OPTIMIZATION OF NONLINEAR SYSTEMS WITH UNCERTAIN OBJECTIVE FUNCTIONS VIA SLIDING-MODE CONTROL

João Carlos Espiúca Monteiro

Tese de Doutorado apresentada ao Programa de Pós-graduação em Engenharia Elétrica, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Doutor em Engenharia Elétrica.

Orientador: Alessandro Jacoud Peixoto

Rio de Janeiro
Março de 2020

OPTIMIZATION OF NONLINEAR SYSTEMS WITH UNCERTAIN OBJECTIVE FUNCTIONS VIA SLIDING-MODE CONTROL

João Carlos Espiúca Monteiro

TESE SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA (COPPE) DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE DOUTOR EM CIÊNCIAS EM ENGENHARIA ELÉTRICA.

Examinada por:

Prof. Alessandro Jacoud Peixoto, D.Sc.

Prof. Amit Bhaya, Ph.D.

Prof. Eduardo Vieira Leão Nunes, D.Sc.

Prof. Ricardo H. C. Takahashi, D.Sc.

Prof. Tiago Roux de Oliveira, D.Sc.

RIO DE JANEIRO, RJ – BRASIL
MARÇO DE 2020

Espiúca Monteiro, João Carlos

Optimization of Nonlinear Systems with Uncertain Objective Functions via Sliding-Mode Control/João Carlos Espiúca Monteiro. – Rio de Janeiro: UFRJ/COPPE, 2020.

XVI, 169 p.: il.; 29, 7cm.

Orientador: Alessandro Jacoud Peixoto

Tese (doutorado) – UFRJ/COPPE/Programa de Engenharia Elétrica, 2020.

Bibliography: p. 145 – 155.

1. real-time optimization. 2. multi-objective optimization. 3. sliding-mode control. 4. extremum-seeking control. 5. unknown control direction. 6. dynamic consensus estimation. I. Jacoud Peixoto, Alessandro. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Elétrica. III. Título.

Acknowledgements

First and foremost, I would like to thank my family, my girlfriend, and my friends. Especially my parents, Denise and Felipe, and Creuza and Fátima, who, alongside my parents, took care of me during my childhood and helped raise me. Without asking for anything in return, you all gave me the support I needed to go through this journey. I owe this achievement to you.

To my grandmother, mother of my mother, Dulcinéa, and my grandparents who are no longer here, parents of my father, Salvador and Lúcia, thanks for all the kindness and love you gave me my entire life.

To my brother and sister, Carlos Felipe and Ana Catarina, and to my dearest and oldest friend, Hugo, thank you for being there for me. I could always count on you to make things better, even in the most stressful times.

To my girlfriend, Luiza, thank you for all the support, especially in the last and most decisive year. We moved to another country, changed our lives completely, and yet you were there to keep me focused and reassure me that all the hard work would pay off. I love you dearly.

I extend my thanks to my colleagues, my advisor, and the professors from the GSCAR group. We had fruitful discussions, came up with creative and innovative ideas to tackle real-life engineering problems, and sometimes agreed to disagree on which one had the best solutions.

Last but not least, I thank CNPq for partly funding this work.

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Doutor em Ciências (D.Sc.)

OPTIMIZATION OF NONLINEAR SYSTEMS WITH UNCERTAIN OBJECTIVE FUNCTIONS VIA SLIDING-MODE CONTROL

João Carlos Espiúca Monteiro

Março/2020

Orientador: Alessandro Jacoud Peixoto

Programa: Engenharia Elétrica

Nesse trabalho, nós propomos técnicas para otimização em tempo-real, estimativa de consenso e rastreamento global de sistemas não-lineares incertos via realimentação de saída. Considerando os problemas de otimização, nós propomos três técnicas distintas: uma para otimização com apenas um objetivo e uma variável de decisão e duas para otimização multi-objetivo e com múltiplas variáveis de decisão. Todos os problemas de otimização são sujeitos à dinâmica de um processo parcialmente desconhecido. As técnicas propostas são baseadas na teoria de modos-deslizantes com funções de chaveamento periódicas e podem ser interpretadas como cascatas de duas malhas de controle. A primeira malha define a dinâmica quando em deslizamento, guiando o erro para zero. A segunda malha é responsável por garantir a ocorrência do modo-deslizante. Nós utilizamos uma abordagem inovadora ao permitir que o projetista selecione funções de chaveamento contínuas na segunda malha, além de permitir que uma classe mais abrangente de controladores possa ser utilizada na primeira malha. Considerando otimização multi-objetivo com funções objetivo convexas e teoria dos jogos, nós mostramos que uma das estratégias atinge um equilíbrio de Nash, enquanto que a outra atinge uma solução eficiente no sentido de Pareto. Para permitir o uso da última em problemas de otimização distribuída, nós desenvolvemos uma estimador dinâmico de consenso capaz de determinar a medição máxima (mínima) em uma rede, onde cada nó na rede realiza uma única medição e pode se comunicar com os seus vizinhos diretos.

No que diz respeito à dinâmica do processo, são consideradas (i) plantas estáticas, (i) afins na entrada de controle e (iii) não-lineares com um único equilíbrio exponencialmente estável. Nós mostramos que as estratégias propostas são robustas a

dinâmicas não-modeladas e, a menos do grau relativo e de algumas hipóteses não-restritivas, não dependem de um conhecimento prévio do modelo do sistema. Ademais, a direção de controle (ganho de alta frequência) é considerada desconhecida e variante com o estado.

Ao longo da Tese nós apresentamos diversos resultados numéricos obtidos através de simulações, tanto teóricas quanto inspiradas em problemas práticos de engenharia.

Abstract of Thesis presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Doctor of Science (D.Sc.)

OPTIMIZATION OF NONLINEAR SYSTEMS WITH UNCERTAIN OBJECTIVE FUNCTIONS VIA SLIDING-MODE CONTROL

João Carlos Espiúca Monteiro

March/2020

Advisor: Alessandro Jacoud Peixoto

Department: Electrical Engineering

In this work, we propose techniques for real-time optimization, consensus estimation, and global output tracking of uncertain nonlinear systems via output-feedback. Regarding the optimization problems, we propose three different techniques: one for single-objective optimization with only one decision variable, and two for multi-objective optimization with multiple decision variables. All optimization problems are subject to the dynamics of a partially unknown process. These techniques are based on the theory of sliding-modes with periodically switching functions and can be interpreted as cascades of two control loops. The first loop defines the sliding-mode dynamics, driving the error towards zero. The second is responsible for ensuring that a sliding-mode occurs. We take a novel approach by allowing the control designer to select continuous switching functions on the second loop, as well as selecting a broader class of controllers (possibly not based on sliding-modes) for the first-loop. Considering multi-objective optimization with convex objective functions and game theory, we show that one of the controllers drives the output toward a Nash equilibrium, while the other drives the output toward a Pareto efficient solution. To enable the application of the former to distributed optimization, we derive a dynamic consensus estimator capable of finding the maximum (minimum) measurement on a network, where each network node makes a single measurement and can communicate with its direct neighbors.

Regarding the process dynamics, we consider (i) static, (ii) input-affine, and (iii) exponentially stable nonlinear systems. We show that the proposed strategies are robust to unmodeled dynamics and, but for the relative degree and some mild assumptions about the unknown dynamics, do not rely on knowledge about the

system model. Furthermore, the control direction (high-frequency gain) is always considered unknown and state-varying.

Throughout the Thesis, we provide several numerical simulation examples, both theoretical and inspired by engineering practice.

Contents

List of Figures xii

List of Tables xv

List of Abbreviations xvi

1 Introduction 1
- 1.1 Objective . . . 9
- 1.2 Literature Review . . . 10
  - 1.2.1 Sliding-Mode Control . . . 10
  - 1.2.2 Extremum-Seeking Control (ESC) . . . 15
  - 1.2.3 Consensus Estimation . . . 26
- 1.3 Contributions . . . 29
  - 1.3.1 Connection with Similar Works . . . 31

2 Output-Feedback Trajectory Tracking under Unknown Control Direction 33
- 2.1 Problem Formulation . . . 34
- 2.2 Continuous Sliding-Mode Controller for Input-Affine Processes . . . 36
  - 2.2.1 Tuning Guidelines . . . 44
  - 2.2.2 Interpretation via High-Gain Proportional Integral Control . . . 45
  - 2.2.3 Application to the Speed Control of Van der Pol Oscillators . . . 47
- 2.3 Extension to Arbitrary Relative Degree . . . 50
  - 2.3.1 Application to the Speed Control of Van der Pol Oscillators with Filtered Velocity . . . 52

3 Single-Objective Real-Time Optimization 55
- 3.1 Problem Formulation . . . 56
- 3.2 Continuous Sliding-Mode Extremum-Seeking-Controller for Input-Affine Processes . . . 57
- 3.3 Stabilization of Systems with Unknown Output Sign . . . 60
  - 3.3.1 Application to Mobile Robots . . . 61

3.4 Extension to Arbitrary Relative Degree . . . 64

4 Multi-Objective Real-Time Optimization with Multiple Decision Variables 65

4.1 Problem Formulation . . . 65
4.1.1 Optimum Characterization . . . 66
4.2 SM-ESC for Nash Equilibrium Seeking in Static Noncooperative Infinite Games . . . 68
4.2.1 The Class of the Input-Output Mapping . . . 70
4.2.2 Extremum-Seeking Controller . . . 73
4.2.3 Modulation Function Design . . . 76
4.2.4 Stability and Convergence Analysis . . . 78
4.2.5 Numerical Simulation Example . . . 82
4.2.6 Application to Raman Optical Amplifiers . . . 83
4.3 Continuous SM-ESC for Nash Equilibrium Seeking in Input-Affine Processes . . . 95
4.3.1 The Class of Input-Output Mappings . . . 97
4.3.2 Extremum-Seeking Controller . . . 99
4.3.3 Modulation Function Design . . . 100
4.3.4 Stability and Convergence Analysis . . . 101

5 Real-Time Optimization with Multiple Decision Variables 105

5.1 The Class of Input-Output Mapping . . . 107
5.2 Extremum-Seeking Controller . . . 107
5.3 Modulation Function Design . . . 108
5.3.1 Matched Input Disturbance . . . 108
5.3.2 Modulation Function Design . . . 108
5.4 Stability and Convergence Analysis . . . 109
5.4.1 Existence of a Family of Integers . . . 109
5.4.2 Existence of a Family of σ-Intervals . . . 112
5.4.3 Attractiveness of Some $\mathcal{D}_{\Delta_j}$ . . . 114
5.4.4 Dominant Configuration of Signs . . . 115
5.4.5 Convergence Towards the Optimum . . . 120
5.5 Numerical Simulation Example . . . 123

6 Dynamic Estimator for Extremum Consensus 127

6.1 Problem Formulation . . . 128
6.2 Dynamic Consensus Estimator . . . 130
6.2.1 Discrete-Time Implementation Guidelines . . . 134
6.2.2 Numerical Simulation Example . . . 135

6.3 Privacy-Preserving Dynamic Consensus Estimator 139

7 Conclusion **141**

7.1 Future Works 142

A Mathematical Background **156**

A.1 Notation 156

A.2 Discontinuous Systems 157

A.2.1 Filippov's Definition of Solutions 158

A.2.2 Nonlinear Systems Affine in the Control Input 160

A.2.3 Nonuniqueness of Solutions 160

A.2.4 Lyapunov Stability and Convergence Properties 161

A.2.5 Discussion on Definitions other than Filippov's 165

B Raman Amplifier Power Dynamics and Numerical Coefficients **167**

List of Figures

1.1 (a) Phase plane portrait of a spring mass system with Coulomb friction ($\ddot{y} + 0.8\ \text{sign}(\dot{y}) + y = 0$), and (b) output of an integrator $\dot{y} = u$ with control law $u = -\text{sign}(y)$. 2
1.2 Phase portrait of the oscillator (1.2a) with control (1.2b). 3
1.3 Phase portraits of systems (a) (1.5b), (b) (1.5a), and (c) (1.4a)-(1.4b), for $\mu = 0.5$, $k = 2$, and $\alpha = 0.4$. 4
1.4 Typical framework for feedback control design. 5
1.5 Block diagram of the perturbation-based extremum-seeking control algorithm published in (Krstić and Wang 2000). 16

2.1 Two switching manifolds S_1 and S_2 dividing $\mathbb{R}^3$ into 4 parts. 33
2.2 Summary of effects of the control parameters $(\delta, T, \epsilon, \bar{\epsilon}, \lambda)$ on the auxiliary tracking error $\bar{e}(t)$, which is $\mathcal{O}(\epsilon)$ close to $e(t)$. 45
2.3 Simulation results for the SISO output tracking example with sigmoid and sign functions. 48
2.4 Comparison of control signals for different values of $\rho(t)$, respectively: $\rho_1(t)$ from (2.42), $\rho_2(t) = 40$ and $\rho_3(t) = 100$. 49
2.5 Sliding variable $\sigma(t)$ and output tracking error $e(t)$ when the HFG of (2.43) changes sign every 0.3 seconds. 50
2.6 Simulation results of the Van der Pol oscillator (2.39) controlled via output feedback ($\mu = 0.005$). To the left, controller with sigmoid functions ($\epsilon = 0.02$ and $\bar{\epsilon} = 0.05$). To the right, controller with sign functions. 53
2.7 Comparison results of the Van der Pol oscillator (2.39) controlled via state and output feedback. To the left, controller via state feedback. To the right, controller via output feedback ($\mu = 0.02$). Both controllers make use of sigmoid functions ($\epsilon = 0.1$ and $\bar{\epsilon} = 0.25$). 54

3.1 Illustration of a mobile robot carrying a laser range-finder and measuring its distance from a wall. 60
3.2 Representation of different slip coefficient impacting controllability. . 62

3.3 Simulation results for the mobile robot example (3.18) with motor dynamics (3.21). 63

4.1 Illustrative block diagram of the proposed multi-variable sliding-mode based extremum-seeking controller. 73
4.2 Simulation results for the MIMO ESC example of section 4.2.5. To the right, the implementation with discontinuous control action. To the left, the implementation using a continuous approximation of the sign function. 84
4.3 Simulation results for the MIMO ESC example of section 4.2.5 with input and output filters (4.52) and the continuous control implementation. To the left, the nominal results without the filters. To the right, the results after adding the filters. 85
4.4 In solid black, the experimental data extracted from (Bromage et al. 2002), for a *TrueWave*® *Reduced Slope* (RS) optical fiber. In blue circles our experimental data characterized by OFS Fitel Denmark ApS for a *TrueWave*® *Reach - Low Water Peak* (RFLWP) optical fiber. The "x" markers correspond to our experimental data multiplied by a factor of 1.2, showing that the *TrueWave*® (RS) and the *TrueWave*® (RFLWP) data differ only by a scale factor. 89
4.5 IO nominal maps $y_1 = h_1(x)$ and $y_2 = h_2(x)$ and the maximum values (in black) corresponding to the set of maximizers Θ_1^* and Θ_2^*, respectively. 90
4.6 Power spectrum of all 34 signals transmitted in the optical fiber (top): power in dBm and wavelength in nm. Zoomed-in views are displayed in the bottom figures: 2 pumps control signals (bottom left), 16 data signals belonging to group 1 (bottom center) and 16 data signals belonging to group 2 (bottom right). Initial power in pink (circle), optimal power in blue (star), and the results of applying maximum pump power in red (square). The black line corresponds to the desired power level (-10 dBm). 92
4.7 Time behavior of the power of all 32 data signals in dBm: 16 data signals belonging to group 1 (top) and 16 data signals belonging to group 2 (bottom). 93
4.8 Time behavior of the pump signals (in blue). Signal $x_1(t)$ (top) and signal $x_2(t)$ (bottom). All signals are expressed in dBm. The black line shows the maximizers values, and the dashed lines are the corresponding *Delta*-neighborhoods. 94

4.9 Time behavior of the objective functions corresponding to the static maps. Top: plant output y_1 and reference trajectory y_{m1}. Bottom: plant output y_2 and reference trajectory y_{m2}. 95
4.10 Phase portrait starting from the pair $(25, 27)$ dBm and converging to the pair $(27.6, 22.5)$ dBm, when the nominal plant is considered (blue); converging subsequently to the pairs $(25, 21.6)$ and $(23.8, 21.2)$ when the perturbed plants are considered (black and red lines). The Δ-regions for each case are the shaded areas. 96
4.11 Robust downstream signal power regulation in the presence of upstream signal power fluctuations (from -10 dBm to -3 dBm) and changes to the reference downstream power level (from -10 dBm to -20 dBm). 96

5.1 Two-dimensional Rosenbrock function at the top, and the systems trajectories when solving the two-dimensional optimization problem of section 5.5 using (left) the control with sign function (5.10) and (right) unit-vector control with a continuous switching element (5.89). 125
5.2 Output (objective function value) and control signals (decision variables) when solving the two-dimensional optimization problem of section 5.5 using (left) the control with sign function (5.10) and (right) unit-vector control with a continuous switching element (5.89). 126

6.1 Examples of (a) an undirected strongly connected graph, (b) a weakly connected but not strongly connected directed graph, and (c) a strongly connected directed graph. 129
6.2 Topology of the two graphs, $\mathcal{G}_a$ on the left and $\mathcal{G}_b$ on the right, considered in the simulations. For simplicity, we draw only 13 nodes of $\mathcal{G}_b$, but the actual simulation runs with 100 nodes. 135
6.3 Maximum consensus of networks $\mathcal{G}_a$ and $\mathcal{G}_b$ shown in fig. 6.2, with inputs (6.25) and (6.26). In blue the value of $\max(u)$ and in light gray the nodes inputs u_j. Only the inputs u_1, u_{33}, u_{66}, and u_{100} are shown for $\mathcal{G}_b$. 137
6.4 Simulation results for networks $\mathcal{G}_a$ and $\mathcal{G}_b$, with inputs (6.25) and (6.26), using the proposed consensus algorithm (6.5). At the top, the solution of $\mathcal{G}_a$. At the bottom, the solution of $\mathcal{G}_b$, where only 25 out of the 100 nodes are shown. 138

A.1 Geometric illustration of Filippov's definition with f discontinuous on a smooth manifold $S = \{x \in \mathbb{R}^n \mid \sigma(x) = 0\}$. 159

List of Tables

1.1 Summary of the main theoretical results on extremum-seeking control (ESC) discussed in section 1.2.2. All works that are labeled as partially applicable to nonlinear systems actually rely on a time-scale separation and strong stability and convergence assumptions, such that the process behaves as a static map. 27

1.2 Summary of theoretical results on multi-variable extremum-seeking control (ESC) listed in chronological order according to their category: perturbation-based or sliding-mode based. Columns MIMO and MISO indicate the type of optimization, while columns Scalarization, Cooperative, and Noncooperative (which do not apply to MISO optimization) further specify the type of problem. Columns Nash and Pareto (also not applicable to MISO optimization) indicate the type of equilibrium. Columns Slow, Weak Decoupling, and Chattering indicate the possible drawbacks of each technique. 28

1.3 Summary describing the applicability and characteristics of the real-time optimization techniques proposed in this Thesis. All columns have the same meanings as in tables 1.1 and 1.2, and two additional columns were added. Columns Exponentially Stable and Unstable refer to the required stability properties of the process. 30

4.1 Control variables and parameters of the sliding-mode extremum-seeking controller for static nonlinear plants, that is capable of finding a Nash equilibrium. 75

6.1 Parameters used to run the simulations on graphs $\mathcal{G}_a$ and $\mathcal{G}_b$ of fig. 6.2.136

List of Abbreviations

2-SMC	Second-order sliding-mode control, p. 11
ESC	Extremum-seeking control, p. 15
FOSMC	First-order sliding-mode control, p. 11
GAS	Globally asymptotically stable, p. 37
HFG	High-frequency gain, p. 35
HGO	High-gain observer, p. 6
HOSMC	Higher-order sliding-mode control, p. 10
HOSMO	Higher-order sliding mode observer, p. 13
MIMO	Multi-input multi-output, p. 12
PDE	Partial differential equation, p. 85
RED	Robust exact differentiator, p. 14
RFLWP	TrueWave Reach - Low Water Peak, p. 88
SISO	Single-input single-output, p. 34
SM-ESC	Sliding-mode based extremum-seeking control, p. 21
SMC	Sliding-mode control, p. 1
UCO	Uniformly completely observable, p. 6
UIOSS	Uniform input-output-to-state stable, p. 38
UVC	Unit vector control, p. 12
VSC	Variable structure control, p. 1
VSS	Variable Structure System, p. 2

Chapter 1

Introduction

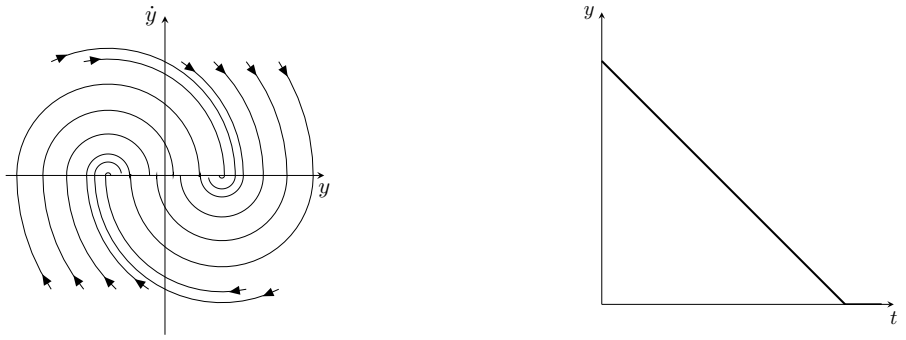
Many of nature's phenomena are mathematically described by ordinary differential equations in the form of

$$\dot{x}(t) = f(t, x) + g(t, x)u(t) \tag{1.1}$$

Equations like (1.1) have been widely used to describe the dynamics of physical systems, from the ratio between predators and preys in an ecosystem and substances in a chemical reaction to the motion of the human body. Functions f, g and u characterize the velocity vector field that establishes how the system evolves through time. This vector field can be smooth (C^∞), have up to n continuous derivatives (C^n) or even be discontinuous. Nonlinear systems with discontinuous right-hand side have several characteristics that can be exploited for control design and are the main object of interest of this Thesis.

When the right-hand side of (1.1) is discontinuous, the system motion is governed by a non-smooth vector field. A classic example is that of a mechanical system subject to Coulomb (dry) friction, which is depicted in Figure 1.1a. Even systems with smooth internal dynamics can be forced to follow non-smooth trajectories by a suitable choice of control input u. As an example, consider the integrator $\dot{y} = u$ with control law $u = -\operatorname{sign}(y)$, see Figure 1.1b.

The study of this class of systems led to the development of the mathematical foundation for variable structure control (VSC) which, in turn, motivate the techniques developed in this work. Under this framework, the theory of sliding-mode control (SMC) were and continues to be developed. In this direction, we propose SMC-based output feedback control laws for linear and nonlinear systems, with the objective of output regulation, output tracking, and real-time optimization. Before presenting our main objective, we provide a brief summary of some techniques from the control theory literature that resonate with the results in this Thesis.



(a) (b)

Figure 1.1: (a) Phase plane portrait of a spring mass system with Coulomb friction ($\ddot{y} + 0.8\ \text{sign}(\dot{y}) + y = 0$), and (b) output of an integrator $\dot{y} = u$ with control law $u = -\text{sign}(y)$.

Variable Structure Systems

As the name suggests, a system is said to be of variable structure when its dynamics varies from one structure to another. Therefore, a variable structure system (VSS) is the composition of more than one continuous subsystem, a *switching variable*, and a *switching logic*. If we revisit the aforementioned spring-mass system with Coulomb friction, the switching variable is $\dot{y}$. The switching logic states that, when the velocity is positive the friction force is negative, and vice-versa, such that it always opposes the movement of the body.

As another example, consider a linear oscillator

$$\ddot{y}(t) + y(t) = u(t) \tag{1.2a}$$

$$u(t) = \begin{cases} 0 & , \quad y > 0 \\ -\dot{y}(t) - 0.5 & , \quad y < 0 \end{cases} \tag{1.2b}$$

which is forced via the control input u, with switching surface $y = 0$. This system can be divided into two continuous subsystems

$$\ddot{y}(t) + y(t) = 0 \qquad , \ y > 0 \tag{1.3a}$$

$$\ddot{y}(t) + \dot{y}(t) + y(t) = -0.5 \ , \ y < 0 \tag{1.3b}$$

Basically, the controller is turned off when $y > 0$ and on when $y < 0$ with the objective of stabilizing the oscillator at $y = -0.5$, as illustrated in fig. 1.2. Note that even if the origin was an unstable focus of the original linear system, there would still exist a region of attraction containing trajectories on $y > 0$.

For the purpose of this work, the switching function will always be a differentiable function of the system states and (possibly) time. The terms switching variable, switching surface, and switching manifold all refer to the equation that defines the

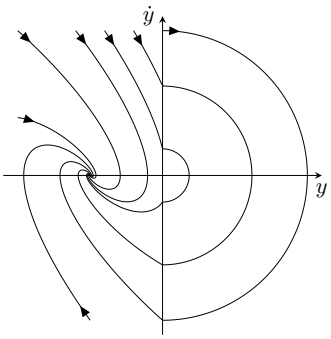


Figure 1.2: Phase portrait of the oscillator (1.2a) with control (1.2b).

subspace of the onto which the system trajectories should be confined. The terms switching function and switching logic refer to the actual control law that enforces this dynamic behavior.

We are particularly interested in a specific type VSS — sliding-mode controllers. Loosely speaking, if a VSS is such that on "opposite sides" of its switching manifold the vector field "points towards" it, then the system movement will become restricted to this manifold. When this happens, the system is said be in sliding-mode. Hence, the manifold is called a sliding manifold and the associated switching function is called sliding-mode control (SMC)[1].

A fundamental property of sliding-modes is that they describe a movement that is not predicted by any of the structures that compose the system.

