Isidori, and A. Praly. On the asymptotic properties of a system arising from the non-equilibrium theory of output regulation. *Mittag Leffler Institute*, 2003.

C. Cai and A. Teel. Characterization of input-to-state stability for hybrid systems. *Systems & Control Letters*, 58:47–53, 2009.

C. Cai and A. Teel. Robust input-to-state stability for hybrid systems. *SIAM Journal of Control and Optimization*, 51:1651–1678, 2013.

C. Cai, A. R. Teel, and R. Goebel. Smooth lyapunov functions for hybrid systems - part I: Existence is equivalent to robustness. *IEEE Transactions on Automatic Control*, 52(7):1264–1277, 2007.

C. Cai, A. Teel, and R. Goebel. Smooth Lyapunov functions for hybrid systems part II: (Pre)asymptotic stable compact sets. *IEEE Transactions on Automatic Control*, 53:734–748, 2008.

S. L. Campbell and C. D. Meyer. *Generalized Inverses of Linear Transformations*. SIAM, 2009.

D. Carnevale, S. Galeani, L. Menini, and M. Sassano. Hybrid output regulation for linear systems with periodic jumps: Solvability conditions, structural implications and semi-classical solutions. *IEEE Transactions on Automatic control*, 61(9), 2016.

Z. Chen and J. Huang. Robust output regulation with nonlinear exosystems. *Automatica*, 41:1447–1454, 2005.

O. Christensen. *Frames and Bases. An Introductory Course.* Birkhäuser, 2008.

A. Cohen, I. Daubechies, and J.-C. Feauveau. Biorthogonal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics*, 45(5): 485–560, 1992.

N. Cox, L. Marconi, and A. R. Teel. Results on non-linear hybrid output regulation. In *52nd IEEE Conference on Decision and Control*, pages 2036–2041, Florence, Italy, 2013.

I. Daubechies. Orthonormal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics*, 41(7):909–996, 1988.

I. Daubechies. *Ten Lectures on Wavelets*. SIAM, 1992.

I. Daubechies and J. C. Lagarias. Two scales difference equations. i. existence and global regularity of solutions. *SIAM Journal on Mathematical Analysis*, 22(5): 1388–1410, 1991.

I. Daubechies and J. C. Lagarias. Two scales difference equations. ii. local regularity, infinite product of matrices and fractals. *SIAM Journal on Mathematical Analysis*, 23(4):1031–1079, 1992.

E. J. Davison. The robust control of a servomechanism problem for linear time-invariant multivariable systems. *IEEE Transactions on Automatic Control*, AC-21(1):25–34, 1976.

F. Delli Priscoli, L. Marconi, and A. Isidori. A new approach to adaptive nonlinear regulation. *SIAM Journal on Control and Optimization*, 45(3):829–855, 2006.

F. Delli Priscoli, A. Isidori, and L. Marconi. Output regulation with redundant measurements. *SICE Journal of Control, Measurement, and System Integration*, 1 (2):92–101, 2008.

F. Deutsch. *Best Approximation in Inner Product Spaces*. Springer, 2001.

Z. Ding. Global stabilization and disturbance suppression of a class of nonlinear systems with uncertain internal model. *Automatica*, 39:471–479, 2003.

M. Farza, M. M'Saad, T. Maatoug, and M. Kamoun. Adaptive observers for nonlinearly parametrized class of nonlinear systems. *Automatica*, 45:2292–2299, 2009.

H. Federer. *Geometric Measure Theory*. Springer-Verlag, New York, 1969.

A. A. Feldbaum. Dual control theory. *Autom. Remote Control*, 21(9):874–880, 1960.

F. Forte, L. Marconi, and A. Isidori. Robust design of internal models by nonlinear regression. In *Proceedings of the 9th IFAC Symposium on Nonlinear Control Systems*, pages 283–288, Toulouse, France, 2013.

F. Forte, L. Marconi, and A. R. Teel. Robust nonlinear regulation: Continuous-time internal models and hybrid identifiers. *IEEE Transactions on Automatic Control*, 62(7):3136–3151, 2017.

B. A. Francis. The linear multivariable regulator problem. *SIAM Journal of Control and Optimization*, 15(3):486–505, 1977.

B. A. Francis and W. M. Wonham. The internal model principle for multivariable regulators. *Applied Mathematics & Optimization*, 2(2):170–194, 1975.

B. A. Francis and W. M. Wonham. The internal model principle of control theory. *Automatica*, 12:457–465, 1976.

C. D. Frith, S. J. Blakemore, and D. M. Wolpert. Abnormalities in the awareness and control of actions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 355, 2000.