The appeal of SMC is that it allows one to divide the stability analysis in two stages: (i) attractiveness of the sliding manifold and (ii) stability of trajectories inside the manifold. It is important to bear this in mind, since this two-step approach will be used throughout this Thesis. A simple example (Utkin 1977) that illustrates this concept of separability consists of a mass-spring-damper system

$$\ddot{y}(t) - \mu\,\dot{y}(t) + k\ \mathrm{sign}(\sigma)\,y(t) = 0 \tag{1.4a}$$

$$\sigma(y,\dot{y}) = \alpha\,y^2 + \dot{y}\,y \tag{1.4b}$$

with negative damping $-\mu < 0$ and sign varying spring constant $k > 0$ dependent on the switching manifold $\sigma = 0$. This particular VSS behaves like two distinct unstable linear systems[2]

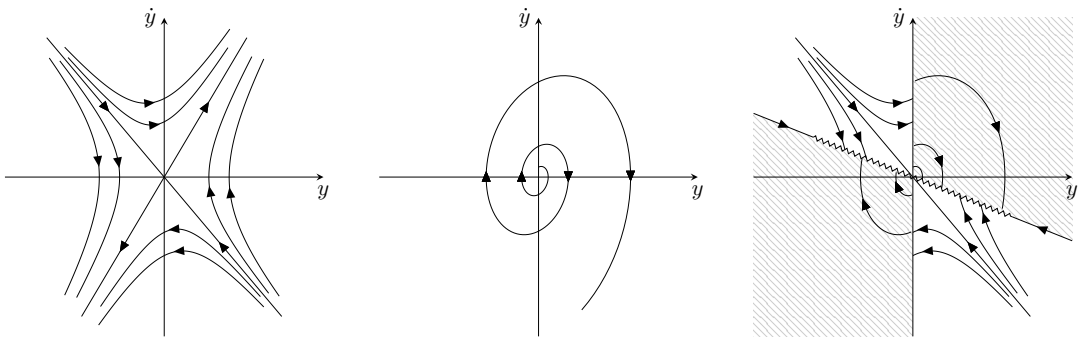
$$\ddot{y}(t) - \mu\,\dot{y}(t) + k\,y(t) = 0 \tag{1.5a}$$

$$\ddot{y}(t) - \mu\,\dot{y}(t) - k\,y(t) = 0 \tag{1.5b}$$

[1] It is, however, not always the case that the switching manifold coincides with the sliding manifold. In this Thesis, however, this holds for all control laws.

[2] When it is clear that a system possesses a single equilibrium point, we use the (not so rigorous) terminology "stable system" to refer to the stability of this point.

depending on the value of sign(σ). It is possible to show that trajectories of (1.4a)-(1.4b) tend to the line $\alpha y + \dot{y} = 0$, as indicated by Figure 1.3c. When the system reaches this manifold, its motion becomes restricted to it, therefore governed by the exponentially stable structure $\alpha y + \dot{y} = 0$. Since the system cannot leave the manifold, it is said to enter a *sliding-mode* on the manifold.



(a) (b) (c)

Figure 1.3: Phase portraits of systems (a) (1.5b), (b) (1.5a), and (c) (1.4a)-(1.4b), for $\mu = 0.5$, $k = 2$, and $\alpha = 0.4$.

This example illustrates a fundamental aspect of SMC — when in sliding-mode, a system motion is governed by a trajectory that is not inherent to any of those present on the variable structure. This new trajectory is solely defined by the sliding manifold $\sigma = 0$. Therefore, when the system is restricted to $\sigma = 0$, its stability is determined from the analysis of $\dot{\sigma} = 0$.

One may use the aforementioned aspects of SMC to design a controller for an arbitrary system. The process is outlined as follows: (i) design of the sliding manifold $\sigma = 0$ such that the origin of $\dot{\sigma} = 0$ is asymptotically stable and (ii) develop a control law that drives the system towards the sliding manifold.

Control of Nonlinear Systems

Control theory provides tools to influence and ultimately alter a process behavior. Typically, this is done in a feedback control framework similar to the one that is conceptually described in Figure 1.4. The process outputs are measured with sensors, compared to a reference, and sent to the controller, which computes the system input based on some knowledge of the process model.

Feedback controllers of this form can be divided into two main classes: output feedback and state feedback. Which strategy is chosen depends strongly on what can be measured. Full-state feedback techniques are the ones that, in general, provide wider stability margins and stronger convergence properties, but require all states to be measured for implementation. This requirement increases the amount of sensors needed to implement the control, which increases the implementation cost. Even

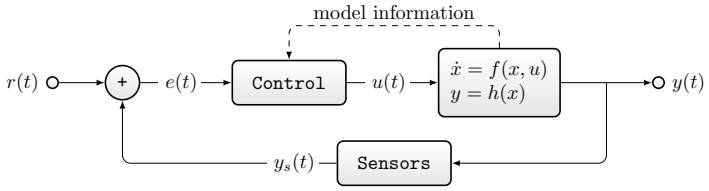


Figure 1.4: Typical framework for feedback control design.

more critical, sensors for some physical quantities are unreliable or nonexistent.

Control via Output Feedback

This work is devoted to the control and real-time optimization of input-affine nonlinear systems via output feedback. Since our results are based on sliding-mode control, when dealing with systems with unit relative degree, there is usually no need to estimate other states. On the other hand, when the system relative degree is greater than one, some of the system states, or at least some estimation of the system states, are usually required to control the output. Therefore, when it is desirable to maintain a feedback control framework, state observers are implemented to produce estimates of the unmeasured states.

Control via Observer-Based Design

The most common and direct approach to adapt a state feedback controller for output feedback is to implement a state observer to estimate the unmeasured states, and use these estimated states in place of the real ones. For linear systems, it can be shown that if both the observer and the ideal state feedback controller are asymptotically stable, so is their interconnection[3]. This property of linear systems is known as *separation principle*, or *separation property* (Chen 1999, page 255).

Although very powerful, this design pattern is not easily extended to nonlinear systems. In fact, the separation principle is (in general) not valid for nonlinear systems, even when the system is globally stabilizable by state feedback, and globally observable. Kokotovic (1992) provides a comprehensive example,

$$\dot{x}_1(t) = -x_1(t) + x_2(t)\, x_1^2(t) + u(t) \tag{1.6a}$$

$$\dot{x}_2(t) = -x_2(t) + x_1^2(t) \tag{1.6b}$$

with equilibrium globally asymptotically stable for $u = -x_2 x_1^2$. If one substitutes in the control equation the second state by the exponentially stable observer

$$\dot{\hat{x}}_2(t) = -\hat{x}_2(t) + x_1^2(t) \tag{1.7}$$

[3]Implicit in this argument is the fact that the controller and the observer must also be linear.

the equilibrium of the system (1.6a)-(1.6b) with control $u = -\hat{x}_2 x_1^2$ becomes only semi-globally asymptotically stable. Mazenc et al. (1994) provides three other examples that also illustrates this concept.

To the best of the authors knowledge, the first result towards a separation principle for nonlinear systems is due to Teel and Praly (1994), and states that: if the unmeasured states are uniformly continuously observable (UCO), global stabilizability by state feedback implies semi-global stability by output feedback. In (Teel and Praly 1995), the authors generalized the results for systems that accept some UCO state feedback controller, and proved semi-global and semi-global practical stability of such systems. These assumptions are further relaxed in (Shim and Teel 2001).

There are several observers in the control literature, each one possessing specific properties. If one seeks a linear observer, the most common ones are Luenberger and Kalman (and its variations) state estimators. Nonlinear observers are more general, accommodating a wider variety of design variations. Usual choices include high-gain observers (HGO) (Khalil and Praly 2014) and sliding-mode observers (SMO). In section 1.2.1, a large selection of sliding-mode observers is reviewed, and the reader is also referred to (Takahashi and Peres 1999) for an insightful comparison between different sliding-mode observers under a unifying approach.

For this Thesis, the possibility of relying on state observers is advantageous. That is because we will show that the output tracking and real-time optimization techniques developed in this work can make use of observers of the states norm, given that some stability and convergence properties hold, to write the control laws using only output-feedback.

Real-time Optimization

The real-time optimization problem being referred to in this Thesis is one in which

- the objective function is unknown,
- it is a function of the decision variable, and
- it can only be manipulated by appropriate choice of the control input,
- which appears on the right-hand side of the ordinary differential equation that governs the behavior of the decision variable.

It is said that the problem is solved in real-time because one must determine and track (in real-time), by choice of a suitable input, the optimum value of a performance index associated to some (partially observable and controllable) dynamical system with control. In the control literature, this problem is usually tackled using techniques from the extremum-seeking or the optimal control literatures.

Real-time Optimization and its Relation to Tracking Control

As defined in the previous section, real-time optimization can also be understood as tracking control of a system with unknown control direction, as we show below.

Let a dynamic system have state $x(t)$ and measurable output $y(t) = h \circ x(t)$, where $h(.)$ is an unknown performance index with a unique optimum $y^* = h(x^*)$. Real-time optimization consists in changing $x(t)$ online, such that y^* is found online. To illustrate this problem, consider

$$\dot{x}(t) = u(t) \tag{1.8a}$$

$$y(t) = \frac{1}{2}\left[x(t) - a\right]^2 \tag{1.8b}$$

with unknown constant $a \in \mathbb{R}$, state $x(t) \in \mathbb{R}$, input $u(t) \in \mathbb{R}$ and, output $y(t) \in \mathbb{R}$.

It is assumed that $y(t)$ is measurable, but its model is unknown. Additionally, assume that it is known that $y^* = 0$, but it is not known that $x^* = a$ nor is the state measurable. With this information, one could attempt to treat the optimization problem as a control one. For that, the output error is defined as

$$e(t) = y(t) \tag{1.9a}$$

$$\dot{e}(t) = \left[x(t) - a\right] u(t) \tag{1.9b}$$

and let $k_p(x) = x(t) - a$ denote the system high-frequency gain. Note that, not only is this gain unknown, but it changes sign as x crosses $x^* = a$. Thus, most of the traditional linear and nonlinear controllers fail to stabilize the error at $e = 0 \iff y = y^*$.

This problem inspires one to derive controllers for systems with time-varying or state-dependent high-frequency gains. If a controller is robust to changes in the system high-frequency gain sign, than it should be capable of reducing the output until it reaches its minimum. If the minimum varies with time, it suffices to show that the output is able to track some given monotonically decreasing reference $y_m(t)$, such that $y_m(t) \leq y^*(t)$ after some time $t = t_m$.

In the sliding-mode control literature, Utkin (1992, section 13.3) and Drakunov and Özgüner (1992) are among the first to consider this interpretation of real-time optimization.

Other algorithms, not only optimization algorithms, may benefit from reinterpretations based on control theory. For instance, in (Bhaya and Kaszkurewicz 2006; Bhaya and Kaszkurewicz 2007; Ferreira et al. 2005) the authors provide well-explained and self-contained descriptions of how algorithms for finding zeros, solving linear systems, and convex programming can be written as control problems.

Real-time Optimization and Optimal Control

The most significant difference between optimal control techniques and the ones developed in this Thesis is that, for optimal control, it is required to know the process model or at least some approximation of it. Using sliding-mode techniques, it is assumed that the plant dynamics follow some simple pre-defined nominal model, and mismatches between the nominal and physical models are treated as disturbances. The sliding-mode control is then designed with the nominal model in mind, but in a way that provides robustness to the unknown disturbances. On the other hand, the plant dynamics is considered as a constraint of the optimization problem in optimal control.

For example, to find the solution that minimizes the problem described by (1.9) through optimal control, the first step would be to formulate the problem as

$$\begin{array}{lll} \min_{u(t)} & y(t) & = \frac{1}{2}\left[x(t) - a\right]^2 \\ \text{s.t.} & \dot{x}(t) & = u(t)\,, \\ & x(t_0) = x_0 & \end{array} \tag{1.10}$$

If the process model is known to the designer, optimal control becomes a good choice since it allows one to explicitly include the process dynamics while also providing for the inclusion of other constraints, such as input bounds[4]. For a concise and elegant introduction to optimal control, the reader is referred to (Liberzon 2011).

Some Considerations

Before presenting the objective that guides the research developed in this thesis, we highlight some aspects of the control theory that have been briefly discussed in this introduction.

Sliding-mode control is a model-free robust control strategy. It is model-free in the sense that it does not require explicit knowledge about the process model to design the feedback controller. Therefore, a sliding-mode control law does not feature terms tailored for canceling or compensating for the process dynamics. Strategies belonging to this vast class of controllers are thus robust to model uncertainties and external disturbances without the need for parameter estimation. Ideally, SMC techniques can be considered completely insensitive to matched input disturbances and model uncertainties, as long as some (usually not restrictive) assumptions are satisfied.

Sometimes it is not possible to measure the system states needed to implement

[4]Concerning the solution method rather than the problem formulation, one thing to bear in mind is that when formulating an optimal control problem, one must solve a Hamilton-Jacobi-Bellman equation, and it must lead to a computable feedback solution(Liberzon 2011).

a state-feedback control. Thus, despite yielding weaker stability properties, output-feedback controllers have great relevance in both engineering practice and control theory.

Real-time optimization is a difficult problem, specially if little or no model information is known. However, these problems can be interpreted under the light of trajectory tracking control and the existing literature shows promising results for single-input single-output systems.

The above reasons are the three cornerstones that motivate the current Thesis. Thus, the results pursued in this manuscript concern (i) robust control of highly uncertain nonlinear systems, with unknown control direction, via sliding-modes, (ii) the extension of these controllers to tackle real-time optimization, and (iii) the generalization of these controllers to perform optimization with multiple decision variables with one or more objective functions.

1.1 Objective

The objective of this Thesis is to provide solutions for real-time multi-objective optimization. We offer solutions for multi-objective optimization seeking both Pareto efficiency and Nash equilibrium. For problems where agents in a distributed network measure the many objectives independently, we provide a solution that finds the maximum (minimum) overall objective value in the network. This overall objective is a combination of each agents individual objective function.

All optimization algorithms utilize sliding-mode based ESC, a technique that can also be used to perform output-tracking of systems with unknown control direction. Thus, one secondary goal of this Thesis is to investigate and highlight this interconnection, hopefully showing that insights from the solution to one problem might help developing solutions to the other.

When implementing the proposed sliding-mode strategies, an unwanted and well-known characteristic of SMC is revealed — chattering, high frequency switching of the control signal. Chattering can hamper control performance, so all techniques developed in this work also aim to overcome this (possible) challenge by an appropriate continuous approximation of the control signal.

Each one of the proposed control laws is followed by mathematical proofs of stability and convergence. When proving stability, our primary attempt is to find conditions for (uniform) global or semi-global asymptotic stability. Practical stability is also obtained when the output error is driven towards an ultimate bound instead of zero. In practice, effects from sampled-data implementation, sensor noise and unmodeled dynamics impact the performance of ideal controllers, such that practical stability is what can be actually achieved in practice. When proving con-

vergence, we assume the existence of a unique global extremum to formulate a proof that is coherent with the global and semi-global stabilities properties of the control algorithm. The results remain valid for locally convex objective functions, but the region of attraction then becomes an open, connected, and invariant set containing the extremum.

All stability and convergence results are developed using Lyapunov's stability theory in conjunction with Filippov and Aizerman-Pyatnitskii (Polyakov and Fridman 2014) definitions for the solution of discontinuous differential equations.

1.2 Literature Review

1.2.1 Sliding-Mode Control

The concept of sliding-mode control dates back to the 1960's and it is still an active research topic. Publications in this area can be divided into two categories, namely (1) first-order sliding-modes and (2) higher-order sliding-modes.

First-order sliding-modes techniques are the foundation of sliding-mode control. In short, given a dynamic system with states $x \in \mathbb{R}^n$, first-order sliding-mode consists of steering the system states towards a surface

$$\sigma(x) = 0 \tag{1.11}$$

such that the system trajectories are restricted to this surface. One of the first complete texts on the subject was published by Utkin (1977). In this work, Utkin discusses the fundamental design and analysis tools used in SMC. As usual in SMC, the author considers systems with relative degree equal to one, and systems affine in the control input. The concept of *equivalent control*, which is formally described in appendix A, is also introduced by the author.

As successful applications of the theory grew, other survey (Hung et al. 1993) and tutorial (DeCarlo et al. 1988; Young et al. 1999) papers were published. In the 1990's, practical implementations of SMC have reached several areas, such as: robotic manipulator control, motor control, aircraft control, and spacecraft control (Hung et al. 1993).

As an extension to the theory of SMC, Levant (1993) presented the concept of higher-order sliding-mode control (HOSMC). In HOSMC, the control signal is continuous and $r-1$ times differentiable, such that only its r-th derivative is discontinuous. With respect to the sliding surface $\sigma(x) = 0$, HOSMC drives not only

σ but also its $r-1$ derivatives to zero,

$$\sigma(x) = \dot{\sigma}(x) = \ldots = \sigma^{(r-1)}(x) = 0 \tag{1.12}$$

Since the derivatives are also driven to zero, this solution yields differentiable control laws, which ideally prevent chattering, e.g. in the absence of unmodeled dynamics, control discretization, and noise. Following the definition in (Levant 1993), techniques based only on the sign of the sliding-variable are called first-order sliding mode control (FOSMC).

A particular case of HOSMC that has received significant attention is that of second-order sliding-mode control (2-SMC). Still in (Levant 1993), the author presents the super-twisting controller, which can be understood as a nonlinear PI controller. To illustrate this controller, consider the input-affine sliding variable dynamics

$$\dot{\sigma}(t) = a(\sigma) + b(\sigma)u(t) \tag{1.13}$$

obtained for some sliding surface $\sigma(x) = 0$. The super-twisting controller for this system is given by

$$u(t) = u_1(t) + \int_0^t u_2(\tau)d\tau \tag{1.14a}$$

$$u_1(t) = \begin{cases} -\kappa_1 \left|\sigma_0\right|^{0.5} \operatorname{sign}(\sigma) & , \quad |\sigma| > \sigma_0 \\ -\kappa_1 \left|\sigma\right|^{0.5} \operatorname{sign}(\sigma) & , \quad |\sigma| \leq \sigma_0 \end{cases} \tag{1.14b}$$

$$u_2(t) = \begin{cases} -u_M & \\ -\kappa_2 \operatorname{sign}(\sigma) & , \quad |u| \leq u_M \end{cases} \tag{1.14c}$$

with positive gains $\kappa_1, \kappa_2 > 0$, sliding surface boundary layer $\sigma_0 > 0$, and control bound $u_M > 0$.

In (Venkataraman and Gulati 1993), the authors propose a technique called terminal sliding-mode control, a 2-SMC law designed to achieve finite-time convergence of the state to the origin after the trajectories have reached the sliding surface. The control law is based on terminal attractors, a concept initially discussed in the field of neural networks (Zak 1988). Considering a system with output $y(t)$ and output error $e(t) = y(t) - y_m(t)$, for some prescribed reference $y_m(t)$, the control law is defined as

$$u(t) = \lambda \left(\beta_n/\beta_d\right) e^{\beta_n/\beta_d - 1}(t)\, \dot{e}(t) + \kappa \operatorname{sign}(\sigma) \tag{1.15a}$$

$$\sigma(t) = \dot{e}(t) + \lambda\, e^{\beta_n/\beta_d}(t) \tag{1.15b}$$

with $\lambda, \kappa > 0$, $\beta_n \in (0,\, \beta_d)$ and $\beta_d = (2k+1)$, $k \in \mathbb{N}$. Note that $\beta_n/\beta_d < 1$, such that $e^{\beta_n/\beta_d - 1} \to \pm\infty$ as $e \to 0$. Thus, terminal sliding-mode assumes $\dot{e} = 0$ when $e = 0$ in order to avoid singularity. Ideally, this is true, but cannot be guaranteed when practical aspects are taken into consideration. To solve this issue, Feng et al. (2002) propose the so-called non-singular terminal sliding-mode control

$$u(t) = \lambda\,(\beta_n/\beta_d)\,\dot{e}^{2-\beta_d/\beta_n}(t) + \kappa\,\mathrm{sign}(\sigma) \tag{1.16a}$$

$$\sigma(t) = \dot{e}(t) + (1/\lambda)e^{\beta_n/\beta_d}(t) \tag{1.16b}$$

Additionally, (Yu and Zhihong 2002) add a proportional term to (1.16b) to ensure fast convergence far from the origin.

A well-known concept in SMC is that of equivalent control. When a system is in sliding-mode, by definition, the control input switches value at infinite frequency in order to keep the system constrained to the sliding manifold. Alternatively, one could compute the equivalent smooth control that would drive the system along the manifold. This equivalent smooth control input is called equivalent control. As an alternative to usual FOSMC and HOSMC, Hsu (1997) proposes the use of a low-pass filter to extract the equivalent control and compose the control law using model-reference adaptive control. Overall, the controller is composed of two elements. First, a nominal control input is computed using the model nominal parameter values, which are updated using model-reference adaptive control. To this signal, the estimated equivalent control, obtained by filtering the switching element, is added.

In the context of multi-input multi-output (MIMO) systems control, the traditional approach of FOSMC is to define n sliding variables σ_j, such that each control input is defined as $u_j = \rho_j\ \mathrm{sign}(\sigma_j)$. Alternatively, unit vector control (UVC) (Gutman and Leitmann 1975; Gutman 1979), also known as unit control, defines the control input as

$$u(t) = \rho(t,x)\frac{\sigma(x)}{\|\sigma(x)\|} \tag{1.17}$$

where $\rho > 0$ and $\sigma = \begin{bmatrix}\sigma_1 & \dots & \sigma_n\end{bmatrix}^\mathsf{T}$. An underlying property of UVC is that it produces a sliding-mode only when all sliding variables converge to zero, i.e. all $\sigma_j = 0$. Thus, there is no chattering during the reaching phase of any of the sliding variables. For some systems, this might produce better transient responses. Unit vector control has been used in the context of model-reference adaptive control (Cunha et al. 2003) and modified to consider systems that can be written in special normal form[5] (Oliveira et al. 2010a).

[5]Systems that are transformable into the special normal form have stable zero dynamics, are

An interesting application of SMC lies in the design of state and disturbance observers. In (Drakunov 1992), the concept of equivalent control (Utkin 1977; Utkin 1992) is used to develop a first-order sliding-mode observer (SMO) for minimal phase autonomous nonlinear systems. Drakunov and Utkin (1995) follow this concept to obtain an SMO for linear systems with nonzero control input. Application of this SMO for non-autonomous nonlinear systems would require the control input derivatives, which might not be acceptable if the control is non-differentiable, e.g. pulse-width modulation motor control. As a viable alternative, Barbot et al. (1996) consider nonlinear systems in triangular input observer form to propose an SMO that does not rely on input derivatives. To obtain an estimate of the equivalent control, Young et al. (1999) use a third-order low-pass butterworth filter. The authors also discuss how fast sensor dynamics impact the observers performance and show applications of SMO for disturbance estimation. In (Xiong and Saif 2001), the authors develop an SMO for a wide class of disturbed nonlinear systems.

These approaches based on first-order SMO converge to a common structure where the system is written in strict-feedback form (Khalil 2002, page 595) and states are estimated sequentially. In each estimation step, a low-pass filter is used to obtain the equivalent control. Assuming an input disturbance d added to the double integrator example, a usual first-order SMO is given by

$$\dot{\hat{x}}_1(t) = \hat{x}_2(t) + \kappa_1 \operatorname{sign}(e_1) \tag{1.18a}$$

$$\dot{\hat{x}}_2(t) = u(t) + \kappa_2 \operatorname{sign}(e_2) \tag{1.18b}$$

$$\hat{d} = \kappa_2 \operatorname{sign}(e_2)|_{\text{eq}} \tag{1.18c}$$

where $(.)_{\text{eq}}$ stands for the equivalent value of a discontinuous element, $e_1 = x_1 - \hat{x}_1$, $e_2 = x_2 - \hat{x}_2 = \kappa_1 \operatorname{sign}(e_1)|_{\text{eq}}$, and $\kappa_1, \kappa_2 > 0$. To avoid peaking, individual state observers are only turned on after the previous observations are in sliding-mode (Sussman and Kokotovic 1991). Regarding the above example, this would require setting $\dot{\hat{x}}_2 = 0$ while $e_1 > \epsilon$, for some small $\epsilon > 0$. In (Haskara and Özgüner 1999), the authors show how the filter time constants and the sampling period impact estimation accuracy.

Higher-order sliding-mode observers (HOSMO) have been successfully designed for state and disturbance estimation of dynamic systems. A natural extension of the results obtained for first-order SMO is provided in (Floquet and Barbot 2006) and consists of replacing the first-order switching functions (sign functions) by the super-twisting algorithm (Levant 1993; Fridman and Levant 2002). For this HOSMO, the same property of sequential convergence of estimated states observed in first-order SMO holds. HOSMO can be used to design exact real-time differentiators. Levant

affine in the control input and have constant high-frequency gain matrix.

(2003) rewrites the differentiation problem as an output tracking one. Via HOSMC, the output of a chain of integrators is set to track the signal to be differentiated, such that differentiation is written as an integration problem. This technique is a generalization of the so called robust exact differentiator (RED) (Levant 1998). Since the RED depends on bounds of the output derivatives, convergence properties of this technique are local. Global exact differentiators based on this concept were obtained in (Nunes et al. 2009; Oliveira et al. 2017a).

Mathematical Aspects of SMC

The importance of this area is not restricted to its practical relevance, but also due to relevant mathematical problems associated with systems governed by differential equations with discontinuous right-hand side. Lack of continuity implies that typical methods for characterizing solutions of differential equations and their stability properties cannot be applied to VSS. For instance, Lyapunov's stability definitions cannot be directly applied to such systems. Hence, a new set of mathematical tools had to be developed to study SMC.

The first rigorous study of VSS was presented in 1960 by Filippov in his celebrated theory of differential equations with discontinuous right-hand side (Filippov 1964; Filippov 1988), which is based on the concept of differential inclusions. However, Filippov's theory fails to describe some discontinuous systems nonlinear in the control input (Bartolini and Zolezzi 1985), and for this reason it was criticized by some authors (Utkin 1992; Aizerman and Pyatnitskii 1974). Naturally, this gave birth to extensions of the theory (Aizerman and Pyatnitskii 1974).

As an alternative to Filippov's theory, many authors use the concept of equivalent control (Utkin 1977; Utkin 1992) to define the solutions of VSS systems subject to sliding-modes, i.e. SMC systems. Note that both formalisms can be connected since Utkin's approach can also be written following the theory of differential inclusions (Utkin 1992).

Considering Lipschitz continuous Lyapunov function, Shevitz and Paden (1994) present stability results and invariance principles based on Filippov's differential inclusions and Clarke's generalized gradient (Clarke 1990). In (Polyakov and Fridman 2014), the authors provide a comprehensive review of mathematical tools used to establish Lyapunov stability properties of a large class of discontinuous systems. The techniques presented in their paper are applicable to piecewise continuous Lyapunov functions, which include Lipschitz continuous functions as a particular case. Furthermore, all mathematical tools are provided based on Filippov's definition. For a review of the theory, refer to appendix A.

Throughout this Thesis, since we restrict our analysis to this class of nonlinear systems, all mathematical tools were .

1.2.2 Extremum-Seeking Control (ESC)

Extremum-seeking control (ESC) is the name given for a wide variety of real-time optimization techniques used to dynamically find the optimum operating condition of an unknown output map. ESC algorithms can be either model-free or model-based and are suited for problems where only limited knowledge of the output map is available.

ESC allows one to treat the optimization problem as a control problem.[6] Thus, optimization is performed dynamically and properties such as disturbance rejection and robustness to modeling errors are easily incorporated to the optimization algorithm. There are mainly four classes of ESC strategies: *perturbation-based*, *model-based*, *sliding-mode based* and *numerical optimization-based*. Apart from numerical optimization-based ESC, these techniques are discussed separately and in chronological order in this section. We choose not to cover numerical optimization because it is substantially different than the techniques covered in this work.

Extremum-seeking control (ESC) received great attention from the control community after the work of Krstić and Wang (2000), which established the first rigorous stability proof of a perturbation-based design. Since then, many significant theoretical results on stability and performance aspects have been published by several authors. This sections highlights some of these works in order to adequately place the contributions of this Thesis later on.

As an interesting remark, most of the works on perturbation-based and model-based ESC can be placed under an unifying framework for analysis purposes, as presented in (Nešić et al. 2010b; Nešić et al. 2010a; Nešić et al. 2012)

In what follows, we try to provide an in-depth review of the most recent works on ESC, which are then summarized in tables 1.1 and 1.2. We encourage the reader to refer to these tables while following through the review. Regarding the works dealing with the multi-variable optimization, summarized in table 1.2, it is essential to note that, even though some authors do not make explicit statements specifying what type of equilibrium is reached, we decide if the solution is a Nash equilibrium or Pareto efficient based on our interpretation of the results. For all works that employ scalarization techniques, Pareto efficiency is assumed because the objective functions considered by the authors are always convex. Therefore, convergence to the scalarized problem implies convergence to a solution on the Pareto front.

[6] Another example of real-time optimization implemented in a control framework is model predictive control.

Perturbation-Based

In (Krstić and Wang 2000), the author established the first stability proof for a perturbation-based design. The results are obtained for a plant composed of a general nonlinear system connected to an input-to-output equilibrium map. This equilibrium map might contain multiple equilibria, but for each equilibrium, the map can be restricted to a region around the equilibrium and, restricted to this region, the map is convex. In other words, the equilibrium map is convex in each equilibrium. The existence of a feedback controller that makes each equilibrium locally stable is assumed. The structure of this controller is shown in Figure 1.5. Even though quite general nonlinear systems are considered, to prove stability the dynamics are assumed fast (quasi-static) with respect to the ESC algorithm. This approach is observed in almost every ESC technique, except those based on sliding-modes. It is worth noting that, prior to (Krstić and Wang 2000), most authors considered the extremum-seeking problem on plants composed by a static map. Few authors had already considered the less general problems where the plant is the cascade of: (i) a linear dynamic system and a nonlinear static map (Wiener model), (ii) a nonlinear static map and a linear dynamic system (Hammerstein model), or (iii) both at the same time (Wiener-Hammerstein model).

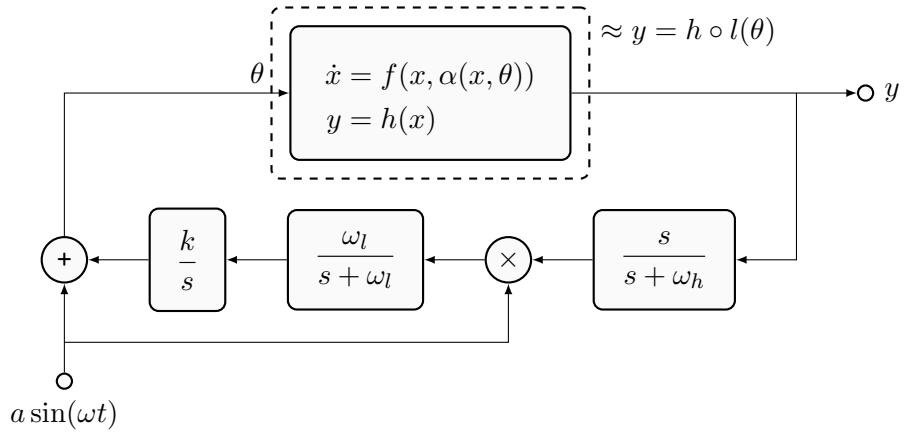


Figure 1.5: Block diagram of the perturbation-based extremum-seeking control algorithm published in (Krstić and Wang 2000).

Rotea (2000) provide a generalization for MIMO plants of the results presented by Krstić and Wang (2000). Each input channel of the MIMO dynamic system is assumed LTI and uncoupled from the other channels (diagonal). Both measurement noise and sensor dynamics are taken into account. During the same conference, Walsh (2000) presented a generalization for MISO plants following the same assumptions originally considered by Krstić and Wang (2000).

The first proof for discrete-time systems appeared in (Choi et al. 2002). The authors considered SISO plants composed of a linear dynamic system connected to

the equilibrium map and linear sensor dynamics (Wiener-Hammerstein systems).

Tunning of perturbation-based extremum-seeking controllers can be quite difficult, specially for nonlinear dynamic plants. For a simple SISO problem, there are five control parameters to be designed:

- high-pass filter cutoff frequency;
- low-pass filter cutoff frequency;
- integrator gain;
- dither (usually sinusoidal) amplitude;
- dither (usually sinusoidal) frequency.

Tan et al. (2006) elaborate on this topic to provide a clearer insight on the performance impacts and overall sensitivity with respect to these parameters. Furthermore, the authors show conditions for semi-global practical stability of the optimal operating condition, in contrast to the usual local results, and present a novel extremum-seeking controller without high-pass and low-pass filters.

Another interesting aspect highlighted in (Tan et al. 2006) is the possibility of choosing excitation (dither) signals other than sine waves, a topic that was then thoroughly studied in (Tan et al. 2008). Three performance indicators are analyzed: speed of convergence, domain of attraction, and accuracy of solutions. It is shown that there exists a trade-off between these indicators inherent to the excitation signal choice. This calls for experimenting with different dither signals when designing an extremum-seeking controller. Overall, an important result is that: for small amplitudes, the square wave provides the fastest rate of convergence. This happens because, as shown by the authors, the convergence rate is proportional to the dither power. Results from both of these works are summarized in (Nešić 2009).

In general, problems of global optimization are hard to solve. A common approach to such problems is to disturb the system when it becomes stuck at a local optimum, making it leave this valley and continue its search towards the global solution. In this direction, the authors of (Tan et al. 2009) present an ESC law for tracking a global extremum of general SISO nonlinear plants in the presence of local extrema. The algorithm is designed without high-pass or low-pass filters and starts from a large initial excitation amplitude, which converges monotonically to zero. The main underlying assumption is that, starting from a sufficiently large dither amplitude that is continuously decreased, the average system is driven towards the global extremum. Although the authors provide a formulation where the amplitude may converge to zero, one should note that this is usually not desirable. Note that, if the perturbation vanishes, then any measurement noise is free to drive the output away from its extremum value.

Perturbation-based ESC usually rely on some sort of gradient estimation inner

loop to update the input average value. In (Moase et al. 2010), the authors propose a novel Newton-like adaptation law based on the estimation of the output map first and second derivatives. By using an estimate of the map Hessian inverse, the algorithm is able to alleviate convergence problems associated with traditional perturbation-based extremum-seeking algorithms, which rely solely on gradient estimation. This adaptation is coupled with a method for varying the dither amplitude, which enables it to increase should the optimal operating condition change over time. Local stability results are obtained for Wiener-Hammerstein plants. One significant advantage of this extremum-seeking method over traditional gradient-based methods, is that its convergence rate is not proportional to the output second derivative. Thus, it is easier to tune the algorithm to achieve a desired performance. This work is significantly generalized for exponentially stable MIMO nonlinear systems in (Ghaffari et al. 2012).

A common drawback of most perturbation-based, model-based and numerical optimization-based ESC approaches, is the required time-scale separation between plant and controller dynamics. Nešić et al. (2010b) develop a unifying framework for the analysis of both perturbation-based (singular perturbation) and model-based (parameter estimation) ESC approaches, showing that a three time-scale separation, with the plant presenting the fastest dynamics and the other two enforced in the control algorithm, is required to achieve stability. Moase and Manzie (2012) are among the first to propose a technique that addresses this issue of requiring a time-scale separation. Their technique yields fast convergence to the optimal operating condition. Similar to the perturbation-based ESC law originally proposed by (Krstić and Wang 2000), which uses low and high-pass filters, Moase and Manzie (2012) apply a Luenberger observer to estimate the objective function gradient. This approach does not rely on a time-scale separation and is able to achieve semi-global stability for Wiener-Hammerstein plants.

Following the results in (Krstić and Wang 2000), the stability proof of perturbation-based extremum-seeking algorithms usually relies on averaging analysis applied to the static map, followed by singular perturbation analysis of the dynamic system. As an alternative, Dürr et al. (2013) show that trajectories of many ESC systems can be approximated by a Lie bracket dynamic system computed from the former. It is established that uniform asymptotic stability of the Lie bracket approximation implies uniform stability of the corresponding ESC system, which is the usual stability property sought for such systems. However, only static MISO plants are considered. This is generalized in (Dürr et al. 2017) for MISO nonlinear systems. This work can be viewed as an alternative framework in which to prove stability of ESC laws, as opposed to the averaging analysis approach proposed by (Krstić and Wang 2000).

Under a different design pattern, (Guay and Dochain 2014a; Guay 2016) proposes a novel extremum-seeking algorithm for minimum-phase nonlinear systems affine in the control input and with unit relative degree. The algorithm is based on proportional-integral action and does not rely on a time-scale separation between plant dynamics and control algorithm. Hence, it is able to provide fast transient response. Also, a neighborhood of the unknown optimum of an order that is inversely proportional to the dither frequency is reached. As pointed out by the authors, this results is counterintuitive in the context of perturbation-based ESC, but follows principles exploited in (Moase and Manzie 2012)[7]. A similar approach is proposed in (Guay and Dochain 2017), but written under an adaptive control framework for parameter identification. This work uses the proportional-integral approach, but an estimate of the objective function gradient is obtained from a time-varying parameter estimation routine (Guay and Dochain 2015). An advantage of using the time-varying parameter estimation routine, as stated by the authors, is that it removes the need to resort to averaging analysis to establish the convergence of the ESC algorithm to the unknown steady-state optimum.

All ESC laws discussed so far are designed for optimization of an unknown objective function. Using a Newton-based extremum-seeking algorithm, Mills and Krstić (2015) and Mills and Krstić (2018) generalize this concept for the optimization of an arbitrary derivative of the output map. In (Mills and Krstić 2015) the authors establish the results for static maps. This result is later generalized in (Mills and Krstić 2018) for a wide class of stable SISO nonlinear systems.

Considering only output delays on a MISO static non-linearity, Oliveira et al. (2015) establish local stability results for gradient-based and Newton-based ESC laws. This work is continued in Oliveira et al. (2017b), where both input and output delays (which need not be equal) are considered. The results in (Oliveira et al. 2015; Oliveira et al. 2017b) are posed for the traditional problem of optimizing an unknown objective function.

In (Rušiti et al. 2016), the authors merge the works of Mills and Krstić (2015) and Oliveira et al. (2015) to produce a Newton-based ESC law for the optimization of higher derivatives of an unknown objective function subject to output delay. An introduction to what is required to generalize these results to the stochastic case is given in (Rušiti et al. 2018), and later generalized the usual class of exponentially stable nonlinear systems considered in the ESC literature, assuming that the plant dynamics are significantly faster than the controller dynamics.

Finally, a very comprehensive overview of gradient-based and Newton-based

[7] As shown in chapter 3, the sliding-mode ESC proposed in this work also reaches a neighborhood inversely proportional to the frequency of a sinusoidal component, which plays the role of switching function instead of dither signal. Besides, it is shown in section 2.2.2 that the proposed sliding-mode ESC behaves as a high-gain proportional-integral one

extremum-seeking techniques is presented in (Krstić 2013).

Model-Based ESC

In contrast to perturbation-based design, where the objective function is assumed completely unknown, model-based design considers only parametric uncertainties. The objective function structure is considered known as a function of unknown parameters and a measurable subset of the system state. Model-based methods rely on parameter adaptation to reach the optimal operating condition. An important distinction between this class of ESC algorithms and perturbation-based ones, is that the objective function value need not be available for online feedback.

The foundations for model-based ESC design were laid by Guay and Zhang (2003). In this work, the authors develop an extremum-seeking scheme for MISO nonlinear systems affine in the control input and in the unknown parameters. Through an adaptation law, the unknown parameters are estimated and their values used to steer the measurable states to their optimal operating condition. Not that this operating condition is known, since the objective function structure is assumed known. For example, consider the optimization of $f(x) = ax^2+bx+c$, with unknown parameters $a, b, c \in \mathbb{R}$. Even though a, b and c are unknown, one already knows a priori that the optimum is at $x^* = -b/2a$. Thus, if b and a are successfully estimated, the ESC problem is solved. For the same class of dynamic systems, Dehaan and Guay (2005) provide a generalization for constrained optimization problems.

In (Adetola and Guay 2006), the authors consider the optimization of SISO Wiener plants. The algorithm distinguishes itself from previous model-based approaches because it consists of a two-stage optimization procedure. Using state and parameter estimations, the set-point that optimizes the current estimate of the objective function is computed. This value is input to a model predictive controller which solves a finite horizon optimal control problem to determine the plant input. This two-stage approach to extremum-seeking is generalized in (Adetola and Guay 2010) for a wide class of nonlinear systems.

A common limitation of model-based ESC is the need for persistence of excitation in order to achieve parametric convergence. Convergence is required to obtain the objective function estimate, which is used to compute the optimal operating condition. This limitation is overcome in (Adetola and Guay 2007) by translating it to a sufficient signal richness condition on the control input.

Similar to perturbation-based ESC, model-based approaches need some sort of time-scale separation between the plant dynamics and the extremum-seeking algorithm, requiring the first to be faster than the latter. In an attempt to relax this requirement, Sharafi et al. (2013) and Sharafi et al. (2015) propose a fast model-based ESC for SISO Hammerstein plants, that is, with LTI output dynamics. The

algorithm uses a high-frequency dither excitation signal within a model-based framework to achieve fast convergence of the parameters estimation.

Sliding-Mode Based ESC

Sliding-mode based ESC (SM-ESC) differs from other traditional extremum-seeking approaches in that it (usually) does not rely on estimation of derivatives nor model knowledge. An appealing aspect of SM-ESC is that it inherits its robustness properties, such as disturbance rejection, from sliding-mode control. Although intimately related to perturbation and model-based extremum-seeking techniques, SM-ESC development occurred somewhat parallel to these areas.

Two of the very first works to consider extremum-seeking in the context of sliding-modes are (Utkin 1992, section 13.3) and (Drakunov and Özgüner 1992). In (Drakunov and Özgüner 1992), the authors use a periodic switching function

$$u(t) = \rho \operatorname{sign}\left(\sin\left[\frac{\pi\sigma(t)}{T}\right]\right) \tag{1.19a}$$

$$\sigma(t) = y(t) + \lambda \int_0^t [y(\tau) - \bar{y}]d\tau \tag{1.19b}$$

with $\bar{y}$ greater than the maximum output value, to track the maximum output value of a process with unit relative degree. This concept was used in (Drakunov et al. 1995) to maximize friction in anti-lock braking systems. This technique distinguishes itself from other ESC designs in that it does not rely on a time-scale separation between plant dynamics and extremum-seeking algorithm. The results are valid for input-affine nonlinear systems with unit relative degree from input to the performance index.

Similar to perturbation-based ESC, the technique mentioned above can be applied to a wider class of exponentially stable nonlinear systems by forcing a time-scale separation in the control system. This is studied in (Haskara et al. 2000; Yu and Özgüner 2002). In (Haskara et al. 2000), the authors propose a two time-scale separation approach by adding an integrator to the control law proposed by Drakunov and Özgüner (1992), such that

$$u(t) = k_I \int_0^t \rho \operatorname{sign}\left(\sin\left[\frac{\pi\sigma(\tau)}{T}\right]\right) \tag{1.20}$$

with small integral gain $k_I > 0$ that guarantees that the control dynamics are slower than the plant dynamics. The authors carry out the stability results for this control strategy through Lyapunov stability methods, without resorting to singular perturbation tools, since the control signal is continuous. Following this design pattern, Yu and Özgüner (2002) show that a similar control strategy, with sliding

surface $\sigma(.)$ written as a function of the states, can be applied to systems with small time-delays, given an appropriate choice of control parameters. An estimation of the objective function gradient is also obtained from filtering of the sliding variable $\sigma(t)$. The natural disadvantage of this control law is the need for state-feedback. Design and parameter tuning guidelines for this type of SM-ESC are given in (Pan et al. 2003).

Similar to (Haskara et al. 2000), Pan et al. (2003) use the two time-scale approach to propose a smooth SM-ESC based on second-order sliding-modes. Stability results are obtained for the same class of systems considered in the former, i.e. exponentially stable nonlinear systems.

From a different perspective, Fu and Özgüner (2009) and Fu and Özgüner (2011) propose an extremum-seeking algorithm based on gradient estimation via sliding-modes. Assuming the objective function depends solely on the plant output and that both of these values are available for measurement, a discrete-time sliding variable is defined to obtain an estimate of the objective function derivative with respect to the system output. Stability results follow for SISO nonlinear systems affine in the control input that can be cast into normal form (Khalil 2002, Theorem 13.1).

Based on the tracking controller of Oliveira et al. (2010b) for strongly nonlinear systems with unknown control direction, Oliveira et al. (2011) and Oliveira et al. (2012) propose an extremum-seeking controller for SISO input-affine nonlinear systems with unit relative degree. The controller is based on that of (Drakunov 1992), such that

$$u(t) = \rho(t)\,\text{sign}\left(\sin\left[\frac{\pi\sigma(t)}{T}\right]\right) \tag{1.21a}$$

$$\sigma(t) = y(t) + \lambda\int_0^t \text{sign}[y(\tau) - \bar{y}]d\tau \tag{1.21b}$$

Note that, in this case, the gain ρ is a function of time. The authors show that, by varying $\rho(t)$ based on an estimate of the states norm (which must be observable), this control law is able to achieve global asymptotic stability with respect to the optimal operating condition. These results are generalized in (Peixoto and Oliveira 2014; Peixoto and Oliveira 2016; Lizarralde et al. 2017; Lara-Cisneros et al. 2017) for system with uniform arbitrary relative degree. In (Peixoto and Oliveira 2014; Peixoto and Oliveira 2016), the authors use high-gain observers with variable gain to obtain estimates of the system external dynamics. This estimate is used to compute the sliding variable $\sigma(t)$ in such a way that, after the trajectories converge, the dynamics from $\sigma(t)$ to $y(t)$ become of unit relative degree. Still considering arbitrary relative degree, Lizarralde et al. (2017) replace sign(.) for tanh(.) in (1.21) and stability is established through traditional singular perturbation theory, such that

the stability results hold only locally. An advantage of this method is that there is no need for additional filters when designing the control algorithm, but it does, however, require the ESC algorithm dynamics to be slower than the process dynamics. In (Lara-Cisneros et al. 2017), high-gain observers are also used to estimate the external dynamics, and this estimate used to compute the sliding variable $\sigma(t)$. However, in contrast to (Peixoto and Oliveira 2014), the modulation function $\rho(t)$ is not computed from a norm observer, but takes the form of a feedback linearizing term.

Multi-Objective Optimization, Multiple Decision Variables, and ESC

When it comes to optimization with one or more objecties and one or more decision variables, the literature is not as extensive as that for optimization problems with only one decision variable and one objective function. Thus, we have decided to dedicate a section for reviewing these works separately.

Let a process be described by the following dynamic equation

$$\dot{x}(t) = f(t, x, u) \tag{1.22a}$$

$$y(t) = h(t, x) \tag{1.22b}$$

with input $u(t) \in \mathbb{R}^{n_u}$, state $x(t) \in \mathbb{R}^n$, and output $y(t) \in \mathbb{R}^{n_y}$. When looking for an optimal behavior of such a process, one must define one or more objective functions $J_i(t, y)$ responsible for measuring the system performance. In doing so, the system dynamics become a constraint to the optimization problem, which is written (with some abuse of terminology) as

$$\min_u : J(y) = \begin{bmatrix} J_1(y) & \dots & J_{n_y}(y) \end{bmatrix}^\mathsf{T} \tag{1.23a}$$

$$\text{st} : \dot{x}(t) = f(t, x, u) \tag{1.23b}$$

$$y(t) = h(t, x) \tag{1.23c}$$

The minimum of a vector in $\mathbb{R}^{n_y}$ is not well-defined. It is used in a broader sense to indicate either Pareto efficiency or Nash equilibrium. We have also chosen to omit additional equality or inequality constraints. Even though these constraints are usually inherent in optimization problems, constraints other than the process dynamics are not covered in this Thesis.

Depending on whether or not optimization is performed assuming access to all decision variables and/or one single objective, there are three large families of optimization problems that can be described by (1.23),

- multi-objective — where one player has access to all decision variables and modifies all objective functions,

- team theory — where multiple cooperative players have access to all decision variables and modify a single objective function,
- game theory — where multiple noncooperative players have access to partitions of the decision variables and modify a partition of (including their own) objective functions.

For short, to make the notation easier, we use multiple inputs multiple outputs (MIMO) optimization to refer to problems with multiple decision variables and multiple objectives, and we use multiple inputs single output (MISO) to refer to problems with multiple decision variables and only one objective, both regardless of the number of players.

When applying game theory to (1.23), it is assumed that each agent (player) in a game is responsible for one control input (action) u_i, which modifies its decision variable y_i through the x-dynamics (1.23b) and, consequently, changes its corresponding objective function J_i value. Interaction between agents can be either *noncooperative*, when there is no information exchange between agents, such as by communicating the value of their objective functions, or *cooperative*, when each agent obtains information from at least one of the other agents.

Scalarization (Marler and Arora 2004; Marler and Arora 2010) consists of casting the multi-objective, single-player problem into an (ideally) equivalent single-objective problem. The simplest choices of scalarization are (i) weighted sums (Marler and Arora 2010), where the scalar objective is obtained by a convex combination of the objectives, and (ii) global criterion (Marler and Arora 2004), where a no-preference criterion based on the vector of objective functions, such as the distance from it to some desired solution, is minimized. Optimization problems with only one objective and multiple decision variables are seldom obtained in practice. However, since scalarization is an effective way of solving MIMO optimization, MISO optimization is also a relevant area of study.

When considering multi-objective optimization problems, the question of whether or not a given point is an optimum must be addressed with care. Depending on the number of players or the type of game, it is said that a point in multi-objective optimization is an optimizer when it is either in Nash equilibrium (when there are multiple players) or in a Pareto optimum set, i.e. when it is Pareto efficient (when there is only one player or multiple players in a cooperative game). Nash equilibrium (Nash 1950) is used to describe optimal solutions in cooperative and noncooperative games. It is achieved when an agent can not benefit (improve its objective J_i) from unilaterally changing its strategy. In a continuous scenario,

Nash equilibrium of a point y^* is verified if

$$\left.\frac{\partial J_i(y)}{\partial y_i}\right|_{y=y^*} = 0, \quad \forall i \in [1, n_y]$$

Alternatively, to be Pareto efficient, the solution must be such that there can be no individual improve in any objective function without deteriorating the performance of another one. For continuous and unconstrained problems, a necessary condition for Pareto efficiency of a point y^* is that matrix $\partial J/\partial y|_{y=y^*}$ must be rank deficient (Miettinen 1998; Kalyanmoy 2001). If the matrix $\partial J/\partial y$ is square, this condition is equivalent to

$$\left|\begin{array}{ccc} \partial J_1/\partial y_1 & \dots & \partial J_1/\partial y_n \\ \vdots & \ddots & \vdots \\ \partial J_n/\partial y_1 & \dots & \partial J_n/\partial y_n \end{array}\right|_{y=y^*} = 0$$

Keeping the concepts of Pareto efficiency, Nash equilibrium, and multi-objective optimization in mind, we proceed to review what we believe are the most relevant works in multi-variable ESC. Note, however, that in the field of ESC, it so happens that some authors do not explicitly state whether their solution is Pareto efficient or in Nash equilibrium. So we take the freedom to interpret the results and decide whether the solution is one or the other. Once again, we encourage the reader to follow this review together with tables 1.1 and 1.2.

Ariyur and Krstic (2002) and Ghaffari et al. (2012) propose similar techniques for MISO optimization. While Ariyur and Krstic (2002) generalize the traditional gradient estimation (Krstić and Wang 2000) approach, Ghaffari et al. (2012) improve upon the Newton-based approach (Moase et al. 2010), which relies on both gradient and Hessian estimation. In (Frihauf et al. 2011), the authors propose an ESC algorithm to achieve Nash equilibrium for classes of noncooperative games with both quadratic and non-quadratic objective functions $J_i(y)$ satisfying a diagonal dominance condition. Assuming some degree of connectivity between each individual agent in a cooperative game, Poveda and Quijano (2013) propose a technique to achieve one Pareto efficient solution using scalarization via convex combination.

Still based on perturbation methods, Guay and Dochain (2014b) propose a MIMO optimization algorithm, which converts the MIMO problem to MISO via scalarization by global criterion. The seeking algorithm is implemented by estimating the objective function gradient using a nonlinear observer, as detailed in (Guay and Dochain 2015), based on the parameter estimation routine of Dhaliwal and Guay (2012). Also based on parameter estimation, Guay et al. (2018) propose a solution for cooperative games using the so-called proportional-integral ESC tech-

nique (Guay and Dochain 2017). The proportional-integral action is used to stabilize the possibly unstable input-affine process, and parameter estimation is performed using the algorithm described in (Guay and Dochain 2015). As in (Poveda and Quijano 2013), scalarization is applied, and each agent tries to optimize the estimate of a convex combination of all objectives, which is obtained using a proportional-integral (unrelated to and not to be confused with the authors' proportional-integral ESC algorithm) consensus approach (Freeman et al. 2006).

Considering sliding-mode based extremum-seeking control (SM-ESC), optimization is usually treated as output-tracking of systems with unknown control direction, with the usual advantage of not relying on separate time-scales between the plant dynamics (fast) and the control dynamics (slow). In this context, Pan et al. (2002) and Peixoto and Oliveira (2012) are one of the first to propose new SM-ESC algorithms for MIMO optimization. In both works, the authors can achieve a Nash equilibrium in noncooperative games. Although their techniques share the advantages above of SM-ESC, only static systems with weak coupling between channels are considered, such that $|\partial J_i/\partial y_i|$ is significantly greater than $|\partial J_i/\partial y_j|$, for $j \neq i$.

Considering cooperative games, Salamah et al. (2018) can remove this weak coupling constraint and achieve a Pareto efficient solution. Once again, scalarization is performed via convex combination, and each agent estimates this common objective through the same proportional-integral consensus (Freeman et al. 2006) algorithm used by (Poveda and Quijano 2013; Guay et al. 2018). Salamah and Özgüncr (2018) propose an SM-ESC algorithm for MISO problems. In this work, optimization is performed one channel at a time, such that when a specific channel is running the control algorithm, the others are frozen at their previous set-point value.

At last, very recently, (Peixoto et al. 2020) published an extension of the earlier results in (Peixoto and Oliveira 2012). In contrast to their earlier conference paper (Peixoto and Oliveira 2012), where only the basic steps of the multi-variable design were discussed, the complete control design and rigorous convergence analysis are detailed in the recent manuscript. Moreover, the authors are able to relax the diagonal dominance condition, replacing it with a significantly less restrictive triangular dominance condition.