J.-P. Gauthier and I. Kupka. *Deterministic Observation Theory and Applications*. Cambridge University Press, 2001.

M. Gevers. Towards a joint design of identification and control? In H. Trentelman and J. Willems, editors, *Essays on Control. Progress in Systems and Control Theory*, volume 14. Birkhäuser, Boston, MA, 1993.

M. Gevers. Identification for control. *Annual Reviews in Control*, 20:95–106, 1996.

M. Gevers. Identification for control: From the early achievements to the revival of experiment design. *European Journal of Control*, 11:1–18, 2005.

R. Goebel and A. R. Teel. Solutions to hybrid inclusions via set and graphical convergence with stability theory applications. *Automatica*, 42:573–587, 2006.

R. Goebel, R. G. Sanfelice, and A. R. Teel. *Hybrid Dynamical Systems*. Princeton University Press, Princeton, N. J., 2012.

G. Gripenberg. Computing the joint spectral radius. *Linear Algebra and its Applications*, 234:43–60, 1996.

R. Grush. The emulation theory of representation: motor control, imaginery, and perception. *Behavioural and Brain Sciences*, 27(3), 2004.

H. Hammouri. Uniform observability and observer synthesis. In G. Besançon, editor, *Nonlinear Observers and Application*. Springer, 2007.

J. P. Hespanha and A. S. Morse. Stability of switched systems with average dwell-time. In *1999 IEEE 38th Annual Conference on Decision and Control*, pages 2655–2660, 1999.

J. P. Hespanha, D. Liberzon, and A. R. Teel. On input-to-state stability of impulsive systems. In *2005 IEEE 44th Conference on Decision and Control*, pages 3992–3997, 2005.

M. W. Hirsch. *Differential Topology*. Springer-Verlag, 1994.

M. R. Hirschorn. Invertibility of multivariable nonlinear control systems. *IEEE Transactions on Automatic Control*, 24(6):855–865, 1979.

H. Hjalmarsson. From experiment design to closed-loop control. *Automatica*, 41: 393–438, 2005.

J. Huang. Asymptotic tracking and disturbance rejection in uncertain nonlinear systems. *IEEE Transactions on Automatic Control*, 40(6):1118–1122, 1995.

J. Huang. Remarks on the robust output regulation problem for nonlinear systems. *IEEE Transactions on Automatic Control*, 46(12):2028–2031, 2001.

J. Huang and Z. Chen. A general framework for tackling the output regulation problem. *IEEE Transactions on Automatic Control*, 49(12):2203–2218, 2004.

J. Huang and C. F. Lin. On a robust nonlinear servomechanism problem. *IEEE Transactions on Automatic Control*, 39(7):1510–1513, 1994.

J. Huang and W. Rugh. On a nonlinear multivariable servomechanism problem. *Automatica*, 26(6):963–972, 1990.

J. Huang, A. Isidori, L. Marconi, M. Mischiati, E. D. Sontag, and W. M. Wonham. Internal models in control, biology and neuroscience. In *2018 IEEE 57th Conference on Decision and Control (CDC 2018)*, 2018.

A. Isidori. *Nonlinear Control Systems. 3rd ed.* Springer-Verlag, 1995.

A. Isidori. *Nonlinear Control Systems II*. Springer-Verlag, 1999.

A. Isidori. *Lectures in Feedback Design for Multivariable Systems*. Springer International Publishing, 2017.

A. Isidori and C. I. Byrnes. Output regulation of nonlinear systems. *IEEE Transactions on Automatic Control*, 35(2):131–140, 1990.

A. Isidori and L. Marconi. Shifting the internal model from control input to controlled output in nonlinear output regulation. In *Proceedings of the 51st IEEE Conference on Decision and Control (CDC 2012)*, pages 4900–4905, Maui, Hawaii, USA, 2012.

A. Isidori, L. Marconi, and A. Serrani. New results on semiglobal output regulation of nonminimum-phase nonlinear systems. In *41st IEEE Conference on Decision and Control*, pages 1467–1472, Las Vegas, Nevada, USA, 2002.

A. Isidori, L. Marconi, and A. Serrani. *Robust Autonomous Guidance. An internal Model Approach*. Springer-Verlag, 2003.

A. Isidori, L. Praly, and L. Marconi. About the existence of locally lipschitz output feedback stabilizers for nonlinear systems. *SIAM Journal on Control and Optimization*, 48(5):3389–3402, 2010.

A. Isidori, L. Marconi, and L. Praly. Robust design of nonlinear internal models without adaptation. *Automatica*, 48:2409–2419, 2012.

A. Isidori, L. Marconi, and G. Casadei. Robust output synchronization of a network of heterogeneous nonlinear agents via nonlinear regulation theory. *IEEE Transactions on Automatic Control*, 59(10):2680–2691, 2014.