1.2.3 Consensus Estimation

Consensus algorithms are of particular interest for distributed team theory optimization problems, where multiple players collectively access the decision variables towards optimizing the same objective function. In fact, this is what motivated us to develop the consensus algorithm for maximum output estimation that we describe later in chapter 6. *Distributed* is used in the sense that the many players (also called

Table 1.1: Summary of the main theoretical results on extremum-seeking control (ESC) discussed in section 1.2.2. All works that are labeled as partially applicable to nonlinear systems actually rely on a time-scale separation and strong stability and convergence assumptions, such that the process behaves as a static map.

		SISO	MIMO	MISO	Wiener	Hammerst.	In. Affine	Nonlinear	In. Delay	Out. Delay
perturbation-based	Krstić and Wang (2000)	●						◒		
	Ariyur and Krstic (2002)			●	●					
	Rotea (2000)		●		●	●				
	Walsh (2000)			●				◒		
	Choi et al. (2002)	●			●	●				
	Tan et al. (2006)	●						◒		
	Tan et al. (2009)	●						◒		
	Moase et al. (2010)	●			●	●				
	Frihauf et al. (2011)		●					◒		
	Ghaffari et al. (2012)			●				◒		
	Moase and Manzie (2012)	●			●	●				
	Poveda and Quijano (2013)		●					◒		
	Guay and Dochain (2014b)		●					◒		
	Guay and Dochain (2015)			●				◒		
	Oliveira et al. (2015)			●					●	●
	Rušiti et al. (2016)	●							●	●
	Guay (2016)	●					●	●		
	Guay and Dochain (2017)			●			●	●		
	Oliveira et al. (2017b)			●					●	●
	Dürr et al. (2017)			●				◒		
	Guay et al. (2018)		●					◒		
	Mills and Krstić (2018)	●						◒		
	Rušiti et al. (2019)	●						◒		●
sliding-mode based	Drakunov and Özgüner (1992)	●					●	●		
	Haskara et al. (2000)	●						◒		
	Yu and Özgüner (2002)	●						◒	◒	
	Yu and Özgüner (2003)	●						◒		
	Fu and Özgüner (2011)	●					●	●		
	Oliveira et al. (2012)	●					●	●		
	Peixoto and Oliveira (2016)	●					●	●		
	Lara-Cisneros et al. (2017)	●					●	●		
	Lizarralde et al. (2017)	●					●	●		
	Salamah and Özgüncr (2018)			●				◒		
	Salamah et al. (2018)		●					◒		
	Peixoto et al. (2020)		●					◒		
model-based	Guay and Zhang (2003)			●			●	●		
	Dehaan and Guay (2005)			●			●	●		
	Adetola and Guay (2006)	●			●					
	Adetola and Guay (2010)			●				◒		
	Sharafi et al. (2013)	●				●				

applicable ● partially applicable ◒

Table 1.2: Summary of theoretical results on multi-variable extremum-seeking control (ESC) listed in chronological order according to their category: perturbation-based or sliding-mode based. Columns MIMO and MISO indicate the type of optimization, while columns Scalarization, Cooperative, and Noncooperative (which do not apply to MISO optimization) further specify the type of problem. Columns Nash and Pareto (also not applicable to MISO optimization) indicate the type of equilibrium. Columns Slow, Weak Decoupling, and Chattering indicate the possible drawbacks of each technique.

		MISO	MIMO	Scalarization	Cooperative	Noncooperative	Nash	Pareto	Slow	Weak Coupling	Chattering
perturbation	Ariyur and Krstic (2002)	●		-	-	-	-	-	●		
	Frihauf et al. (2011)		●			●	●		●	●	
	Ghaffari et al. (2012)	●		-	-	-	-	-	●		
	Poveda and Quijano (2013)		●	●	●			●	●		
	Guay and Dochain (2014b)		●	●				●	●		
	Guay et al. (2018)		●	●	●			●	◒		
	Ye et al. (2020)		●			●	●		●		
sliding-mode	Pan et al. (2002)		●			●	●		●	●	●
	Peixoto and Oliveira (2012)		●			●	●		-	●	●
	Salamah and Özgüncr (2018)	●		-	-	-	-	-	◒		●
	Salamah et al. (2018)		●	●	●			●	-		●
	this Thesis	●	●	●		●	●	●	◒		◒

applicable ● partially applicable ◒ not applicable -

nodes in the consensus literature) collaborate towards estimating a common performance index, but each player only has access to a subset of the measurements and estimates of the other players. The players belong to a network, and the connections between them determine the measurements and estimates that a player has access to. *Consensus* is used in the sense that each player aims at estimating the common performance index, and all the players converge to the same estimate.

Problems such as this one are extensively studied in the area of networked systems. They usually appear in two forms: static consensus and dynamic consensus. In *static consensus*, a snapshot of the nodes' inputs at a given time is used to initialize the algorithm, but changes to these inputs are ignored. In *dynamic consensus*, algorithms are designed to track the desired network performance as the nodes' inputs change through time.

Some of the pioneering works on consensus estimators are due to Spanos et al. (2005), Ren and Beard (2005), and Olfati-Saber et al. (2007). There are many works on consensus estimators dealing with average consensus. In the static case, authors

have proposed many solutions, encompassing, for example, privacy-preserving algorithms (Manitara and Hadjicostis 2013; Mo and Murray 2016), robustness to switching topologies and time-delays (Olfati-Saber and Murray 2004), and disturbance rejection (Bauso et al. 2009). Since the literature on static consensus is quite mature, one might be tempted to repeatedly apply a static algorithm over fixed periods of time. As discussed and exemplified by Kia et al. (2019), this is usually not the best approach.

In contrast, dynamic consensus algorithms are explicitly developed to deal with time-varying inputs. Although the literature is not as extensive as the one on static consensus, for specific consensus algorithms (mainly average consensus), authors have already tackled problems such as robustness to additive disturbances (Shi and Johansson 2013), privacy-preserving schemes (Kia et al. 2015), and robustness to communication delays (Moradian and Kia 2018). Very recently, Kia et al. (2019) wrote a survey paper on various applications and theoretical foundations of dynamic average consensus algorithms. For a thorough review of the state-of-the-art of consensus estimators, we strongly recommend (Kia et al. 2019) and the references therein. Also, a field of study that uses many of the techniques that come from the dynamic consensus literature is leader-follower networks of mobile agents.

Regarding maximum-value consensus, some authors have studied this problem to solve time synchronization in wireless sensor networks. In such a network, each sensor performs measurements in a given time and publishes this information to its neighbors. Thus, all logical times must be synchronized across the network. For this problem, authors have proposed appealing static consensus algorithms tackling the most common challenges of wireless sensor networks — privacy-preservation (Wang et al. 2019), security against malicious attacks (He et al. 2014a), and robustness to network delays (He et al. 2014b).

1.3 Contributions

In this section, we highlight the main contributions of the present work. These contributions appear in more than one form: (i) novel control algorithms, (ii) extension of current sliding-mode extremum-seeking control frameworks to embrace different design choices, and (iii) a mathematical foundation on which to establish the stability and convergence properties of similar control strategies yet to be developed. Concerning the novel and the extended control algorithms, we list them in table 1.3 to assist the reader in comparing the results to the existing literature.

Following the main objective of this Thesis — provide solutions for real-time multi-objective optimization — we can list four main contributions.

- Sliding-mode based ESC algorithms for real-time *single-objective optimization*

Table 1.3: Summary describing the applicability and characteristics of the real-time optimization techniques proposed in this Thesis. All columns have the same meanings as in tables 1.1 and 1.2, and two additional columns were added. Columns Exponentially Stable and Unstable refer to the required stability properties of the process.

	SISO	MISO	MIMO	Cooperative	Noncooperative	Nash	Pareto	Static	Input Affine	Nonlinear	Exp. Stable	Unstable	Slow	Weak Coupling	Chattering
Chap. 3.2	●	-	-	-	-	-	-		●		●		●		
Chap. 3.4	●	-	-	-	-	-	-			◒	●	●			
Chap. 4.2			●		●	●		●			-	-	-	◒	◒
Chap. 4.3			●		●	●			●			●		◒	
Chap. 4.3			●		●	●				◒	●		●	◒	
Chap. 5		●		-	-	-	-	●			-	-	-		◒

applicable ● partially applicable ◒ not applicable -

and output-feedback trajectory tracking of input-affine nonlinear systems with unknown control direction. There are two versions:

(1) Described in sections 2.2 and 3.2, applies to input-affine nonlinear systems with uniform unit relative degree. This algorithm is an extension of usual sliding-mode based controllers for processes with unknown control direction, enabling the control designer to implement a continuous, possibly chattering-free control algorithm.

(2) Published in (Lizarralde et al. 2017) and described in sections 2.3 and 3.4, implements the previous control law on exponentially stable nonlinear systems with arbitrary relative degree, relying on a time-scale separation between the process dynamics and the controller dynamics.

- Sliding-mode based ESC algorithms for real-time, *multi-objective*, and distributed optimization problems, capable of achieving *Nash equilibrium* when all objective functions are convex. There are four versions.

(1) Published in (Peixoto et al. 2020) and described in chapter 4, applies to static maps.

(2) Described in section 4.3, extends the first controller, enabling the control designer to implement a continuous, possibly chattering-free control algorithm, for input-affine nonlinear system with uniform unit relative degree, without relying on any time-scale separation. Nonetheless, if one forces a time-scale separation between the control algorithm proposed in

this section and the plant dynamics, it follows directly from the results of section 2.3 that the algorithm proposed in section 4.3 works for a class of exponentially stable nonlinear systems with arbitrary relative degree

- Sliding-mode based ESC algorithms for real-time *multi-objective* optimization problems, capable of achieving *Pareto efficient* solutions when all objective functions are convex.

  (1) Described in chapter 5, applies to static maps and relies on the scalarization of the multiple objective functions, assuming a non-distributed scenario, to achieve a Pareto efficient solution.

- A sliding-mode based consensus estimator, published in (Monteiro and Peixoto 2020) and described in chapter 6, capable of determining the maximum (minimum) output value among all outputs measured in a network of agents that cooperate with each other. The necessary and sufficient assumption is that the graph describing the network is strongly connected.

1.3.1 Connection with Similar Works

Regarding output-tracking and real-time optimization of single-input single-output systems, our proposed sliding-mode controller is heavily based on the earlier works of Drakunov and Özgüner (1992), Haskara et al. (2000), Oliveira et al. (2010b), and Oliveira et al. (2012). There are, however, some key aspects that tell our work apart from the former. These works implement a two-layer cascade sliding-mode strategy. The second layer is responsible for locking on the unknown high-frequency gain sign, while the first implements a first-order sliding mode controller to drive the error towards zero. We, on the other hand, provide a general framework that enables the control designer to remove chattering from the second sliding-mode layer and also select several other control algorithms (other than sliding-mode controllers) to drive the error to zero in the first layer. Throughout the manuscript we highlight these differences and show how the stability proof is modified to account for these changes.

The extension of the aforementioned controllers to systems with arbitrary relative degree relates to the ones in (Peixoto and Oliveira 2014) and (Lara-Cisneros et al. 2017). Nonetheless, there are key differences between ours and the aforementioned works. As already mentioned, the difference between (Lara-Cisneros et al. 2017) and (Peixoto and Oliveira 2014) is that the former computes the modulation function based on feedback linearization, while the latter uses a norm observer. Both algorithms make use of high-gain observers to obtain an estimate of the system

external dynamics when written in normal form (Khalil 2002, page 541), which implies the existence of a global diffeomorphism (or a series of local diffeomorphisms) that transforms the system into this form. We, on the other hand, show that, under the assumption that the system admits a two time-scale separation and that its quasi-steady-state model has unit relative degree, a simpler control law may be used.

The stabilization controller for SISO systems with unknown output sign, exemplified in section 3.3, can be related to (Scheinker and Krstić 2013; Scheinker and Krstić 2016). However, our control technique is fundamentally different, since Scheinker and Krstić require the control algorithm dynamics to be slower than the plant dynamics, and is based on perturbation-based ESC. Either way, Scheinker and Krstić (2013) and Scheinker and Krstić (2016) provide more general results based on the minimization of control Lyapunov functions, which remain an open topic for future development of the SISO sliding-mode controller proposed in this Thesis.

Regarding multi-objective optimization, our work on solving noncooperative problems via extremum-seeking shares a deep relation with (Peixoto and Oliveira 2012). In fact, the start of our study on multi-objective optimization derives from this work. Nonetheless, in contrast to this earlier conference paper, where only the basic steps of the multi-variable design were discussed, the complete control design and rigorous convergence analysis are detailed in this Thesis, following our published results in (Peixoto et al. 2020). Moreover, we are also able to relax the diagonal dominance condition considered in (Peixoto and Oliveira 2012) to a triangular dominance condition, which is significantly less restrictive, as well as generalize the results for input-affine nonlinear systems. The former diagonal dominance condition was generally assumed in works about multi-agents in noncooperative games (Pan et al. 2002; Frihauf et al. 2011). Another similarity of our results to those of Pan et al. (2002) and Frihauf et al. (2011) is that our algorithm also drives the system to a Nash equilibrium. Regarding the algorithm sensitivity to its initial conditions, the method based on periodic switching functions overcomes traditional perturbation-based approaches, since global convergence to a small neighborhood of the extremum (practical stability) is demonstrated rather than local stability results, given that all objective functions are convex. We say that global convergence is achieved because any initial conditions are allowed.

Finally, regarding cooperative multi-objective optimization, we take inspiration from the recent works of (Guay et al. 2018; Salamah and Özgüncr 2018; Salamah et al. 2018). Our contributions to this area are two-fold: not only do we develop a novel sliding-mode-based extremum-seeking controller (chapter 5), but we also propose a novel dynamic consensus estimator (chapter 6) that can be used in a variety of other optimization algorithms, and even outside the optimization and control fields.

Chapter 2

Output-Feedback Trajectory Tracking under Unknown Control Direction

Sliding-mode control (SMC) consists of modifying a process behavior by inducing a movement that is restricted to a sub-space of its phase space. To achieve this behavior, SMC strategies use control laws that force a discontinuity in one of the derivatives of the process output.

Two concepts are crucial for defining SMC laws; *sliding-surfaces* (*sliding-manifolds*) and *switching-functions*. The sliding-manifold defines the sub-space of the phase space to which the system movement should be restricted, while the switching-function defines the control law, calculated as a function of the sliding-manifold, to ensure that the above restriction is satisfied.

For SISO systems, the sliding-manifold splits the phase-space in two parts, each one with its own velocity vector field as a function of the switching-function. On the other hand, when there are multiple (n_u) inputs, the n_u sliding-manifolds S_1 to S_{n_u} divide the phase-space into 2^{n_u} parts, again with individual velocity vector fields as functions of the n_u switching-functions.

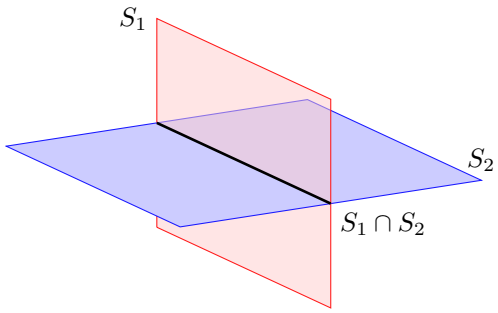


Figure 2.1: Two switching manifolds S_1 and S_2 dividing $\mathbb{R}^3$ into 4 parts.

In this chapter, we commence our study on sliding-mode control and its ap-

plications to robust output-feedback trajectory tracking and real-time model-free optimization. Regarding robust tracking, robustness is used in the sense that the control law is capable of recovering from changes in the control direction. Although Utkin (1992), Drakunov and Özgüner (1992), Haskara et al. (2000), Oliveira et al. (2010a), Oliveira et al. (2012), Aminde et al. (2013), and Lizarralde et al. (2017), among others, have provided relevant contributions to this area, we believe the results we present in this chapter further improve upon the works of these authors. Not only do we allow the control designer to choose among a broader range of controllers, but we also provide strong stability and convergence results.

The main objectives of this Thesis, discussed in section 1.1, are related to optimization, nonetheless, as indicated in the very first publications in the subject of tracking control under unknown control direction, techniques robust to these changes are usually applicable to perform real-time optimization. Therefore, the results we develop in this chapter serve as cornerstones to those detailed in chapters 3 and 4. We present solutions for trajectory tracking of uncertain nonlinear systems with unknown high-frequency gain. These solutions are based on sliding-mode control (SMC), and are inherently discontinuous[1]. Thus, there is a need to properly define what properties should their solutions satisfy. These definitions follow Filippov's theory of ordinary differential equations with discontinuous right-hand side (Filippov 1964; Filippov 1988) and its implications (Polyakov and Fridman 2014). Considering the proper mathematical tools, solutions to SMC problems and their stability proofs are presented. The reader is referred to appendix A for a review of these tools.

2.1 Problem Formulation

In this section we deal with single-input single-output (SISO) input-affine nonlinear systems that can be modeled by a finite number of coupled first-order differential equations,

$$\dot{x}(t) = f(x) + g(x)u(t) \tag{2.1a}$$

$$y(t) = h(x) \tag{2.1b}$$

where $t \in \bar{\mathbb{R}}_+$ denotes time, $x : \bar{\mathbb{R}}_+ \mapsto \mathbb{R}^n$ denotes the state variables, $\dot{x}$ their derivatives with respect to time, $f, g : \mathbb{R}^n \mapsto \mathbb{R}$ denote the state and input functions, $y : \bar{\mathbb{R}}_+ \mapsto \mathbb{R}$ denotes the system output and $h : \mathbb{R}^n \mapsto \mathbb{R}$ the output function. Discontinuities are only allowed in the control input. Therefore, functions f, g and

[1] Continuity of a system is said in terms of its velocity vector field. Note that, as defined in appendix A, solutions are always absolutely continuous.

h are at least C^0. The control input $u : \bar{\mathbb{R}}_+ \mapsto \mathbb{R}$ is piecewise continuous in its arguments.

Without loss of generality, throughout this manuscript, all signals that evolve through time are defined on $[0,\ t_f] \subset \bar{\mathbb{R}}_+$. If a solution exists $\forall t \geq 0$, then $t_f = \infty$. Furthermore, we omit the dependence of f, g, and h on time, since time-variant systems can be considered by augmenting the state vector with a fictitious state $x_{n+1}(t) = t$.

The control law presented in this section is based on the work of Oliveira et al. 2011 and was published under a less general framework in (Lizarralde et al. 2017). Before jumping into it, we elaborate a little more on what properties that systems described by (2.1) must satisfy.

Definition 1 (Relative Degree and High-Frequency Gain)**.** *System* (2.1) *has relative degree* $r \in \mathbb{N} > 0$ *at* $x' \in \mathbb{R}^n$ *iff there exists an open neighborhood* $U : x' \in U$ *such that,* $\forall x \in U$ *and* $\forall k \in \{0, 1, \ldots, r-2\}$*, the first* $r-1$ *Lie derivatives exist and they satisfy*

$$L_g L_f^k h(x) = 0 \quad \text{and} \quad L_g L_f^{r-1} h(x) \neq 0 \tag{2.2}$$

From this relation, the system high-frequency gain (HFG) is defined as

$$k_p(x) = L_g L_f^{r-1} h(x) \tag{2.3}$$

One may understand the relative degree as the number of times one must differentiate the system output with respect to time such that the control action appears. For instance, most mechanical devices are governed by second-order dynamics, such that it is only possible to directly act on the system acceleration, the second derivative of its position. Thus, if one takes the force exerted on such mechanical devices as the input and their position as the output, the system has relative degree equal to two.

Assumption 1 (Uniform Relative Degree)**.** *The relative degree* r *of system* (2.1) *is constant* $\forall x \in \mathbb{R}^n$*.*

Assumption 2 (Bounded HFG)**.** *The system high-frequency gain is continuous* $\forall x \in \mathbb{R}^n$ *and its absolute value has a lower-bound*

$$0 < \underline{k}_p \leq |k_p(x)| \tag{2.4}$$

that holds uniformly on x*.*

These assumptions are not very restrictive, since they just require the relative

degree to remain constant and the high-frequency gain to remain bounded. Note that we do not impose any previous knowledge on the HFG sign. A slightly more restrictive assumption, but common in SMC literature and required to develop the stability proofs presented in this work, is given below.

Assumption 3 (Unit Relative Degree)**.** *System* (2.1) *is assumed to have relative degree* $r = 1$*, such that*

$$k_p(x) = L_g h(x) = \left(\frac{\partial h}{\partial x}\right)^{\mathsf{T}} g(x) \tag{2.5}$$

holds for every $x \in \mathbb{R}^n$*.*

Processes with a unit relative degree are those that allow the input to influence the output velocity directly, which, for instance, is not the case for most mechanical systems. We do relax this assumption further in section 2.3, but we then require strong stability properties from (2.1).

2.2 Continuous Sliding-Mode Controller for Input-Affine Processes

Suppose system (2.1) is required to follow a known reference trajectory

$$y_m(t), \dot{y}_m(t) \in \mathbb{R} \tag{2.6}$$

with pre-specified velocity $\dot{y}_m(t)$. Define the tracking error

$$e(t) = y(t) - y_m(t) \tag{2.7a}$$
$$\dot{e}(t) = L_f h(x) + k_p(x)u(t) - \dot{y}_m(t) \tag{2.7b}$$

where u is the control input yet to be specified.

Remark 1. *If* $\operatorname{sign}(k_p)$ *is known, finite-time stabilization of the origin of system* (2.7) *is relatively simple. It suffices to choose* $u = -\rho \operatorname{sign}(k_p)\operatorname{sign}(e)$*, with sufficiently high* $\rho > 0$*, such that disturbances caused by* $L_f h(x)$ *are dominated. The proof follows from the simple Lyapunov function* $V(e) = e^2$ *and it is well-known in the SMC literature.*

Since $L_f h(x)$ and $k_p(x)$ are both unknown, stabilization of $e(t)$ is not a trivial task. The main idea here, which was first published by Drakunov et al. (1995) and later generalized by Oliveira et al. (2011), is to define a sliding variable $\sigma(t)$ as a function of $e(t)$, such that when σ enters a sliding-mode with $\dot{\sigma} = 0$ the error is led

to zero $e(t) \to 0$. For that, let

$$\sigma(t) = e(t) + \int_0^t f_e(e) d\tau \tag{2.8a}$$

$$\dot{\sigma}(t) = \dot{e}(t) + f_e(e) \tag{2.8b}$$

with $f_e : \mathbb{R} \mapsto \mathbb{R}$ restricted by the following assumption.

Assumption 4 (Stable Error Dynamics Design)**.** *Function $f_e : \mathbb{R} \mapsto \mathbb{R}$ is chosen such that*

- *the origin of $\dot{e} + f_e(e) = 0$ is globally asymptotically stable (GAS);*
- *solutions to $\dot{e} + f_e(e + \mathcal{O}(\epsilon)) = 0$ with small disturbances of order $\mathcal{O}(\epsilon)$ in f_e are globally ultimately bounded, with ultimate bound also of order $\mathcal{O}(\epsilon)$.*

In (Oliveira et al. 2011), this function is defined as $f_e(e) = \lambda \operatorname{sign}(e)$ with $\lambda > 0$, which satisfies Assumption 4. The demonstration is easily verified with $V(e) = e^2$. Another valid (and more general) choice is $f_e(e) = \lambda \operatorname{sigmoid}_\epsilon(e)$, where $\operatorname{sigmoid}_\epsilon(e)$ denotes a smooth implementation of $\operatorname{sign}(e)$. It is important to let the control designer choose $f_e(e)$, since the differential equation $\dot{e} + f_e(e) = 0$ will be the one driving the error to zero.

From Assumption 4 and (2.8b), it follows that: if σ enters a sliding regime on a constant value, its derivative goes to zero, and the error is also led to zero. Thus, $e = 0$ is at least GAS, depending on the selection of f_e. If σ enters a real sliding-mode regime, i.e. approaches the sliding surface and stays around its vicinity, then the error trajectories are globally ultimately bounded, with ultimate bound of order $\mathcal{O}(\epsilon)$.

Definition 2 (Real Sliding-Mode)**.** *A system is said to be in real sliding-mode when its trajectories can be driven, by appropriate choice of design parameters, arbitrarily close to the sliding manifold.*

Definition 3 (Matched Input Disturbance Form)**.** *A system is in matched input disturbance form when it is written as*

$$\dot{\sigma} = k_p \left(u + d\right) \tag{2.9}$$

where terms that do not belong to the control input can be grouped and written as an additive disturbance that influences the system together with the control input.

To find the control law (switching function) that makes σ enter a real sliding-mode on $\sigma = kT$, thus drives σ arbitrarily close to kT, for some integer k and

the control parameter $T > 0$, (2.7) is substituted in (2.8) and the σ-dynamics is rewritten in an equivalent matched input disturbance form

$$\dot{\sigma}(t) = k_p(x)\left[u(t) + d_\sigma(x, e, \dot{y}_m)\right] \tag{2.10a}$$

$$k_p(x)\, d_\sigma(x, e, \dot{y}_m) = \left[L_f h(x) + f_e(e) - \dot{y}_m(t)\right] \tag{2.10b}$$

where $d_\sigma(x, e, \dot{y}_m)$ is a matched input disturbance with respect to u, which grows infinitely as the HFG approaches zero $k_p(x) \to 0$. As usual in SMC theory, the control law that achieves the desired sliding-mode relies on a known disturbance d_σ bound.

Assumption 5 (Disturbance Boundedness)**.** *There exist known functions $\alpha_1 \in \mathcal{K}$ and $\beta_1 \in \mathcal{KL}$ such that $\|L_f h(x)\| \leq \alpha_1(\|x\|) + \beta_1(\|x(0)\|, t)$ and*

$$|d_\sigma| \leq \frac{1}{\underline{k}_p}\left[\alpha_1(\|x\|) + |f_e(e) - \dot{y}_m(t)| + \beta_1(\|x(0)\|, t)\right] \tag{2.11}$$

In equation (2.11), α_1 depends on the norm of the full state x. From a practical point of view, if the problem is such that α_1 depends only on the output y, then an upper-bound $\bar{d}_\sigma$ for $|d_\sigma|$ can be computed directly from (2.11). Nonetheless, in the general case, some estimation of $\|x\|$ must be obtained to compute $\bar{d}_\sigma$. Thus, consider the following assumption valid.

Assumption 6 (Norm Observability)**.** *System (2.1) is uniformly input-output-to-state stable (UIOSS) (Krichman et al. 2001, Definition 2.1) and, therefore, admits a norm observer (Krichman et al. 2001, Proposition 5.3)*

$$\|x\| \leq \bar{\alpha}_3(\|\eta\|) + \beta_2(\|x\|, t) \tag{2.12a}$$

$$\dot{\eta} = -\eta + \alpha_3(|u|) + \alpha_4(|y|), \;\; \eta(0) = 0 \tag{2.12b}$$

with known functions $\bar{\alpha}_3(s) = \bar{\alpha}_1^{-1}(2s) \in \mathcal{K}_\infty$, $\beta_2(s, t) = \bar{\alpha}_1(2\bar{\alpha}_2(s)e^{-t}) \in \mathcal{KL}$, and $\alpha_3, \alpha_4 \in \mathcal{K}$, where $\bar{\alpha}_1, \bar{\alpha}_2 \in \mathcal{K}_\infty$ are the bounds of a UIOSS-Lyapunov function (Krichman et al. 2001, Definition 2.2)

$$\bar{\alpha}_1(\|x\|) \leq V(x) \leq \bar{\alpha}_2(\|x\|) \tag{2.13}$$

of system (2.1).

This assumption might be restrictive in some cases and it can, in fact, be ignored if one does not require global stability results. When local stability results are enough, it is usual to select the maximum control effort as the norm bound for d_σ. Nonetheless, Assumption 6 allows us to rewrite (2.11) in terms of η, which depends

solely on u and y,

$$\begin{aligned}
\underline{k}_p\,|d_\sigma| &\leq \alpha_1\big(\bar{\alpha}_3\left(\|\eta\|\right) + \beta_2(\|x\|\,,t)\big) + |f_e(e) - \dot{y}_m(t)| + \beta_1(\|x(0)\|\,,t) \\
&\leq \alpha_1\left(2\bar{\alpha}_3(\|\eta\|)\right) + \alpha_1(2\beta_2(\|x\|\,,t)) + |f_e(e) - \dot{y}_m(t)| + \beta_1(\|x(0)\|\,,t) \\
&= \alpha_5\left(\|\eta\|\right) + \beta_3(\|x\|\,,t) + |f_e(e) - \dot{y}_m(t)| + \beta_1(\|x(0)\|\,,t) \\
&\leq \alpha_5\left(\|\eta\|\right) + \beta_3(2\,\|\eta\|\,,t) + \beta_3(2(\|x\| - \|\eta\|),t) + |f_e(e) - \dot{y}_m(t)| + \beta_1(\|x(0)\|\,,t) \\
&= \bar{d}_\sigma(\eta, e, \dot{y}_m) + \beta_3(2(\|x\| - \|\eta\|),t) + \beta_1(\|x(0)\|\,,t) \\
&= \bar{d}_\sigma(\eta, e, \dot{y}_m) + \beta_4(|e_\eta(0)|\,,t) + \beta_1(\|x(0)\|\,,t)
\end{aligned} \tag{2.14a}$$

$$\bar{d}_\sigma = \alpha_5\left(\|\eta\|\right) + \beta_3(2\eta,t) + |f_e(e) - \dot{y}_m(t)| \tag{2.14b}$$

where $\beta_3(x,t) = \alpha_1(2\beta_2(\|x\|\,,t))$ is known, $\alpha_5(\eta) = \alpha_1(2\bar{\alpha}_3(\|\eta\|))$, $\beta_4(|e_\eta(0)|\,,t)$, and $e_\eta = \|x\| - \|\eta\|$ is the norm observation error. Once again, we reinforce that, in practice, the upper-bound (2.14b) can be replaced by a constant upper-bound lower or equal to the maximum control effort.

With the bound provided by (2.14), the controller that produces sliding-mode $\sigma = kT$, for some integer k, is ready to be presented. This controller constitutes the second layer of the two-layer cascade sliding-mode controller for output tracking. It is important that the reader understands this first control law, because it serves as a basis (not in a mathematical sense) for all techniques that are presented in this Thesis.

Proposition 1 (Sliding-Mode Controller Design)**.** *Consider the σ-dynamics written in input disturbance form, system (2.10) with input u, matched input disturbance d_σ and unknown high-frequency gain $k_p(x)$, repeated below to ease redability*

$$\dot{\sigma}(t) = k_p(x)\left[u(t) + d_\sigma(x,e,\dot{y}_m)\right] \tag{2.15a}$$

$$k_p(x)\,d_\sigma(x,e,\dot{y}_m) = \left[L_f h(x) + f_e(e) - \dot{y}_m(t)\right] \tag{2.15b}$$

Also, consider Assumption 2, such that the high-frequency gain is bounded away from zero, and the disturbance bound $\bar{d}_\sigma$ from (2.14). Define the input

$$u(t) = \rho(\eta,e,\dot{y}_m)\,\mathrm{sigmoid}_\epsilon\left(\sin\left[\frac{\pi}{T}\sigma(t)\right]\right) \tag{2.16a}$$

$$\rho(\eta,e,\dot{y}_m) = \frac{\kappa}{\underline{k}_p}\left[\bar{d}_\sigma(\eta,e,\dot{y}_m) + \delta\right] \tag{2.16b}$$

with $T,\delta > 0$, $\kappa \geq 1$ and non-decreasing, odd function $\mathrm{sigmoid}_\epsilon : \mathbb{R} \mapsto [-1,1]$, such that $\mathrm{sigmoid}_\epsilon(s) \geq 1/\kappa$ for $s > \epsilon$, where $0 < \epsilon \ll 1$. Then, (i) no finite-time escape occurs in the system signals, x, y, σ, and u, and (ii) the σ-dynamics reach a $\mathcal{O}(\epsilon)$ real sliding-mode on the sliding manifold $\sigma = kT$ in finite time, for some integer k.

Proof. The proof is done in two steps. First, we assume that property (ii) is satisfied and that there is a time $t = t_1$ for which σ enters real sliding-mode. This is used to prove property (i). Second, property (ii) is proved. Furthermore, there are two equations in this proof that deserve a special mention, they are equations (2.18) and (2.22). The former allows us to write the proof simultaneously for continuous and discontinuous control laws, there is, if we let $\beta(\sigma) = 1$ in (2.20), then the sigmoid$_\epsilon(.)$ in (2.16a) becomes sign(.). The latter makes this proof easier to follow and to be used in conjunction to Assumption 4 to prove global tracking and real-time optimization, which we do later in Theorems 1 and 3.

Property (i) Since property (ii) is assumed valid, the sliding variable $\sigma(t)$ must be bounded, which in turn implies that the control signal $u(t)$ is bounded. Note that $\sigma(t)$ can be seen as an input to the error dynamics (2.8), which from Assumption 4 is GAS and, therefore, BIBO stable. Hence, both $e(t)$ and $y(t) = e(t) + y_m(t)$ are bounded. Finally, since $u(t)$ and $y(t)$ are bounded and the system (2.1) is UIOSS (Assumption 6), the system state $x(t)$ is also bounded.

Property (ii) The objective is to show that $\sigma(t)$ converges to kT, for some integer k, such that the distance $|\sigma(t_0) - kT|$ is minimal. Additionally, if sign$(k_p) < 0$, k must be even and if sign$(k_p) > 0$, k must be odd. With this in mind, note that the inequality

$$k\pi \leq \frac{\pi}{T}\sigma(t) \leq (k+1)\pi \ , \quad k(\sigma) = \left\lfloor \frac{\sigma(t)}{T} \right\rfloor \tag{2.17}$$

holds $\forall t \geq 0$, where k depends on σ. For any real s, $\lfloor s \rfloor$ denotes the greatest integer lower or equal to s. Moreover, $\sin(\pi\sigma/T) > 0$ for k even and $\sin(\pi\sigma/T) < 0$ for k odd. Therefore, the control input (2.16) can be written as

$$u = \beta(\sigma)\rho(-1)^k \tag{2.18}$$

for some nonnegative function $\beta(\sigma)$[2], and the following inequality is valid:

$$\beta(\sigma) \geq 1/\kappa \quad \Longleftrightarrow \quad \left|\sin\left[\frac{\pi}{T}\sigma(t)\right]\right| > \epsilon \tag{2.19}$$

Without loss of generality, consider $\epsilon \ll 1$, such that the above inequality becomes

$$\beta(\sigma) \geq 1/\kappa \quad \Longleftrightarrow \quad \left|\frac{\sigma(t)}{T} - \tilde{k}\right|\pi > \epsilon \tag{2.20}$$

where $\tilde{k} = \text{round}(\sigma(t)/T)$ denotes the rounding of $\sigma(t)/T$ to the nearest integer value. To show that an $\mathcal{O}(\epsilon)$ sliding-mode occurs on the manifold $\sigma = k^*T$, for some

[2]For sigmoid$_\epsilon(.) = $ sign$(.)$ function $\beta(\sigma) \equiv 1$ is a constant.

integer k^*, assume there is a time instant $t_0 \geq 0$ for which $\sigma(t)$ is not in sliding motion, and define

$$k_0 = \lfloor \sigma(t_0)/T \rfloor \tag{2.21a}$$

$$k^* = k_0 + \frac{\text{sign}(k_p) + (-1)^{k_0}}{2} \tag{2.21b}$$

$$s_{k^*} = \left(\frac{\sigma(t)}{T} - k^* \right) \pi \tag{2.21c}$$

where $k^* = k_0$ or $k^* = k_0 + 1$, depending on $\text{sign}(k_p)$ and the parity of k_0. Essentially, k^* must be odd when $\text{sign}(k_p) > 0$ and even when $\text{sign}(k_p) < 0$. When k_0 does not comply with these conditions, k^* is set to the following integer value $k^* = k_0 + 1$.

These results are preliminary and will be used to prove $\dot{V}(s_{k^*}) < 0$, for $|s_{k^*}| > \epsilon$, considering the candidate Lyapunov function

$$V(s_{k^*}) = \frac{T}{\pi} |s_{k^*}| \quad , \quad \dot{V}(s_{k^*}) = \frac{T}{\pi} \text{sign}(s_{k^*}) \dot{s}_{k^*} \tag{2.22}$$

which is differentiable for every $s_{k^*} \in \mathbb{R} \setminus \{0\}$, hence, usual Lyapunov analysis can be carried on, according to Corollary 3, even if $\text{sigmoid}_\epsilon(.) = \text{sign}(.)$. To this end, consider the time derivative of s_{k^*} along the trajectories of (2.10)

$$\dot{s}_{k^*} = \dot{\sigma} = k_p \left(u + d_\sigma \right) = k_p \left[\beta(\sigma) \rho (-1)^k + d_\sigma \right] \tag{2.23}$$

where the last equality is a consequence of (2.18). By using $k_p = |k_p| \, \text{sign}(k_p)$,

$$\dot{s}_{k^*} = |k_p| \left[\beta(\sigma) \rho \, (-1)^k \, \text{sign}(k_p) + d_\sigma \, \text{sign}(k_p) \right] \tag{2.24}$$

and, consequently,

$$\dot{V}(s_{k^*}) = |k_p| \left[\beta(\sigma) \rho \, (-1)^k \, \text{sign}(k_p) \, \text{sign}(s_{k^*}) + d_\sigma \, \text{sign}(k_p) \, \text{sign}(s_{k^*}) \right] \tag{2.25}$$

The first term in brackets corresponds to the control action, while the second corresponds to the system disturbances. To obtain $s_{k^*} \dot{s}_{k^*} < 0$, for $|s_{k^*}| > \epsilon$, the first term should be negative and it must dominate the second one. Therefore, we aim at showing that

$$(-1)^k \, \text{sign}(k_p) \, \text{sign}(s_{k^*}) = (-1)^k \, \text{sign}(k_p) \frac{s_{k^*}}{|s_{k^*}|} = -1 \tag{2.26}$$

independent of $\text{sign}(k_p)$ and the parity of k.

First, note that k^* is chosen such that $\sigma(t)/\epsilon \in [k^* - 1, \; k^* + 1]$, at least at $t = t_0$.

Verify that

$$k = \left\lfloor \frac{\sigma(t)}{T} \right\rfloor = k^* - 1 \implies \frac{\sigma(t)}{T} \in [k^* - 1,\ k^*] \tag{2.27a}$$

$$k = \left\lfloor \frac{\sigma(t)}{T} \right\rfloor = k^* \implies \frac{\sigma(t)}{T} \in [k^*,\ k^* + 1] \tag{2.27b}$$

and consider both cases:

- $k = k^* - 1 \implies \sigma(t)/T < k^*$

$$\begin{aligned}(-1)^k s_{k^*} &= (-1)^k \left(\frac{\sigma}{T} - k^*\right)\pi = (-1)^k(-1)\left|\frac{\sigma}{T} - k^*\right|\pi \\ &= (-1)^{k+1}\left|\frac{\sigma}{T} - k^*\right|\pi = (-1)^{k^*}\left|\frac{\sigma}{T} - k^*\right|\pi \\ (-1)^k s_{k^*} &= (-1)^{k^*}\,|s_{k^*}|\end{aligned} \tag{2.28}$$

- $k = k^* \implies \sigma(t)/T > k^*$

$$\begin{aligned}(-1)^k s_{k^*} &= (-1)^k \left(\frac{\sigma}{T} - k^*\right)\pi = (-1)^k\left|\frac{\sigma}{T} - k^*\right|\pi = (-1)^{k^*}\left|\frac{\sigma}{T} - k^*\right|\pi \\ (-1)^k s_{k^*} &= (-1)^{k^*}\,|s_{k^*}|\end{aligned} \tag{2.29}$$

Equations (2.28) and (2.29) substituted into $(-1)^k \operatorname{sign}(k_p)\operatorname{sign}(s_{k^*})$ yield

$$(-1)^k \operatorname{sign}(k_p)\operatorname{sign}(s_{k^*}) = (-1)^{k^*}\operatorname{sign}(k_p) \tag{2.30}$$

Therefore, to obtain (2.26), one must find $(-1)^{k^*}\operatorname{sign}(k_p) = -1$, which is a direct consequence of k^* definition (2.21b). In fact, as previously stated, k^* is odd for $\operatorname{sign}(k_p) > 0$ which implies $(-1)^{k^*}\operatorname{sign}(k_p) = (-1)^{k^*} = -1$. For $\operatorname{sign}(k_p) < 0$, k^* is even, implying $(-1)^{k^*}\operatorname{sign}(k_p) = (-1)^{k^*+1} = -1$. Thus, equation (2.26) holds and $\dot{V}(s_{k^*})$ in (2.25) is rewritten as

$$\begin{aligned}\dot{V}(s_{k^*}) &= |k_p|\left[-\beta(\sigma)\rho + d_\sigma \operatorname{sign}(k_p)\operatorname{sign}(s_{k^*})\right] \\ &\leq -\underline{k}_p\left[\beta(\sigma)\rho - |d_\sigma|\right] \\ &\leq -\left[\beta(\sigma)\left(\kappa\bar{d}_\sigma + \delta\right) - \bar{d}_\sigma\right] + \underline{k}_p\left[\beta_4(|e_\eta(0)|, t) + \beta_1(\|x(0)\|, t)\right]\end{aligned}$$

with the help of (2.14). Finally, since β_4 and β_1 are class $\mathcal{KL}$ functions which vanish exponentially, and taking inequality (2.20) into account,

$$\dot{V}(s_{k^*}) \leq -\delta < 0\ , \quad |s_{k^*}| > \epsilon \tag{2.31}$$

This concludes the proof, implying that real sliding-mode occurs in $s_{k^*} = 0 \iff$

$\sigma = k^* T$, with ultimate bound

$$|\sigma| \leq k^* \ T + \frac{\epsilon \, T}{\pi} \leq k^* \ T + \mathcal{O}(\epsilon) \tag{2.32}$$

■

Proposition 1 implies that the second layer of the two-layer cascade controller is capable of producing real or ideal sliding-modes on a constant manifold $\sigma = k^* T$, for some integer k^*, depending on the choice of sigmoid$_\epsilon$. Therefore, the output error dynamics, after some finite time, is governed by one of the equations in Assumption 4, and the control designer is free to choose the desired transient response by a proper selection of $f_e(e)$.

Theorem 1. *Consider system (2.1) with output error (2.7), reference model $y_m(t)$, and control law (2.8), (2.16). Then, the output tracking error origin is uniformly globally practical asymptotically stable (UGPAS), with ultimate bound $|e| \leq \mathcal{O}(\epsilon)$, and all system states remain bounded.*

Proof. The proof is a direct consequence of Proposition 1 and Assumption 4. From Proposition 1, equation (2.32) , it follows that $|\sigma(t) - k^* \ T| \leq (\epsilon/\pi)$, for $t \geq t_1$ and some integer k^*. Hence, consider the change of variables $\bar{e} = e - \sigma + k^* \ T$, such that, for $t \geq t_1$, equation (2.8) is rewritten as

$$k^* \ T = \bar{e} + \int_0^t f_e(\bar{e} + \mathcal{O}(\epsilon)) d\tau \tag{2.33a}$$

$$0 = \dot{\bar{e}} + f_e(\bar{e} + \mathcal{O}(\epsilon)) \tag{2.33b}$$

From Assumption 4, it follows that the origin of $\bar{e}$ is UGPAS, with ultimate bound $|\bar{e}| \leq \mathcal{O}(\epsilon)$. Since $|e| \leq |\bar{e}| + |\sigma - k^* \ T|$, the origin of e is also UGPAS with ultimate bound $|e| \leq \mathcal{O}(\epsilon)$. To conclude the proof, since the system is UIOSS and its input and output are bounded, the state x must also be bounded.

■

Theorem 1 concludes the general results on global output tracking under unknown control direction. Below, we work on this results to provide two useful choices for $f_e(e)$ that will be used throughout this manuscript.

Corollary 1. *Consider Theorem 1 and let $f_e(e) = \lambda \, \text{sat}_{\bar{\epsilon}}(e)$, with*

$$\text{sat}_{\bar{\epsilon}}(e) = \begin{cases} e/\bar{\epsilon} & , \quad |e| < \bar{\epsilon} \\ \text{sign}(e) & , \quad \text{otherwise} \end{cases} \tag{2.34}$$

where $\lambda > 0$ and $\bar{\epsilon} \geq 2\epsilon \, T/\pi$. Then, the output tracking error origin is UGPAS, with ultimate bound $|e| \leq 2\epsilon \, T/\pi \leq \mathcal{O}(\epsilon)$.

Proof. Consider (2.33) with $f_e(.) = \lambda \, \mathrm{sat}_{\bar{\epsilon}}(.)$ and (2.34). Using the Lyapunov function $V(\bar{e}) = \bar{e}^2$, with $\bar{e}$ from the proof of Theorem 1, and the fact that $\bar{\epsilon} > \epsilon \, T/\pi$, it follows that $\bar{e} = 0$ is UGPAS with ultimate bound $|\bar{e}| \leq \epsilon \, T/\pi$. Therefore, since $e = \bar{e} + (\sigma - k^* \, T)$, the tracking error origin $e = 0$ is UGPAS, with ultimate bound $|e| \leq \epsilon \, T/\pi + \epsilon \, T/\pi = 2\epsilon \, T/\pi \leq \mathcal{O}(\epsilon)$. ■

Corollary 2. *Consider Theorem 1 and let* $\mathrm{sigmoid}_\epsilon(.) = \mathrm{sign}(.)$ *and* $\kappa = 1$. *Then, (i) an ideal sliding-mode* $\sigma = k^*T$ *is attained in finite time, for some integer* k^*, *and (ii) if* $f_e(.) = \lambda \, \mathrm{sign}(.)$, *with* $\lambda > 0$, *the output tracking error origin is UGAS, and* $e = 0$ *is reached in finite time.*

Proof. By taking $\mathrm{sigmoid}_\epsilon(.) = \mathrm{sign}(.)$, one has $\epsilon = 0$, which, from Proposition 1, implies $\sigma = kT$ and $\dot{\sigma} = 0$ for $t \geq t_1$ and some integer k. The proof of property (i) is established by considering Assumption 4. To prove property (ii), note that $f_e(.) = \lambda \, \mathrm{sign}(.)$ and $\epsilon = 0$ imply $\dot{e} + \mathrm{sign}(e) = 0$. Hence, UGAS and finite-time convergence follow from the Lyapunov function $V(e) = |e|$. ■

Note that, if one specializes the proposed control law by selecting $\mathrm{sigmoid}_\epsilon(.) = \mathrm{sign}(.)$ and $f_e(.) = \lambda \, \mathrm{sign}(.)$, the controller is the exact same as the one in Oliveira et al. 2011; Oliveira et al. 2012.

2.2.1 Tuning Guidelines

In general, we recommend using the controller proposed in Corollary 1, since simulations considering implementation constraints, such as sampling rate and actuator dynamics, indicate that usual switching functions based on $f_e(.) = \lambda \, \mathrm{sign}(.)$ suffer from severe performance deterioration in these cases. On the other hand, continuous approximations, such as $f_e(.) = \lambda \, \mathrm{sat}_{\bar{\epsilon}}(.)$ and $\mathrm{sigmoid}_\epsilon(.) = \mathrm{sat}_\epsilon(.)$, offer smaller gains near $e = 0$, reducing the error sensitivity with respect to these constraints. Assuming $f_e(.) = \lambda \, \mathrm{sat}_{\bar{\epsilon}}(.)$ and $\mathrm{sigmoid}_\epsilon(.) = \mathrm{sat}_\epsilon(.)$, directives for choosing the controller parameters are given in this section.

Essentially, there are 5 parameters available for the designer: δ, T, ϵ, $\bar{\epsilon}$ and λ. Together, they control all the major convergence and robustness properties of the proposed controller, as discussed below and illustrated in fig. 2.2.

δ **(excess in control) and** T **(sliding manifold position)** Together they make for the maximum amount of time taken for σ to enter sliding motion, given by T/δ. Therefore, one should compromise when attempting to reduce the "open-loop transient" duration by choosing when to reduce T, which increases chattering and requires a smaller sampling period, and when to increase δ, which increases the control signal. It is desirable that T/δ be as small as possible so that the controller recovers promptly from: strong disturbances

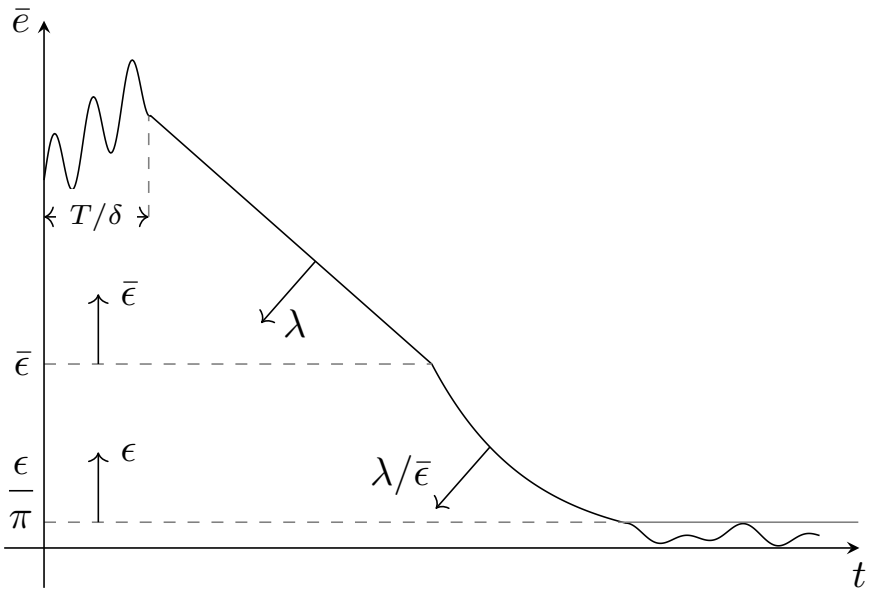


Figure 2.2: Summary of effects of the control parameters $(\delta, T, \epsilon, \bar{\epsilon}, \lambda)$ on the auxiliary tracking error $\bar{e}(t)$, which is $\mathcal{O}(\epsilon)$ close to $e(t)$.

and possible losses of controllability. One should note, however, that this estimate is usually very conservative and small transient responses can be achieved with a much softer (but not guaranteed) upper bound $0.1T/\delta$.

ϵ **(steady-state error)** Determines the maximum steady-state output tracking error, given by $|e| \leq 2\epsilon T/\pi$. It should satisfy $0 < \epsilon \ll 1$. Care should be taken when decreasing ϵ, because smaller values of ϵ usually produce chattering.

$\bar{\epsilon}$ **(chattering reduction)** Reduces the input gain seen by the noise coming from nonzero $\sigma - k^* \; T$ and defines a lower error bound, above which $\dot{e} = -\lambda$. Thus, by proper choice of $\bar{\epsilon}$ (big enough), the output becomes less sensitive to imperfections from real sliding motion, which are inevitable. Nonetheless, big values of $\bar{\epsilon}$ imply slower error decay rates when $|e| < \bar{\epsilon}$.

λ **(error decay rate)** Controls the error decay rate $\dot{e}$, which is (i) linear, with $|e(t)| = |e(t_1)| - \lambda(t - t_1)$ for $|e| \geq \bar{\epsilon}$ or (ii) exponential, with $|e(t)| \approx |e(t_1)| \exp\left[-(\lambda/\bar{\epsilon})(t - t_1)\right]$ for $|e| < \bar{\epsilon}$, where t_1 is the time at which σ enters real sliding-mode. Although λ should be big for higher decay rates, it should not be so big as to increase chattering.

2.2.2 Interpretation via High-Gain Proportional Integral Control

The proposed controller with $\text{sigmoid}_\epsilon(.) = \text{sat}_\epsilon(.)$ and $f_e(.) = \lambda \, \text{sat}_{\bar{\epsilon}}(.)$, when in real sliding-mode, reduces to a high-gain proportional integral controller (where the inte-

gral action is even higher than the proportional action). This interpretation is very enlightening and might aid one in developing and interpreting control algorithms based on different choices of $f_e(.)$. Note that, when σ enters real sliding-mode, $\sigma(t) = k^* T + \mathcal{O}(\epsilon)$, for some integer k^* and $\mathcal{O}(\epsilon) < \epsilon T/\pi$. Therefore,

$$\begin{aligned}
\sin(\pi\sigma/T) &= \sin(\pi k^* + \pi\mathcal{O}(\epsilon)/T) \\
&= \cos(\pi k^*)\sin(\pi\mathcal{O}(\epsilon)/T) \\
&= -\operatorname{sign}(k_p)\sin(\pi\mathcal{O}(\epsilon)/T) \\
\sin(\pi\sigma/T) &\approx -\operatorname{sign}(k_p)\pi\mathcal{O}(\epsilon)/T
\end{aligned}$$

and

$$u(t) = \rho(t)\,\mathrm{sat}_{\,\epsilon}\left(\sin\left[\frac{\pi}{T}\sigma(t)\right]\right) \approx -\operatorname{sign}(k_p)\frac{\pi\rho(t)}{\epsilon T}\mathcal{O}(\epsilon) \tag{2.35}$$

From equation (2.8), when σ is in real sliding-mode and $|e| < \bar{\epsilon}$ (which is true for some finite time), $k^* T + \mathcal{O}(\epsilon) = e + (\lambda/\bar{\epsilon})\int e d\tau$. Thus,

$$u(t) \approx -\operatorname{sign}(k_p)\frac{\pi\rho(t)}{\epsilon T}\left[e(t) + \frac{\lambda}{\bar{\epsilon}}\int e(\tau)d\tau - k^* T\right] \tag{2.36}$$

Finally, let $d_e = d_\sigma - f_e/k_p$, such that

$$\dot{e}(t) = k_p(x)\left[u(t) + d_e(t)\right] \tag{2.37}$$

which, for small tracking error $|e| < \bar{\epsilon}$ and σ in real sliding-mode, becomes

$$\dot{e}(t) \approx -\frac{|k_p(x)|}{\mu}\left[e(t) + \frac{\lambda}{\bar{\epsilon}}\int e(\tau)\,d\tau - k^* T + \mu\, d_e(t)\right] \tag{2.38}$$

with $\mu = \epsilon T/[\pi\rho(t)]$.