M. Jeannerod. *Motor Cognition*. Oxford University Press, 2006.

Z.-P. Jiang, A. R. Teel, and L. Praly. Small-gain theorem for ISS systems and applications. *Mathematics of Control, Signals, and Systems*, 7:95–120, 1994.

J. Jouffroy and T. I. Fossen. A tutorial on incremental stability analysis using contraction theory. *Modeling, Identification and Control*, 31(3):99–106, 2010.

M. Kawato, T. Kuroda, H. Imamizu, E. Nakano, S. Miyauchi, and T. Yoshioka. Internal forward models in the cerebellum: fMRI study on grip force and load force coupling. In C. Prablanc, D. Pélisson, and Y. Rossetti, editors, *Progress in Brain Research, Vol. 142*, chapter 11. Elsevier Science B. V., 2003.

H. Khalil. Robust servomechanism output feedback controllers for feedback linearizable systems. *Automatica*, 30:1587–1599, 1994.

H. Khalil. On the design of robust servomechanisms for minimum phase nonlinear systems. In *IEEE 37th Conference on Decision and Control (CDC 1998)*, pages 3075–3080, Tampa, Florida, USA, Dec. 1998.

H. Khalil and L. Praly. High-gain observers in nonlinear feedback control. *International Journal of Robust and Nonlinear Control*, 24(6):993–1015, 2013.

H. K. Khalil. *Nonlinear Systems. Third Edition*. Prentice Hall, Upper Saddle River, New Jersey 07458, 2002.

D. Liberzon. Output-input stability implies feedback stabilization. *Systems & Control Letters*, 53(3):237–248, 2004.

D. Liberzon, A. S. Morse, and E. D. Sontag. Output-input stabiity and minimum-phase nonlinear systems. *IEEE Transactions on Automatic Control*, 47(3):422–436, 2002.

D. Liberzon, D. Nešić, and A. R. Teel. Lyapunov-based small-gain theorems for hybrid systems. *IEEE Transactions on Automatic Control*, 59:1395–1410, June 2014.

L. Ljung. *System Identification: Theory for the user*. Prentice Hall, 1999.

L. Ljung and T. Söderström. *Theory and Practice of Recursive Identification*. MIT Press, 1985.

W. Lohmiller and J.-J. E. Slotine. On contraction analysis for non-linear systems. *Automatica*, 34(6):683–696, 1998.

M. Lu and J. Huang. A class of nonlinear internal models for global robust output regulation problem. *International Journal of Robust Nonlinear Control*, 25:1831–1843, 2015.

L. Marconi and A. Isidori. Mixed internal model-based and feedforward control for robust tracking in nonlinear systems. *Automatica*, 36:993–1000, 2000.

L. Marconi and L. Praly. Uniform practical nonlinear output regulation. *IEEE Transactions on Automatic Control*, 53:1184–1202, 2008a.

L. Marconi and L. Praly. *Essential and Redundant Internal Models in Nonlinear Output Regulation*, pages 263–283. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008b.

L. Marconi and A. R. Teel. Internal model principle for linear systems with periodic state jumps. *IEEE Transactions on Automatic Control*, 58(11), 2013.

L. Marconi, L. Praly, and A. Isidori. Output stabilization via nonlinear Luenberger observers. *SIAM Journal of Control and Optimization*, 45:2277–2298, 2007.

R. Marino and G. L. Santosuosso. Regulation of linear systems with unknown exosystems of uncertain order. *IEEE Transactions on Automatic Control*, 52(2): 352–359, 2007.

R. Marino and P. Tomei. Global adaptive observers for nonlinear systems via filtered transformations. *IEEE Transactions on Automatic Control*, 37(8):1239–1245, 1992.

R. Marino and P. Tomei. Adaptive observers with arbitrary exponential rate of convergence for nonlinear systems. *IEEE Transactions on Automatic Control*, 40 (7):1300–1304, 1995.

R. Marino and P. Tomei. Output regulation for linear systems via adaptive internal model. *IEEE Transactions on Automatic Control*, 40(12):2199–2202, 2003.

R. Marino, G. L. Santosuosso, and P. Tomei. Robust adaptive observers for nonlinear systems with bounded disturbances. *IEEE Transactions on Automatic Control*, 46(6):967–972, 2001.

N. K. McGregor, C. I. Byrnes, and A. Isidori. Results on nonlinear output regulation for MIMO systems. In *2006 American Control Conference (ACC 2006)*, 2006.

R. C. Miall, D. J. Weir, D. M. Wolpert, and J. F. Stein. Is the cerebellum a Smith predictor? *Journal of Motor Behavior*, 25:203–216, 1993.