Equation (2.38) shows that the proposed controller behaves like a high-gain proportional integral controller with matched input disturbances. Since $k^* T$ is constant, it is perfectly compensated by the integral action. The other disturbance component $d_e(t)$ is significantly attenuated. In fact, from equations (2.16b), (2.14b), when σ is in real sliding-mode, $|d_e(t)| \leq |\rho(t)|$, which implies $\mu\,|d_e(t)| \leq \epsilon T/\pi$.

Hence, the proposed control law functions in two stages. First, σ is driven towards the sliding manifold $\sigma = k^* T$ and the high-frequency gain sign is obtained from $\operatorname{sign}(k_p) = -\cos(\pi k^*)$, since k^* is odd for $\operatorname{sign}(k_p) > 0$ and even otherwise. Then, a high-gain proportional integral controller drives the tracking error to zero while compensating for constant and small time-varying disturbances.

2.2.3 Application to the Speed Control of Van der Pol Oscillators

The Van der Pol oscillator model is widely used in control theory and applications. It consists of a mass-spring-damper system with nonlinear damping coefficient. In this work, we write it as

$$m\ddot{y}(t) + b(y)\dot{y}(t) + ky(t) = u(t) \tag{2.39}$$

with spring constant $k > 0$, mass $m > 0$ and damping $b(y) = b_1[(y/b_0)^2 - 1]$, $b_0, b_1 > 0$. The nonlinear damping $b(y)$ is negative for $|y| < b_0$ and positive otherwise. To write (2.39) in state-space form, let $x = [y\ \dot{y}]^\mathsf{T}$, such that

$$\dot{x}(t) = \begin{bmatrix} 0 & 1 \\ -k/m & -b(x_1)/m \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 1/m \end{bmatrix} u(t) \tag{2.40}$$

where both y and $\dot{y}$ are assumed measured and $y' = \dot{y}$ is chosen as output, where y' is just a new variable to denote the output. The assumption of measurable velocity $\dot{y}$ is relaxed in section 2.3, where only the position y is measured and a lead filter is used to estimate $\dot{y}$. For now, this assumption is required to obtain unit relative degree from input to output. To obtain the modulation function $\rho(t)$, consider the following bounds:

$$b(y) \leq \bar{b}_1 \left[(y/\bar{b}_0)^2 - 1\right] = \bar{b}(y) \tag{2.41a}$$

$$k \leq \bar{k} \tag{2.41b}$$

$$m = 1 \tag{2.41c}$$

where the mass m is assumed exactly known for simplicity. Finally, combining the previous expressions yield

$$\rho(t) = \kappa \left(\left| \bar{k}y + \bar{b}(y)\dot{y} \right| + |f_e(e) - \dot{y}'_m(t)| + \delta \right) \tag{2.42}$$

For simulation purposes, we set $m = 1$ kg, $k = 100$ N/m, $b_1 = 10$ N/(m/s) and $b_0 = 0.05$ m. To compute the lower bounds, we consider $\bar{k} = 1.25k$, $\bar{b}_1 = 1.25b_1$, and $\bar{b}_0 = 1.25b_0$, that is, a 25% safety margin on each parameter. Assume the following specifications for the velocity controller: following error below 0.01 m/s, settling time of 0.2 s, and real sliding-mode achieved in 0.05 s. Following the procedure of section 2.2.1, the controller gains are set to: $\epsilon = 0.02$, $T = 1$ m/s, $\delta = 2$ m/s^2, $\bar{\epsilon}$, $\lambda = 5$ m/s^2, and $\bar{\epsilon} = 0.05$ m/s.

Implementing the controllers in discrete-time with a sampling period of 10 μs, a

reference signal $y'_m(t) = \sin(2\pi t)$, and solving the simulation with first order Runge-Kutta with a step size of 1 μs, the closed-loop system was simulated and the results are shown in fig. 2.3. To show the results with ideal sliding-modes, a simulation using sign(.) functions was also performed. Ideally, if one reduces the simulation step size even further, the output tracking error can be driven exactly to zero. Nonetheless, it is evident that the control signal suffers from severe chattering when ideal sliding-mode controllers are considered, while absolutely no chattering is observed when using the saturation function.

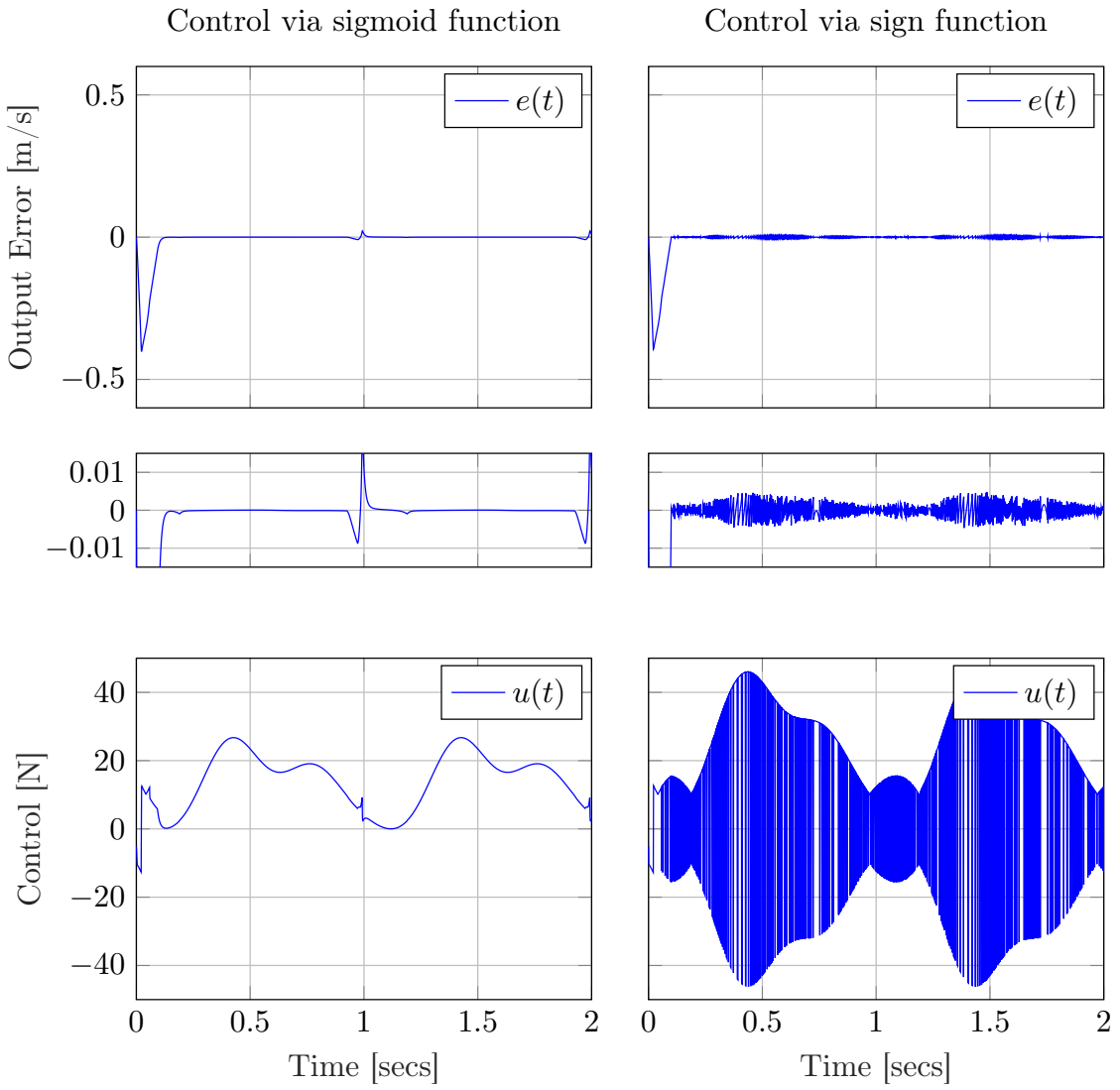


Figure 2.3: Simulation results for the SISO output tracking example with sigmoid and sign functions.

Another interesting result observed during simulations is that: even when one increases the modulation function $\rho(t)$, e.g. due to a conservative choice, the control via sigmoid(.) functions is able to adaptively adjust the control signal and stabilize the system without requiring any increase in control action. This happens because

when σ enters real sliding motion, the output $u(t)$ becomes (almost) equal to the equivalent control, which is unique and does not depend on $\rho(t)$. Figure 2.4 exemplifies this property with 3 different choices of $\rho(t)$: $\rho_1(t)$ from equation (2.42), $\rho_2(t) = 40$ and $\rho_3(t) = 100$. This is a major advantage over usual sign(.) implementations, where the control signal switches with an amplitude exactly equal to $\rho(t)$.

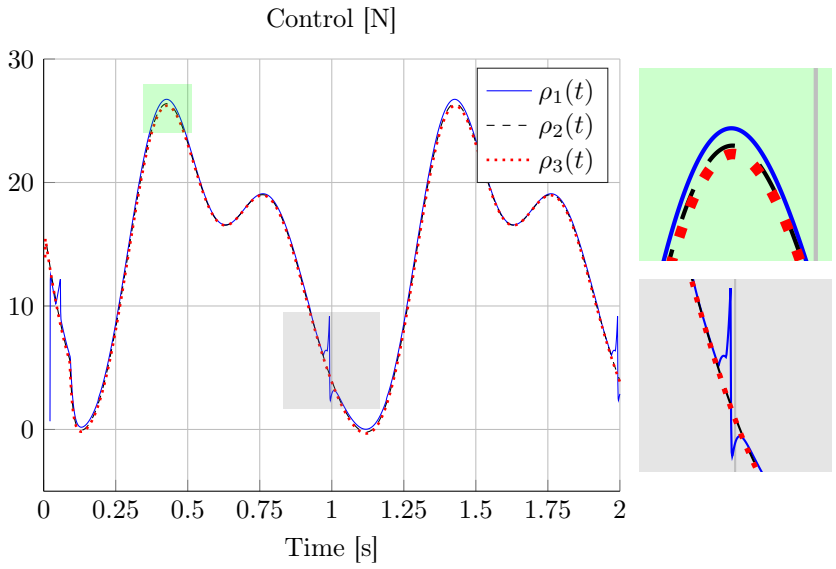


Figure 2.4: Comparison of control signals for different values of $\rho(t)$, respectively: $\rho_1(t)$ from (2.42), $\rho_2(t) = 40$ and $\rho_3(t) = 100$.

The only significant difference between possible modulation function choices lies in the system transient response. For instance, assume that at some $t = t_0$, $\sigma(t_0)$ is in sliding motion and k_p abruptly changes signal. Or, equivalently, assume that $t_0 = 0$ and the initial condition $\sigma(t_0)$ is such that $\sin(\sigma(t_0)\pi/T) = \pm 1$. Until σ enters real sliding-mode again at $t = t_1$, the control signal decreases from $u(t_0) = \pm\rho(t_0)$ to $u(t_1) = u_{eq}(t_1)$, where $u_{eq}(t)$ denotes the equivalent control at time t.

Additionally, to display the controller ability to track changes to the high-frequency gain, a final example is considered. Let system (2.39) be modified with the inclusion of $\gamma(t) = \text{sign}[\sin(2\pi t/0.6)]$ pre-multiplying $u(t)$,

$$m\ddot{y}(t) + b(y)\dot{y}(t) + ky(t) = \gamma(t)u(t) \tag{2.43}$$

such that the HFG signal is changed every 0.3 seconds. No controller parameter is changed and a simulation is performed using the controller via sigmoid functions. Figure 2.5 shows the σ switches from one sliding manifold to the other in order to track the HFG sign. Furthermore, every time $\gamma(t)$ changes sign, σ exits its sliding motion and the error starts to grow, until σ slides again, which takes less than 20 milliseconds.

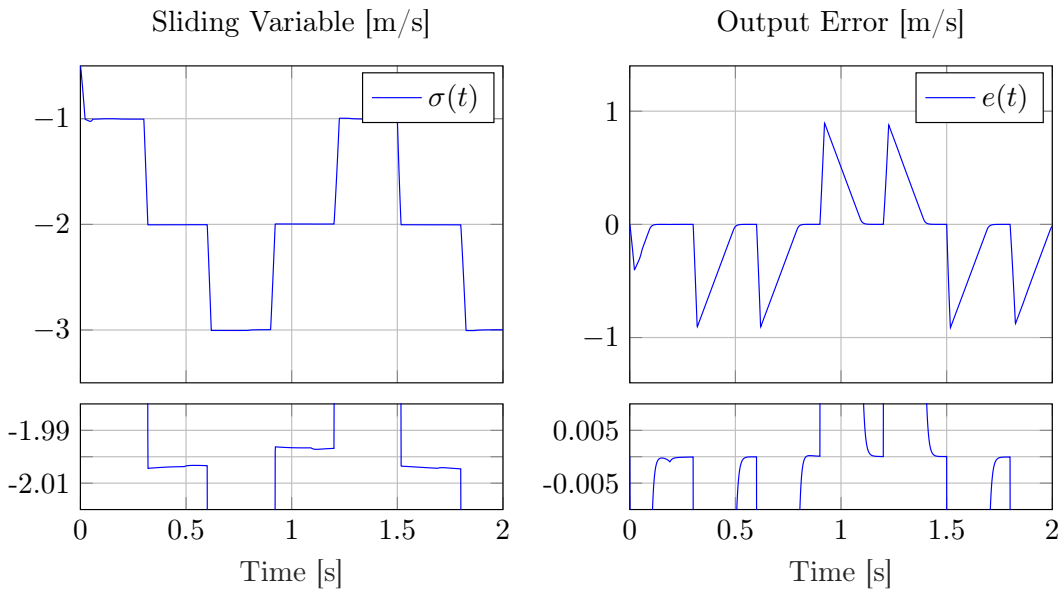


Figure 2.5: Sliding variable $\sigma(t)$ and output tracking error $e(t)$ when the HFG of (2.43) changes sign every 0.3 seconds.

2.3 Extension to Arbitrary Relative Degree

The results presented so far are restricted to SISO systems with unit relative degree from input u to output y. In this section, the unit relative degree assumption is slightly relaxed by considering systems with arbitrary relative degree from input to output, but with fast dynamics. The fast dynamics are such that the corresponding slow model has relative degree equal to one. Thus, consider system

$$\dot{x}(t) = f(x, z, \mu) + g(x, z, \mu)u(t) \tag{2.44a}$$
$$\mu\dot{z}(t) = f_z(x, z, \mu) \tag{2.44b}$$
$$y(t) = h(z) \tag{2.44c}$$

with small constant $0 < \mu \ll 1$ such that $x(t) \in \mathbb{R}^n$ varies slowly when compared to $z(t) \in \mathbb{R}^m$. In other words, we have $|\dot{x}(t)| \leq \mathcal{O}(1)$ and $|\dot{z}(t)| \leq \mathcal{O}(\mu)$.

Assumption 7 (Arbitrary Relative Degree)**.** *The relative degree of* (2.44) *from its input u to its output y is arbitrary, finite, and holds uniformly in x.*

Assumption 8 (Slow Model)**.** *The model* (2.44b) *is in standard form (Khalil 2002, page 424). Therefore, the algebraic equation $0 = f_z(x, z, 0)$ has $k \geq 1$ real roots. For simplicity, we consider $k = 1$. The slow model is defined as*

$$\dot{\bar{x}}(t) = f(\bar{x}, \varphi(\bar{x}), 0) + g(\bar{x}, \varphi(\bar{x}), 0)u \tag{2.45a}$$
$$\bar{y}(t) = (h \circ \varphi)(\bar{x}) \tag{2.45b}$$

Assumption 8 can be interpreted as substituting the fast dynamics (2.44b) in (2.44) by its quasi-steady-state form. For this reason, model (2.45) is often called a quasi-steady-state model of (2.44). In contrast, while z approaches its quasi-steady-state value, the movement is said to be contained in a boundary layer.

Since, implicitly, we assume the existence of an equilibrium $\bar{z} = \varphi(x)$ of z, it is only natural to pose some stability assumption on this equilibrium. For that, consider the change of variables $e_z(t) = z(t) - \varphi(x)$ that denote the distance from $z(t)$ to its equilibrium and the fast time variable $\tau = t/\mu$. The dynamics e_z in the τ time-scale is

$$\frac{de_z}{d\tau} = f_z(x, e_z + \varphi(x), \mu) - \mu \frac{\partial \varphi}{\partial x} f(x, e_z + \varphi(x), \mu) \tag{2.46}$$

which, by setting $\mu = 0$ and allowing x to vary slowly (instead of freezing it at its initial value), simplifies to (Khalil 2002, page 432-433)

$$\frac{de_z}{d\tau} = f_z(x, e_z + \varphi(x), 0) \tag{2.47}$$

Model (2.47) is called the boundary layer system.

Assumption 9 (Exponentially Stable Boundary Layer)**.** *The equilibrium $e_z = 0$ of the boundary layer system (2.47) is exponentially stable.*

Assumption 10 (Exponentially Stable Slow Model)**.** *Let the slow model (2.45) be controlled by (2.16), with reference output $y_m(t)$, such that $|\dot{y}_m(t)| \leq \mathcal{O}(1)$. Then, for every $y_m(t)$, there is $x_s(t, y_m)$ that is a solution to (2.45a) and the equilibrium $\bar{x} - x_s = 0$ is exponentially stable.*

Assumptions 9 and 10 are important for the application of Tikhonov's theorem, which is used below to prove local asymptotic stability of the proposed tracking and extremum-seeking controllers. Nonetheless, they are also reasonable assumptions if one attempts to develop specific Lyapunov-based stability proofs for such controllers. For this reason we highlight them separately from the following assumption, which is specific for the application of Tikhonov's theorem.

Assumption 11. *System (2.44) satisfies the conditions to apply Tikhonov's theorem on the infinite time interval (Khalil 2002, Theorem 11.2). Hence,*

$$x(t) = \bar{x}(t) + \mathcal{O}(\mu), \ \forall t \tag{2.48a}$$

$$z(t) = \bar{z}(t) + \mathcal{O}(\mu), \ \forall t \geq t_b \tag{2.48b}$$

for some $t_b > 0$.

Tikhonov's theorem guarantees that both $x(t)$ and $z(t)$, trajectories of the actual

process, will remain close to $\bar{x}(t)$ and $\bar{z}(t)$, trajectories of the slow model. To ensure that the output $y(t)$ also remains close to its quasi-steady-state counterpart $\bar{y}(t)$, one final assumption is in order.

Assumption 12. *Let $D_y \subseteq \mathbb{R}$ be compact and $y_m : \bar{\mathbb{R}}_+ \mapsto D_y$. The output function $h(.)$ is locally Lipschitz continuous in D_y, with Lipschitz constant $L_h > 0$, such that $\|y - \bar{y}\| = \|h(z) - h(\bar{z})\| \leq L_h \|z - \bar{z}\| \leq L_h \mathcal{O}(\mu)$. Thus,*

$$y(t) = \bar{y}(t) + \mathcal{O}(\mu), \ \forall t \geq t_b \tag{2.49}$$

holds for every $y(t) \in D_y$.

Note that the above assumptions allow for the direct applications of Tikhonov's theorem to prove that the proposed output tracking controller (Theorem 1) is applicable to system (2.44). This is summarized in the following theorem, which, for the above reasons, is given without proof.

Theorem 2. *Consider system (2.44) with slow model (2.45). Let the control law (2.16) be designed such that Theorem 1 is satisfied for the slow model. Then, the output error $e(t) = y(t) - y_m(t)$ origin is practical asymptotically stable (PAS), with ultimate bound $|e| \leq \mathcal{O}(\epsilon) + \mathcal{O}(\mu)$. Therefore, the output tracks the reference, but for a small residual error of order $\mathcal{O}(\epsilon) + \mathcal{O}(\mu)$.*

It is only possible to consider Theorem 2, obtained after the direct application of Tikhonov's theorem because our control formulation supports the use of both continuous and smooth control actions. Otherwise, it would not be possible to consider Tikhonov's theorem, since it does not cover systems with a discontinuous right-hand side.

2.3.1 Application to the Speed Control of Van der Pol Oscillators with Filtered Velocity

Having shown the validity of the proposed controller to systems with fast dynamics and arbitrary relative degree, the Van der Pol oscillator example of section 2.2 is revisited. Only this time, an output feedback implementation is pursued. For that, every occurrence of $\dot{y}(t)$ in the control law is replaced by $z(t)$, where $z(t)$ is a velocity estimate obtained from the lead filter

$$\mu \dot{z}(t) = -z(t) + \dot{y}(t) \tag{2.50a}$$

$$\mu z(t) = \mu z(0) - y(0) - \int_0^t z(\tau) d\tau + y(t) \tag{2.50b}$$

For simulation purposes, the filter time constant is set to $\mu = 0.005$ and the same

parameters used when resorting to state feedback are used, i.e. $\epsilon = 0.02$, $\bar{\epsilon} = 0.05$, $T = 1$, $\lambda = 5$, $\delta = 2$, and sampling period of 10 μs. The results are shown in fig. 2.6, where the continuous controller via saturated gains is compared to the discontinuous controller via sign functions. It is evident that the discontinuous implementation is unable to maintain the desired performance, which severely deteriorates when the lead filter is considered. The continuous controller via sigmoid function, on the other hand, is able to track the output. Small oscillations can be seen in the transient response due to the delay induced by the lead filter.

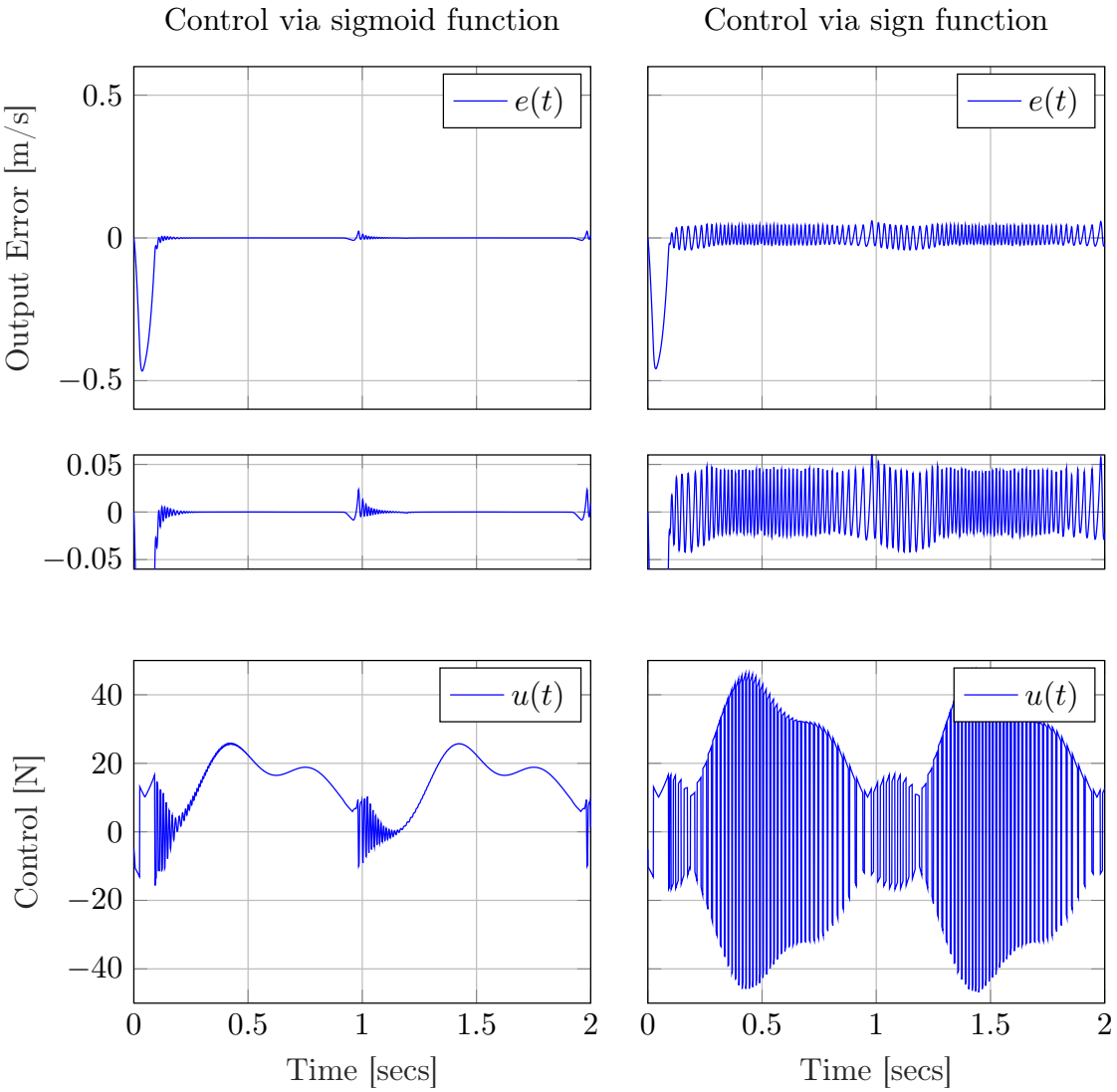


Figure 2.6: Simulation results of the Van der Pol oscillator (2.39) controlled via output feedback ($\mu = 0.005$). To the left, controller with sigmoid functions ($\epsilon = 0.02$ and $\bar{\epsilon} = 0.05$). To the right, controller with sign functions.

If one wishes to increase the filter time constant, the control parameters ϵ and $\bar{\epsilon}$ should also be increased. For instance, by selecting $\epsilon = 0.1$ and $\bar{\epsilon} = 0.25$, it is possible to set $\mu = 0.02$, which is 4 times higher than what was previously needed. Note that all parameters were proportionally increased. Using this new selection of

control parameters, the sigmoid controllers via state feedback and output feedback with lead filter are compared, see fig. 2.7. By increasing the filter time constant, the time-scale separation between the fast-dynamics (filter) and the slow-dynamics (tracking controller) decreases, explaining the deterioration in control performance. This deterioration is significantly more severe in the discontinuous implementation, which is expected. Intuitively, when the right-hand side is discontinuous, the velocity vector field changes abruptly, making it harder for the slow-model to match the process dynamics. This gives an intuition into why Tikhonov's theorem does not apply to systems with a discontinuous right-hand side.

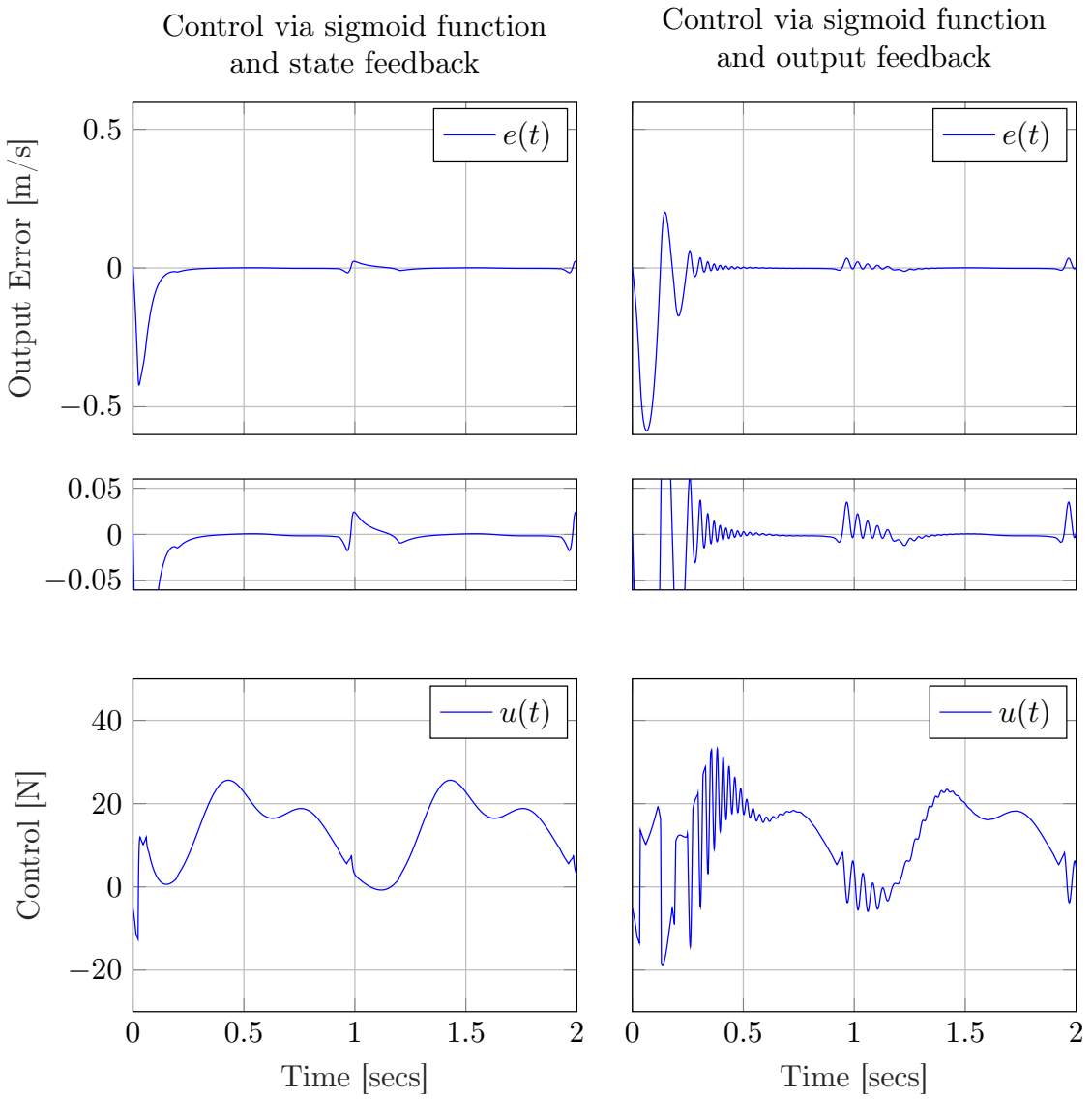


Figure 2.7: Comparison results of the Van der Pol oscillator (2.39) controlled via state and output feedback. To the left, controller via state feedback. To the right, controller via output feedback ($\mu = 0.02$). Both controllers make use of sigmoid functions ($\epsilon = 0.1$ and $\bar{\epsilon} = 0.25$).

Chapter 3

Single-Objective Real-Time Optimization

In this chapter, the previous results on output-tracking via output-feedback are applied to real-time optimization. The optimization technique developed here is based on earlier results by Drakunov et al. (1995) and Oliveira et al. (2012), and may be placed under the category of extremum-seeking control (ESC).

Optimization problems arise in most engineering applications. Given a set of objective functions and system constraints, optimization is used to determine the optimal process input and design parameters that are needed to meet these conditions. Constraints are related to the system's optimal performance, but they can also drive the choice of an optimization algorithm. For instance, when modifications must be made in real-time, and the objective functions are measured but unknown, the problem becomes harder to solve since process information and computational power are limited. In the optimal control literature, extremum-seeking control (ESC) algorithms are a suitable choice when it comes to dealing with real-time optimization of systems with unknown (or partially unknown) models.

When the objective function is known, or at least its model is known, but the parameters are unknown, it is common to use other techniques from the optimal control literature, such as model predictive control. Optimal control is an exciting and vast branch of control theory, from which, unfortunately, we cover only the small subset related to ESC. For an in-depth view of optimal control, the reader is referred to (Liberzon 2011).

In the following section, we describe the optimization problem and how it relates to the output-tracking of processes with unknown control direction. The objective of this section is to help the reader understand this connection while also writing the underlying assumptions needed to establish it.

3.1 Problem Formulation

In this Thesis, we write single-objective optimization with one decision variable as a standard optimal control problem[1]:

$$\min_{u} : \nu = J(y) \tag{3.1a}$$

$$\text{s.t.} : \dot{x}(t) = f(x) + g(x)u(t) \tag{3.1b}$$

$$y(t) = h(x) \tag{3.1c}$$

The function $J : \mathbb{R} \mapsto \mathbb{R}$ is called the objective function, and represents a performance that must be optimized, e.g. the power output in power plants, the temperature in heat exchangers, and a company's profit. In (3.1), the objective is written as a function the process output. Several variations of this problem include the control effort $u(t)$ in the objective function, in order to waste less control energy, or an explicit dependence on time. On the other hand, we only take the control effort into account in the process dynamics, which is seen as an equality constraint to the optimization problem. It is usual to consider other equality and inequality constraints, but they are not accounted for in this work.

To consider the above optimization problem solved, one must find the control input $u(t)$ that drives the process output such that the objective function $J(y)$ is minimized over time. The two usual approaches to this problem are (i) to solve it offline and then apply the pre-computed control effort over a fixed period of time, or (ii) to continuously update the control input. A widely adopted technique that can tackle both of the approaches above simultaneously is model predictive control (Camacho and Bordons Alba 2007). Extremum-seeking falls into the category of continuously updating the control input.

If we take the objective function and differentiate it with respect to time,

$$\dot{\nu}(t) = \frac{\partial J(y)}{\partial y} \left[L_f h(x) + L_g h(x) u(t) \right] \tag{3.2}$$

the connection to output-tracking of systems with unknown control direction becomes apparent. With respect to the process output y, the high-frequency gain is $L_g h(x)$. However, if we take the objective as the process output, the high-frequency gain becomes

$$k_p(x) = \frac{\partial J(y)}{\partial y} L_g h(x) \tag{3.3}$$

[1] We write optimization as a minimization problem, but all results are readily applied to maximization problems.

which is only a function of x, since the mapping from y to x is static. Thus, even if the process HFG does not change signal, the HFG with respect to the objective changes every time the output crosses the optimal value. For example, when dealing with minimization, $\partial J/\partial y < 0$ to the left of the minimum and $\partial J/\partial y > 0$ to the right.

Remark 2. *For simplicity, in what follows we ignore the variable ν and use only $y = h(x)$ as the objective function. This simplification does not impose any loss of generality, since we might consider $\nu = J \circ h(x) = h'(x)$. We use the notation $y = h(x)$ to keep the results closer to the ones already established for output-tracking.*

To adapt the previous output-tracking results from extremum-seeking, some changes are in order since Assumption 2 is no longer valid. Also, to derive global results on ESC, some assumptions on the output function are needed. Otherwise, convergence to local optima can still be established, as demonstrated at the very end of section 3.2. These assumptions are provided below.

Assumption 13 (Unique Extremum)**.** *The output function $h(.)$ in (2.1b) is continuous and differentiable almost everywhere in its domain and have a unique strict global minimizer x^*, with minimum $y^* = h(x^*)$.*

Assumption 14 (Known Gradient Bound)**.** *For any chosen $\Delta > 0$, there exists $\bar{\Delta}(\Delta) > 0$ and known constants $L(\Delta)$ and $\underline{k}_p(L)$ such that*

$$\begin{aligned} &L \leq \left\| \frac{\partial h(x)}{\partial x} \right\| \quad \text{and} \quad 0 < \underline{k}_p(L) \leq |k_p(x)| = |L_g h(x)| \\ &\forall x \notin D_\Delta = \left\{ x \in \mathbb{R}^n : \|x - x^*\| \leq \bar{\Delta}\,,\, |y(x) - y^*| \leq \Delta \right\} \end{aligned} \tag{3.4}$$

holds uniformly on t. This means that a lower bound L for the derivative can be established for any Δ-neighborhood of the optimum y^* and that the system remains controllable outside this vicinity.

Note that Assumption 14 replaces Assumption 2 close to the optimizer and is equivalent to the latter when the states are outside of the Δ-neighborhood, i.e. $\forall x \notin D_\Delta$.

3.2 Continuous Sliding-Mode Extremum-Seeking-Controller for Input-Affine Processes

To implement the extremum-seeking solution, the control law developed in section 2.2 is used to track a monotonic function. Since the solution presented here is for minimization problems, the output reference $y_m : \mathbb{R}_+ \mapsto \mathbb{R}$ is monotonically

decreasing, converges to a constant value, i.e. $\exists \lim_{t\to\infty} y_m(t)$, and chosen such that there exists $t^* \geq 0$ for which $t \geq t^* \implies y_m(t) \leq y^*$. This last requirement ensures that the model reaches the optimum, and it is reasonable since bounds for the optimal value are usually known in practice.

Remark 3. *To adapt this control law for maximization problems, one need only change $y_m(.)$ from a monotonically decreasing to a monotonically increasing function, such that for $t \geq t^* \implies y_m(t) \geq y^*$.*

Below, we enunciate and prove the theorem that states the results of the sliding-mode controller for real-time optimization. Considering our general formulation with a sigmoid function and the function $f_e(e)$ in the σ-dynamics, our previous choice of Lyapunov function (2.22) based on the absolute value makes the proof of the next theorem easier.

Theorem 3. *Consider system (2.1) with output error (2.7) based on a differentiable and monotonically decreasing output model $y_m(t)$, and control law (2.8), (2.16),*

$$f_e(e) = \lambda \operatorname{sigmoid}_{\bar{\epsilon}}(e) \tag{3.5}$$

with $\lambda > 0$ and $\bar{\epsilon} > 0$. Then, the extremum-seeking error

$$e^*(t) = y(t) - y^* \tag{3.6}$$

is uniformly globally practical asymptotically stable, with ultimate bound $|e^*| \leq \Delta + \big(1 + \lambda/\delta\big)T \leq \Delta + \mathcal{O}(T)$.

Proof. To prove that oscillations above the minimum y^* are bounded by $|y - y^*| \leq \Delta + \mathcal{O}(T)$, it is necessary and sufficient to show that after regaining controllability, i.e. $x \notin D_\Delta \iff |y - y^*| > \Delta$, at a time t_0, the sliding variable σ will enter real sliding-mode at t_1, such that $|y(t_0) - y(t_1)| \leq \mathcal{O}(T)$. Thereafter, the error decreases until $|y - y^*| \leq \Delta$ and the system loses controllability once again. Hence, $y(t_1)$ is the maximum value the output is allowed to reach before decreasing, and it is already a worst case estimate.

This proof is made in two steps. First, it is shown that the amount of time needed for σ to enter real sliding-mode is $t_1 - t_0 \leq \mathcal{O}(T)$. Second, this implies that the output distances itself from the Δ-neighborhood by $|y(t_1) - y(t_0)| \leq \mathcal{O}(T)$.

At time t_0, when the system leaves D_Δ, controllability is regained and, assuming that σ is not in real sliding-mode, the Lyapunov function (2.22) derivative satisfies (2.31). Therefore,

$$V(t_1) - V(t_0) = \int_{t_0}^{t_1} \dot{V}(\tau) d\tau \leq -\delta(t_1 - t_0)$$

$$t_1 - t_0 \leq [V(t_0) - V(t_1)] / \delta$$

from the proof of Theorem 1, we know that $V(t_0) - V(t_1) \leq T$, therefore

$$t_1 - t_0 \leq T/\delta \leq \mathcal{O}(T) \tag{3.7}$$

Recall the sliding variable σ definition (2.8). From this equation and the time difference inequality (3.7),

$$\begin{aligned}
\sigma(t_1) - \sigma(t_0) &= e(t_1) - e(t_0) + \int_{t_0}^{t_1} f_e(e) d\tau \\
|e(t_1) - e(t_0)| &\leq |\sigma(t_1) - \sigma(t_0)| + \int_{t_0}^{t_1} |f_e(e)| \, d\tau \\
|e(t_1) - e(t_0)| &\leq |\sigma(t_1) - \sigma(t_0)| + \lambda \int_{t_0}^{t_1} d\tau \\
|e(t_1) - e(t_0)| &\leq (1 + \lambda/\delta) \, T \leq \mathcal{O}(T)
\end{aligned} \tag{3.8}$$

where $|\sigma(t_1) - \sigma(t_0)| \leq T$ was used. From the output error definition (2.7),

$$|y(t_1) - y(t_0)| \leq |e(t_1) - e(t_0)| + |y_m(t_1) - y_m(t_0)| \tag{3.9a}$$

Since $y_m(t)$ maps to a complete set (the real line) and converges to a constant value, it forms a Cauchy sequence, implying $|y_m(t_1) - y_m(t_0)| \to 0$ as $t \to \infty$. Thus, ultimately,

$$|y(t_1) - y(t_0)| \leq |e(t_1) - e(t_0)| \leq (1 + \lambda/\delta) \, T \leq \mathcal{O}(T) \tag{3.10}$$

Since $y(t_1) - y^* = y(t_1) - y(t_0) + y(t_0) - y^* = y(t_1) - y(t_0) + \Delta$, it follows that

$$|y(t_1) - y^*| \leq (1 + \lambda/\delta)T + \Delta \leq \mathcal{O}(T) + \Delta \tag{3.11}$$

which concludes the proof. ■

Remark 4. *If for some time $t_2 \in (t_0, t_1)$ the system state happens to enter D_Δ such that $|y - y^*| \leq \Delta$, the distance $|y(t_2) - y(t_0)| \leq |y(t_1) - y(t_0)|$, since it is assumed that the output distances itself from $y(t_0)$ during the interval $[t_0, t_1]$. Therefore, the bound $|y - y^*| \leq \Delta + \mathcal{O}(T)$ remains valid.*

If there are multiple isolated optima, convergence to one of these local optima can still be established. For that, the bound L computed for the Δ-vicinity in Assumption 14 must hold in a neighborhood

$$D_\Delta = \left\{ x \in \mathbb{R}^n : \|x - x^*\| \leq \bar{\Delta} \, , \, |y - y^*| \leq \Delta + (1 + \lambda/\delta)T \right\} \tag{3.12}$$

This guarantees that the output will remain bounded to the Δ-vicinity of one of the optima and will not exit this vicinity.

3.3 Stabilization of Systems with Unknown Output Sign

Constraints on a mobile robot sensor configuration or sensor malfunction are some of the reasons for uncertain output measurements. For instance, imagine a mobile robot carrying a single range-finder and placed at a known position, but unknown orientation, from some wall. The control objective is to make the robot face the wall. This scenario is depicted in fig. 3.1. Note that, if the distance from the wall is given $l_w > 0$, and the sensor measures a distance $l_s > 0$, the vehicle measured orientation with respect to the wall, denoted θ_s, is

$$\theta_s = \arccos(l_w/l_s) = |\theta| \tag{3.13}$$

where θ is the actual orientation.

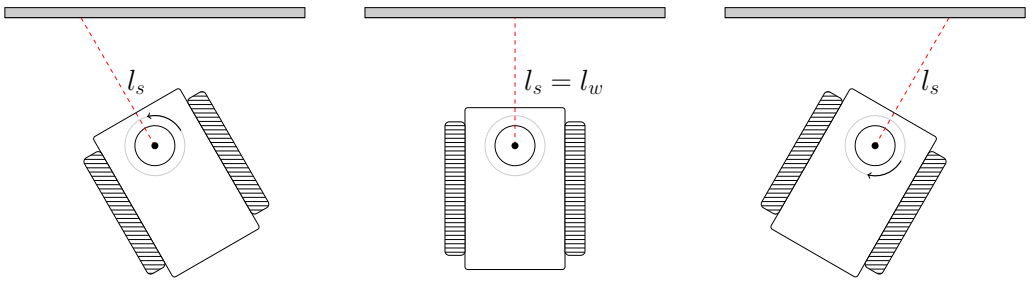


Figure 3.1: Illustration of a mobile robot carrying a laser range-finder and measuring its distance from a wall.

This example motivates the development of a control law for cases where only the output absolute value is available for measurement. Thus, consider the dynamics

$$\dot{x}(t) = f(x) + g(x)u(t) \tag{3.14a}$$
$$y(t) = h(x) \tag{3.14b}$$
$$y_s(t) = |y(t)| \tag{3.14c}$$

with state vector $x : \bar{\mathbb{R}}_+ \mapsto \mathbb{R}^n$, unmeasured output $y : \bar{\mathbb{R}}_+ \mapsto \mathbb{R}$ and measurement $y_s(t) : \bar{\mathbb{R}}_+ \mapsto \mathbb{R}_+$. Let $L_g h(x)$ be bounded away from zero such that Assumption 2 holds, and define the high-frequency gain with respect to the measured output

$$k_p(x) = \text{sign}(y) L_g h(x) \tag{3.15}$$

The control objective is to asymptotically stabilize $y(t)$ at $y = 0$ using only the measurable signal $y_s(t)$. On a side note, we could also consider $L_g h(x)$ changing sign over time, but then the problem would fall back to the same case as the one considered in section 3.2.

Since $y_s = 0$ if and only if $y = 0$, and the absolute value function is continuous, it follows that asymptotic stabilization of $y_s = 0$ implies asymptotic stabilization of $y = 0$. Hence, a straightforward approach is to treat the control problem as an extremum-seeking problem. The major differences from this formulation to other extremum-seeking schemes mentioned in this work are that (i) the minimum output value is known ($y^* = 0$) a priori and (ii) the HFG $k_p(x)$ is discontinuous at $y = 0$, where it changes sign. These imply the following lemma, which replaces Assumption 14.

Lemma 1. *There exists a known constant L such that*

$$L \leq \left| \frac{\partial h(x)}{\partial x} \right|, \ \forall x \in D_\Delta = \{x \in \mathbb{R}^n : |y| > 0\} \tag{3.16}$$

holds uniformly on t. Therefore, Δ in Assumption 14 can be set to zero.

With the above lemma and the already mentioned assumptions, the following theorem can be established.

Theorem 4. *Consider system (3.14) with output error $e(t) = y_s(t)$ and control law (2.8), (2.16),*

$$f_e(e) = \lambda \operatorname{sigmoid}_\epsilon(e) \tag{3.17}$$

with $\lambda > 0$. Then, the output origin $y = 0$ is UGPAS, with ultimate bound $|e| \leq (1 + \lambda/\delta)T \leq \mathcal{O}(T)$.

Proof. The proof is straightforward by considering the proof of Theorem 3 and letting $y(t_0) = y^*$, where t_0 is the time at which controllability is lost. ■

Remark 5. *Since controllability is only lost at $y(t_0) = y^*$ and immediately regained afterwards, different from the ESC results presented in chapter 3, the error ultimate bound does not include a parcel Δ. This is the same bound experienced by a tracking controller if one lets the HGF to change its sign, without crossing $k_p(x) = 0$, as illustrated in the Van der Pol oscillator example of section 2.2.3.*

3.3.1 Application to Mobile Robots

The results presented in this section are quite general and apply to a wide class of input-affine nonlinear plants with (almost) arbitrary output functions. A subclass

of such plants that also plays an important role in control theory is given below

$$\dot{x}(t) = b(x)\left[u(t) + d(t)\right] \tag{3.18a}$$

$$y(t) = |x| \tag{3.18b}$$

with state $x : \bar{\mathbb{R}}_+ \mapsto \mathbb{R}$, output $y : \bar{\mathbb{R}}_+ \mapsto \mathbb{R}$ and control input $u : \bar{\mathbb{R}}_+ \mapsto \mathbb{R}$. An example system that follows (3.18) is the orientation dynamics of differential drive mobile robots, as depicted in fig. 3.1.

For simplicity, we assume that $d(t)$ is not estimated and that the following bounds are known

$$\underline{b}(y) \leq |b(x)| \tag{3.19a}$$

$$d(t) \leq \bar{d}(t) \tag{3.19b}$$

Using these bounds, the modulation function $\rho(t)$ is defined as

$$\rho(t, y) = \frac{\kappa}{\underline{b}(y)}(\lambda + \delta) + \bar{d}(t) \tag{3.20}$$

For simulation purposes, let $b(x) = 1$ deg/s for $x \in [0, \pi/2] \cup [\pi, 3\pi/2]$, $b(x) = 0.4$ deg/s for $x \in (\pi/2, \pi) \cup (3\pi/2, 2\pi)$, $d(t) = 0$ for $t \in [0, 2] \cup [4, \infty)$ and $d(t) = 10$ for $t \in (2, 4)$. The change in $b(x)$ represent a loss of traction due to slippage on the mobile robot wheels as illustrated in fig. 3.2.

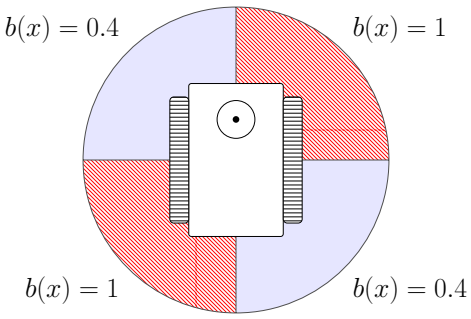


Figure 3.2: Representation of different slip coefficient impacting controllability.

As a worst case scenario, we take $\underline{b} = 0.3$ deg/s and $\bar{d}(t) = 15$. To recover from a 10 deg error in 1 second, let $\lambda = 10$ deg/s and in order to achieve practical stability with at most 2 deg error, let $\delta = 5$ deg/s and $T = 2/3$ deg. The controller is implemented via saturated gain sigmoid functions, with $\epsilon = 0.5$ deg, $\bar{\epsilon} = 1$ deg and $\rho(t, y) = 65$. Finally, to meet the 2 deg error specification, $(1 + \lambda/\delta)T \leq 2$, which, for $\lambda = 10$ and $\delta = 5$, implies $T = 2/3$ deg.

The simulation output is shown in fig. 3.3, where the controller is implemented

in discrete-time with sampling period of 10 ms and the first-order motor dynamics

$$H(s) = \frac{1}{0.015s + 1} \tag{3.21}$$

is considered. The sampling period is selected to be 10 ms to bring the simulation closer to typical processing constraints of controllers used in mobile robotics. The results are compared with a traditional sigmoid $_\epsilon(.) = \text{sign}(.)$ controller to show that the continuous approximation via saturated gain improves performance and reduces chattering.

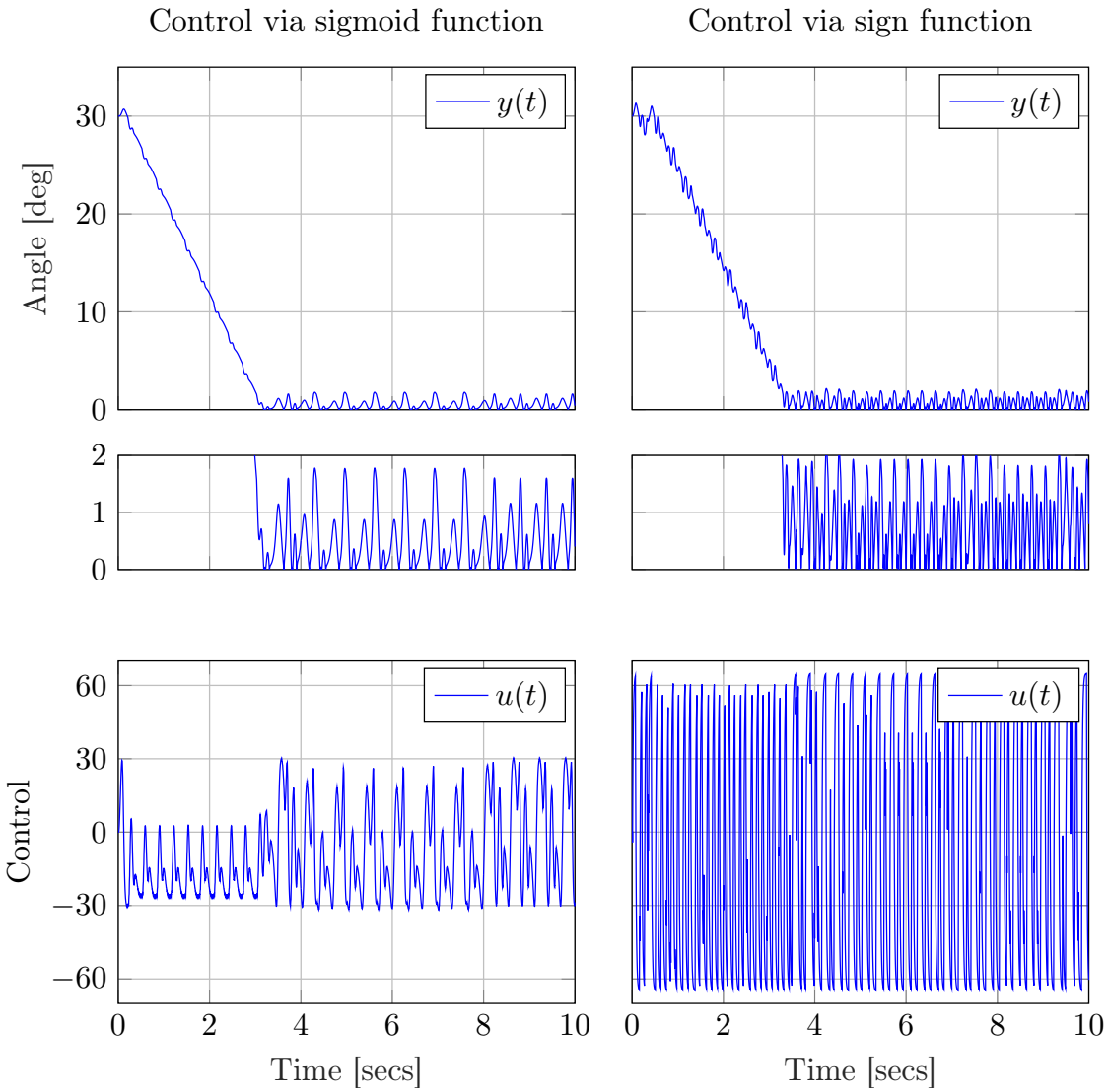


Figure 3.3: Simulation results for the mobile robot example (3.18) with motor dynamics (3.21).