R. C. Miall, H. Imamizu, and S. Miyauchi. Activation of the cerebellum in coordinated eye and hand tracking movements. *Experimental Brain Research*, 135 (1):22–33, 2000.

I. Muzimoto and Z. Iwai. Adaptive output regulation of unknown MIMO linear systems with unknown exosystems. In *European Control Conference 2007*, Kos, Greece, 2007.

V. O. Nikiforov. Adaptive non-linear tracking with complete compensation of unknown disturbances. *European Journal of Control*, 4:132–139, 1998.

Q. Petit, M. Bin, and L. Marconi. Necessary conditions for output regulation with exosystem modelled by differential inclusions. IFAC Second Conference on Modelling, Identification and Control of Nonlinear Systems 2018, 2018.

A. Pyrkin and A. Isidori. Output regulation for robustly minimum-phase multivariable nonlinear systems. In *2017 IEEE 56th Conference on Decision and Control*, pages 873–878, 2017.

R. I. Schubotz. Prediction of external events with our motor system: towards a new framework. *Trends in Cognitive Science*, 11(5), 2007.

A. Serrani and A. Isidori. Global robust output regulation for a class of nonlinear systems. *Systems & Control Letters*, 39:133–139, 2000.

A. Serrani, A. Isidori, and L. Marconi. Semiglobal nonlinear output regulation with adaptive internal model. *IEEE Transactions on Automatic Control*, 46(8): 1178–1194, 2001.

S. N. Singh. A modified algorithm for invertibility of nonlinear systems. *IEEE Transactions on Automatic Control*, 26(2):595–598, 1981.

J. Sjöberg, Q. Zhang, L. Ljung, A. Benveniste, B. Delyon, P.-Y. Glorennec, H. Hjalmarsson, and A. Juditsky. Nonlinear black-box modeling in system identification: a unified overview. *Automatica*, 31(12):1691–1724, 1995.

T. Söderström. Errors-in-variables methods in system identification. *Automatica*, 43:939–958, 2007.

E. Sontag. On the input-to-state stability property. *European Journal of Control*, 1 (1):24–36, 1995.

E. D. Sontag. Smooth stabilization implies coprime factorization. *IEEE Transactions on Automatic Control*, 34(4):435–443, 1989.

E. D. Sontag. Adaptation and regulation with signal detection implies internal model. *Systems & Control Letters*, 50(2):119–126, 2003.

J. C. Sprott. *Elegant Chaos. Algebraically Simple Chaotic Flows*. World Scientific Publishing Co. Pte. Ltd., 2010.

G. Strang and T. Nguyen. *Wavelets and Filter Banks*. Wellesley-Cambridge Press, 1996.

A. Tangirala. *Principles of System Identification. Theory and Practice*. CRC Press, 2015.

A. R. Teel and L. Praly. Tools for semiglobal stabilization by partial state and output feedback. *SIAM Journal on Control and Optimization*, 33:1443–1488, 1995.

M. Twain. *Life on Mississippi river*. James R. Osgood & Co., Boston, 1883.

I. Y. Tyukin, E. Steur, H. Nijmeijer, and C. van Leeuwen. Adaptive observers and parameter estimation for a class of systems nonlinear in the parameters. *Automatica*, 49:2409–2423, 2013.

D. Walnut. *An Introduction to Wavelets*. Birkhäuser, 2002.

L. Wang, A. Isidori, and H. Su. Global stabilization of a class of invertible MIMO nonlinear systems. *IEEE Transactions on Automatic Control*, 60(3):616–631, 2015a.

L. Wang, A. Isidori, and H. Su. Output feedback stabilization of MIMO systems having uncertain high-frequency gain matrix. *Systems & Control Letters*, 83: 1–8, 2015b.

L. Wang, A. Isidori, H. Su, and L. Marconi. Nonlinear output regulation for invertible nonlinear MIMO systems. *International Journal of Robust and Nonlinear Control*, 26:2401–2417, 2016.

L. Wang, A. Isidori, Z. Liu, and H. Su. Robust output regulation for invertible nonlinear MIMO systems, 2017.

P. Wieland, R. Sepulchre, and F. Allgöwer. An internal model principle is necessary and sufficient for linear output synchronization. *Automatica*, 47:1068–1074, 2011.

D. M. Wolpert, R. C. Miall, and M. Kawato. Internal models in the cerebellum. *Trends in Cognitive Science*, 2(9), 1998.

Q. Zhang. Adaptive observer for multiple-input-multiple-output (MIMO) linear time-varying systems. *IEEE Transactions on Automatic Control*, 47(3):525–529, 2002.