3.4 Extension to Arbitrary Relative Degree

To conclude this chapter, we provide one final result that shows that the proposed extremum-seeking control law, the same as the output-tracking control law for systems with unknown control direction, can be used for systems with arbitrary relative degree. This is a short section, since it is a direct consequence of the results already discussed in section 2.3. Since the results also follow from the direct application of Tikhonov's theorem under the same assumptions as in section 2.3, we give the following theorem without proof.

Theorem 5. *Consider system* (2.44) *with slow model* (2.45). *Let the control law* (2.16) *be designed such that Theorem 3 is satisfied for the slow model, for some objective function satisfying Assumptions 13 and 14. Then, the extremum-seeking error* $e^*(t) = y(t) - y^*$ *origin is PAS, with ultimate bound* $|e^*| \leq \Delta + \big(1 + \lambda/\delta\big)T + \mathcal{O}(\mu) \leq \Delta + \mathcal{O}(T) + \mathcal{O}(\mu)$.

An example that illustrates this extension to an arbitrary relative degree is the previous one for mobile robots, with results shown in fig. 2.6. In this example, the relative degree is equal to two, and given that the motor dynamics is faster than the ESC dynamics, stabilization (minimization of the output absolute value) is still possible given an appropriate choice of control parameters.

Finally, our approach to arbitrary relative degree follows the usual generalizations made in the model-based and perturbation-based extremum-seeking literature. Nonetheless, one important advantage of our results, which are based on sliding-mode control, is that they are do not rely on any time-scale separation when the process has unit relative degree.

Chapter 4

Multi-Objective Real-Time Optimization with Multiple Decision Variables

In chapter 3, we have discussed optimization in the sense of single-objective problems with only one decision variable. Even though the results are significant, single-objective problems and, especially, problems with only one decision variable are not that common. In practice, more than one objective needs to be satisfied, *as best as possible*, and there are multiple decision variables capable of changing these objectives. While the notion of *best* is intuitive in single-objective problems, this is not necessarily the case in multiple-objective optimization. When there are multiple objectives, there might be several equally good solutions, and selecting among them is usually a matter of establishing some preference measurement.

In the next section, we formulate the multi-objective optimization problems considering two distinct concepts of optimal solutions that are commonly used in multi-objective optimization — Pareto efficiency and Nash equilibrium.

4.1 Problem Formulation

Similarly to the problem formulation for single-objective optimization problems with only one decision variable, we write the multi-objective optimization problem with multiple decision variables, or MIMO optimization for short, as[1]

$$\min_{u} : \nu = J(y) = \begin{bmatrix} J_1(y) & \dots & J_{n_J}(y) \end{bmatrix}^{\mathsf{T}} \tag{4.1a}$$

$$\text{s.t.} : \dot{x}(t) = f(x) + g(x)u(t) \tag{4.1b}$$

$$y(t) = h(x) \tag{4.1c}$$

[1] We define what we mean by minimum of a vector in the next section.

This time, however, we have $n_J \geq 1$ objectives functions , with $J : \mathbb{R}^{n_y} \mapsto \mathbb{R}^{n_J}$, $n_y \geq 1$ outputs, with $h : \mathbb{R}^n \mapsto \mathbb{R}^{n_y}$, and $n \geq 1$ states, with $f : \mathbb{R}^n \mapsto \mathbb{R}^n$, and $n_u \geq 1$ inputs (decision variables), with $g : \mathbb{R}^n \mapsto \mathbb{R}^{n_u}$. The optimization problem described by (4.1), just as it was in the SISO case, does not cover equality or inequality constraints other than the process dynamics. Regarding the process constraints, for a part of this chapter, we will consider static systems, such that the optimization problem is rewritten as

$$\min_x : \nu = J(y) = \begin{bmatrix} J_1(y) & \dots & J_{n_J}(y) \end{bmatrix}^\mathsf{T} \tag{4.2a}$$

$$\text{s.t.} : y(t) = h(x) \tag{4.2b}$$

4.1.1 Optimum Characterization

We consider two interpretations of optimum of multi-objective problems in this Thesis — Pareto efficiency and Nash equilibrium. In this chapter, the algorithms described in sections 4.2 and 4.3 produce solutions in Nash equilibrium, while the last algorithm, described in chapter 5, is capable of reaching Pareto efficient solutions.

Nash Equilibrium

Nash equilibrium, as defined and proved by Nash (1950), applies to *n-person games*. First, we quote the definition of an n-person game from Nash (1951): "*For us an n-person game will be a set of n players, or positions, each with an associated finite set of pure strategies; and corresponding to each player, i, a payoff function, p_i, which maps the set of all n-tuples of pure strategies into the real numbers. When we use the term n-tuple we shall always mean a set of n items, with each item associated with a different player.*".

A pure strategy is deterministic and defines how a person plays a game. To establish a parallel between our approach to multi-objective optimization and n-person games, consider again the optimization problem described by (4.1). Interpreting equation (4.1) as an n-person game, the payoffs are the objective functions $J_i(.)$ and the pure strategies are the control algorithms, which modify the game through the control actions $u_i(.)$, components of the control vector $u(.)$. The process dynamics is a constraint to the optimization problem. Since in n-player games there are as many payoffs as players and strategies, it also follows that $n_J = n_y = n_u$, i.e. there are as many objective functions as process outputs and process inputs.

In an n-person game, it is said that an n-tuple of pure strategies counters another tuple when it yields the highest possible payoff against the $n - 1$ strategies of the other players (which remain unchanged).

Definition 4 (Nash Equilibrium (Nash 1950))**.** *In an n-person game, an n-tuple of pure strategies is a Nash equilibrium points if it counters itself. In other words, an n-tuple is a Nash equilibrium when no player can benefit from unilaterally changing its strategy.*

Definition 5. *Let $J \in C^1$. In terms of the derivatives of the objective functions $J_i(.)$ and of the way we established the parallel between n-person games and multi-objective optimization, a solution y^* is a Nash equilibrium point if and only if*

$$\left.\frac{\partial J_i(y)}{\partial y_i}\right|_{y=y^*} = 0, \quad \forall i \in [1, n_y]$$

Pareto Efficiency

Compared to Nash equilibrium, reached when no improvement can be achieved by unilaterally changing one's strategy, Pareto efficiency is a state of allocation of resources where there can be no improvement in one objective function without deteriorating the performance of another. We introduce the concept using the term "allocation of resources" because, initially, Pareto efficiency was developed as a tool to study economic efficiency, and it is now widely used throughout microeconomics, welfare economics, and multi-objective optimization.

Pareto improvements allow for multiple strategies to change at once such that at least one improves, and none deteriorates. In terms of multi-objective optimization and, specifically, problem (4.1), finding a Pareto efficient solution involves executing Pareto improvements through the decision variables y_i until no further Pareto improvement is possible.

Definition 6 (Pareto Efficiency)**.** *A solution is said to be Pareto efficient if it is not strictly dominated by any other solution. A solution y' is strictly dominated by some other solution y'' if $y''_j \leq y'_j$ ($y''_j \geq y'_j$), for every j, and $y''_k < y'_k$ ($y''_k > y'_k$), for at least one k.*

Definition 7 (Pareto Front)**.** *The Pareto front is the set of all Pareto efficient solutions. This set is denoted by $P(J)$.*

Once again, in terms of the derivatives of the objective functions J_i, we define below the conditions for a solution to belong to the Pareto front $P(J)$.

Definition 8. *Let $J \in C^1$. For a problem with an identical number $n_J = n_y$ of objective functions J_i and decision variables y_j, a necessary condition for $y^* = h(x^*) \in$*

$\mathbb{R}^{n_y}$ *to be Pareto optimal, thus belong to the Pareto front* $P(J)$*, is that*

$$\left|\begin{matrix} \dfrac{\partial J_1}{\partial y_1} & \dfrac{\partial J_1}{\partial y_2} & \cdots & \dfrac{\partial J_1}{\partial y_n} \\ \dfrac{\partial J_2}{\partial y_1} & \dfrac{\partial J_2}{\partial y_2} & \cdots & \dfrac{\partial J_2}{\partial y_n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial J_n}{\partial y_1} & \dfrac{\partial J_n}{\partial y_2} & \cdots & \dfrac{\partial J_n}{\partial y_n} \end{matrix}\right|_{x=x^*} = 0 \tag{4.3}$$

In other words, the Jacobian matrix must be rank deficient. Additionally, if the $h(x)$ *is strictly convex, then the above condition becomes both necessary and sufficient.*

4.2 SM-ESC for Nash Equilibrium Seeking in Static Noncooperative Infinite Games

In this section, we describe our results (Peixoto et al. 2020) on real-time multi-objective optimization via sliding-mode extremum-seeking. These results apply to static plants, with optimization problems described by (4.2). In the optimization literature, such problems are called static noncooperative infinite games (Basar and Olsder 1999).

To solve the optimization problem, we follow the same methodology used in chapter 3. Let $n_J = n_y = n$, and write

$$\nu(h(x)) = J(h(x)) = \begin{bmatrix} J_1 \circ h_1(x) & \dots & J_n \circ h_n(x) \end{bmatrix}^\intercal \tag{4.4}$$

Combining the objective function with the output map to make the notation less cumbersome, we use only

$$y(x) = h(x) = \begin{bmatrix} h_1(x) & \dots & h_n(x) \end{bmatrix}^\intercal \tag{4.5}$$

to denote the objective function image $\nu(h(x))$. By differentiating (4.5) and setting $\dot{x} = u$, one obtains

$$\dot{x}(t) = u(t) \tag{4.6a}$$

$$\dot{y}(t) = \frac{\partial h}{\partial x}(x)\, u(t) \tag{4.6b}$$

where $u(t) \in \mathbb{R}^n$ is a control signal yet to be specified. To ensure that output-feedback controllers can be used and no signal escapes in finite time, we make the following assumption.

Assumption 15 (Unboundedness Observability (Angeli and Sontag 1999))**.** *The closed-loop system* (4.6) *possesses an unboundedness observability property, such that if any internal signal escapes in some finite time, then all other signals escape at the same time.*

In order to accurately define the minimum-seeking problem considered here, we make the following assumptions, which are quite common in multi-objective optimization problems.

Assumption 16 (Unique Optimizer)**.** *Let $\theta_j \in \mathbb{R}^{n-1}$ denote all elements of $x \in \mathbb{R}^n$, but for the j-th entry, and let $h \in C^2$. Then, each function $h_j(\cdot)$ is unimodal w.r.t. x_j, thus, for every fixed $\theta_j \in \mathbb{R}^{n-1}$ there exists a unique minimizer $x_j^*(\theta_j)$ such that*

$$\left.\frac{\partial h_j}{\partial x_j}\right|_{x_j=x_j^*} = 0 \quad \text{and} \quad \left.\frac{\partial^2 h_j}{\partial x_j^2}\right|_{x_j=x_j^*} > 0 \tag{4.7}$$

where $x_j^* : \mathbb{R}^{n-1} \mapsto \mathbb{R}$ is a continuous function of the $n-1$ variables θ_j.

The objective of the controller we propose in this section is to achieve a Nash equilibrium. Since the controller behaves as an extremum-seeking controller, the expected result is to converge to a Nash equilibrium, and remain arbitrarily close to it. Mathematically, we describe this behavior in the following definition.

Definition 9 (Δ-Neighborhood)**.** *The Δ-neighborhood of the minimizer $x_j^*(\theta_j)$ is the region of diameter Δ along $x_j^*(\theta_j)$, i.e., the set*

$$\mathcal{D}_{\Delta_j} = \left\{ x \in \mathbb{R}^n : |x_j - x_j^*(\theta_j)| \leq \frac{\Delta}{2} \right\} \tag{4.8}$$

where $\Delta > 0$ is a constant.

Assumption 17 (Bounded Jacobian)**.** *Outside the Δ-neighborhood, the main diagonal of the HFG matrix are bounded away from zero,*

$$0 < \left| \frac{\partial h_j}{\partial x_j} \right|, \quad \forall x \notin \mathcal{D}_{\Delta_j} \tag{4.9}$$

with $j \in \{1, \ldots, n\}$.

Note that, for simplicity, only one parameter Δ is considered for all regions $\mathcal{D}_{\Delta_j}$. The minimum-seeking problem considered in this section can now be stated as follows:

(I) *Extremum Points:* Let Θ_j^* denote the nonempty set of minimizers

$$\Theta_j^* = \left\{ x^* \in \mathbb{R}^n : \frac{\partial h_j(x^*)}{\partial x_j} = 0\,,\ \frac{\partial^2 h_j(x^*)}{\partial x_j^2} < 0 \right\} \subseteq \mathcal{D}_{\Delta_j} \tag{4.10}$$

which is the same as the set $\mathcal{D}_{\Delta_j}$ with $\Delta = 0$. We say that $y_j^* = h_j(x^*)$ is an extremum (minimum) of the smooth mapping $h_j(\cdot)$ when $x^* \in \Theta_j^*$. Moreover, we say that x^* is a minimizing point of $h(\cdot)$ when x^* is sufficiently close to

$$\Theta^* = \bigcap_{j=1}^{n} \Theta_j^* \tag{4.11}$$

where sufficiently close means that $x^* \in \mathcal{D}_{\Delta_j}$. Set Θ^* contains all Nash equilibrium points for the optimization problem (4.2).

(II) *Parametric Uncertainties:* We assume that Θ^*, x^*, $h(\cdot)$ and its gradient are unknown to the control designer.

Remark 6 (Equilibrium Points)**.** *When $x \to x^* \in \Theta^*$ one has that $\partial h_j/\partial x_j \to 0$, for every j. Thus, the set Θ^* contains all Nash equilibrium points. Note, however, that when the mapping $h(\cdot)$ is weakly coupled, x^* is also close to Pareto efficient, since solutions belonging the Pareto optimum set satisfy*

$$\left| \frac{\partial h}{\partial x} \right| = \begin{vmatrix} \dfrac{\partial h_1}{\partial x_1} & \cdots & \dfrac{\partial h_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial h_n}{\partial x_1} & \cdots & \dfrac{\partial h_n}{\partial x_n} \end{vmatrix} = 0 \tag{4.12}$$

which for weakly coupled systems, with (4.12) in dominant triangular or dominant diagonal form, implies $\partial h_j/\partial x_j$ close to zero. If (4.12) is triangular or diagonal, then both sets (Pareto efficient and Nash equilibrium points) coincide.

4.2.1 The Class of the Input-Output Mapping

The first derivative of the output y with respect to time can be written as in (4.6), where the high-frequency gain matrix is written as

$$\frac{\partial h}{\partial x}(x) = \begin{bmatrix} \dfrac{\partial h_1}{\partial x_1} & \dfrac{\partial h_1}{\partial x_2} & \cdots & \dfrac{\partial h_1}{\partial x_n} \\ \dfrac{\partial h_2}{\partial x_1} & \dfrac{\partial h_2}{\partial x_2} & \cdots & \dfrac{\partial h_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial h_n}{\partial x_1} & \dfrac{\partial h_n}{\partial x_2} & \cdots & \dfrac{\partial h_n}{\partial x_n} \end{bmatrix} = k_p(x) + \tilde{k}_p(x) \tag{4.13a}$$

$$k_p(x) = \begin{bmatrix} \dfrac{\partial h_1}{\partial x_1} & 0 & \dots & 0 \\ 0 & \dfrac{\partial h_2}{\partial x_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \dfrac{\partial h_n}{\partial x_n} \end{bmatrix} \quad , \quad \tilde{k}_p(x) = \begin{bmatrix} 0 & \dfrac{\partial h_1}{\partial x_2} & \dots & \dfrac{\partial h_1}{\partial x_n} \\ \dfrac{\partial h_2}{\partial x_1} & 0 & \dots & \dfrac{\partial h_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial h_n}{\partial x_1} & \dfrac{\partial h_n}{\partial x_2} & \dots & 0 \end{bmatrix} \tag{4.13b}$$

where $k_p(x)$ holds the diagonal and $\tilde{k}_p(x)$ the off-diagonal entries (coupling elements) of $\partial h(x)/\partial x$. Note that the main diagonal of $\tilde{k}_p(x)$ is zero. Thus, the j-th line of $\tilde{k}_p(x)$ corresponds to the coupling of channels u_i $(i \neq j)$ in the channel u_j.

Recall that, from Assumption 17, the inverse matrix $k_p^{-1}(x)$ is well defined outside the Δ-neighborhood. Therefore, one can rewrite (4.6b) as

$$\dot{y}(t) = k_p(x) \left[u(t) + k_p^{-1}(x)\tilde{k}_p(x)u(t) \right] , \quad \forall x_j \notin \mathcal{D}_{\Delta_j} \tag{4.14}$$

Note that the matrix $k_p^{-1}\tilde{k}_p$ also has a zero main diagonal.

As a generalization with respect to (Peixoto and Oliveira 2012), we can cover input-output mappings $h(\cdot)$ such that $k_p^{-1}\tilde{k}_p$ is triangular, so that no hard restriction is assumed for the class of triangular input-output mapping $h(\cdot)$. Moreover, for the non-triangular case, we consider the so-called "triangular dominant case", which represents a remarkably relaxed assumption in contrast to (Peixoto and Oliveira 2012), where diagonal dominance was assumed. This is the approach we followed in (Peixoto et al. 2020).

First, we consider the change of variables $u = S(x)v$ with $S(x)$ a pre-multiplier state-dependent diagonal matrix with continuous positive functions $s_j : \mathbb{R}^n \mapsto \mathbb{R}^+$ to be defined later on. Thus, one can rewrite (4.14) as

$$\dot{y}(t) = k_p(x)S(x) \left[v(t) + S^{-1}(x)k_p^{-1}(x)\tilde{k}_p(x)S(x)v(t) \right] \quad \forall x \notin \mathcal{D}_{\Delta_j} \tag{4.15}$$

To avoid the loss of controllability near the Nash equilibrium, as in Assumption 17, the following assumption is taken as granted, where $k_{pj} = \partial h_j/\partial x_j$.

Assumption 18. *Outside the Δ-neighborhood, there exists a known constant $L_j(\Delta) > 0$ such that*

$$L_j \leq |k_{pj}(x)s_j(x)| , \quad \forall x \notin \mathcal{D}_{\Delta_j} \tag{4.16}$$

with $j = 1, \dots, n$, where Δ can be made arbitrary small by allowing a smaller $L_j(\Delta)$.

Assumption 18 guarantees a lower-bound for the diagonal elements of the HFG matrix $(k_p + \tilde{k}_p)S$ in the v-coordinate system. Now, we consider the following as-

sumption regarding the class of input-output mappings $h(\cdot)$ for non-triangular cases:

Assumption 19 (Bounded Matched Disturbances Pre-Multiplier)**.** *There exists a known state-dependent diagonal matrix $S(x) \in \mathbb{R}^{n\times n}$, such that*

$$\left|S^{-1}(x)k_p^{-1}(x)\tilde{k}_p(x)S(x)\right| \leq P \ , \quad \forall x \notin \mathcal{D}_{\Delta_j} \tag{4.17}$$

where P is a known constant matrix with non-negative elements. The inequality is understood element-by-element and all entries of the main diagonal of $S(x)$ are continuous positive functions $s_j : \mathbb{R}^n \mapsto \mathbb{R}^+$.

Note that, since $S^{-1}k_p^{-1}\tilde{k}_pS$ has zero main diagonal, then P also has a zero main diagonal. Assumption 19 allows that, but for scaling factors $s_i(x)$, every off-diagonal entry of the Jacobian matrix is dominated by the corresponding diagonal entry $k_{pj} = \frac{\partial h_j}{\partial x_j}$, so that one can write:

$$\left|\frac{\partial h_j}{\partial x_i}\right| \leq p_{ij}\,|k_{pj}(x)|\,\frac{s_j(x)}{s_i(x)}\,, \ \forall i\,, \ \forall x \notin \mathcal{D}_{\Delta_j} \tag{4.18}$$

where p_{ij} denotes the ij-th components of matrix P in Assumption 19. Assumption 19 is a little restrictive, but it still allows to encompass a large class of functions $h(\cdot)$, for instance any polynomial function, functions that admit a local polynomial approximation, and functions with a triangular Jacobian matrix.

Remark 7. *If each component of the objective function is polynomial (at least locally) of order r_j, and $\partial h/\partial x$ is upper-triangular, a choice of $s_j(x)$, other than the trivial choice $s_j(x) = 1$, that satisfies (4.18) is*

$$s_j(x) = s_{j+1}(x)\max\left(\frac{a_j}{|x_j - x_{dj}|^{r_j-1} + b_j}\,,\ 1\right) \tag{4.19a}$$

$$s_n(x) = 1 \tag{4.19b}$$

where $x_d = \begin{bmatrix} x_{d1} & \dots & x_{dn} \end{bmatrix}^\mathsf{T}$ is some rough estimate of the optimizer x^*, e.g. the initial guess $x(0)$, and $a_j, b_j > 0$. If $\partial h/\partial x$ is lower-triangular, then

$$s_j(x) = s_{j-1}(x)\max\left(\frac{a_j}{|x_j - x_{dj}|^{r_j-1} + b_j}\,,\ 1\right) \tag{4.20a}$$

$$s_0(x) = 1 \tag{4.20b}$$

Finally, we introduce one other condition which characterizes the coupling between input and output channels of the process, relaxing the previous diagonal

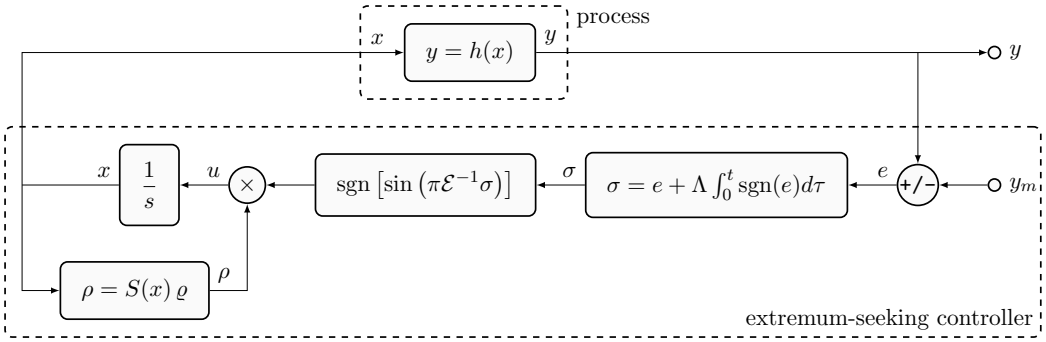


Figure 4.1: Illustrative block diagram of the proposed multi-variable sliding-mode based extremum-seeking controller.

dominance property assumed in (Peixoto and Oliveira 2012).

Assumption 20 (Upper-Boundedness of (4.17))**.** *In Assumption 19, the constant matrix P is "triangular dominant", such that $\forall i, j$ with $i < j$ (or $\forall i, j$ with $i > j$), the entries $p_{ij} \in \mathbb{R}_+$ satisfy*

$$p_{ij} < 1 \tag{4.21}$$

with sufficiently small values.

In (Frihauf et al. 2011; Peixoto and Oliveira 2012), it was considered that all entries p_{ij} $\forall i \neq j$ were sufficiently small. This is the mentioned "Jacobian diagonal dominance" condition. Here, in contrast, we consider only this constraint $\forall i < j$ $(i > j)$, named "triangular dominance condition", which requires **only** p_{ij} to be sufficiently small for $i < j$ $(i > j)$. Each element **below (above) the main diagonal of the Jacobian matrix** must be sufficiently small while no hard restriction is imposed on each element above (below) the main diagonal.

In what follows, for simplicity and without loss of generality, the triangular dominant case is considered only when the elements below $(i < j)$ the Jacobian main diagonal are smaller than one.

4.2.2 Extremum-Seeking Controller

In this section, we describe our sliding-mode extremum-seeking-controller for static nonlinear plants (Peixoto et al. 2020), which we later illustrate by considering the signal power control of Raman optical amplifiers modeled as a static I/O mapping, as well as numerical simulation examples.

The developed output-feedback based ESC law with *periodic switching function* is written as

$$u(t) = \rho(t)\,\mathrm{sign}\left(\sin\left[\pi \mathcal{T}^{-1}\sigma(t)\right]\right) \tag{4.22a}$$

$$\rho(t) = S(x)\varrho(t) \tag{4.22b}$$

with $\rho(t) = \text{diag}\left(\rho_1(t)\,, \quad \ldots\,, \quad \rho_n(t) \right)$ being the *modulation function* (continuous in t) to be defined later on by designing $\varrho(t)$,

$$S(x) = \text{diag}\left(s_1(x)\,, \quad \ldots\,, \quad s_n(x) \right) \tag{4.23}$$

with $s_i(x)$ satisfying Assumptions 18 and 19,

$$\sigma(t) = e(t) + \Lambda \int_0^t \text{sign}[e(\tau)]d\tau \tag{4.24}$$

and $\Lambda, \mathcal{T} > 0$ being appropriate constant matrices in the form $\Lambda = \text{diag}\left(\lambda_1, \ \ldots, \ \lambda_n \right)$ and $\mathcal{T} = \text{diag}\left(T_1\,, \ \ldots\,, \ T_n \right)$.

The error signal e is given by

$$e(t) = y(t) - y_m(t) \tag{4.25}$$

where $y_m : \bar{\mathbb{R}}_+ \mapsto \mathbb{R}^n$ is a simple ramp time function. The structure of the proposed multi-variable sliding-mode based ESC algorithm is shown in fig. 4.1.

For analysis purposes, such a vector function is produced by the reference model

$$\dot{y}_m = -K_m\,, \quad y_m(0) = y_{m0} \tag{4.26}$$

with $K_m \in \mathbb{R}^n > 0$, $y_{m0} \in \mathbb{R}^n$ a design constant vector, and the elements K_{mj} $(j = 1, \ldots, n)$ of $K_m = \left[\begin{array}{ccc} K_{m1} & \ldots & K_{mn} \end{array} \right]^\mathsf{T}$ positive constants (decreasing ramps). The modulation function ρ is designed to make $y(t)$ track $y_m(t)$ as long as possible. In this sense, y is brought to the neighborhood of the extremum $y^* = h(x^*)$. One can saturate the model output at a rough known norm bound of y^* in order to avoid an unbounded reference signal $y_m(t)$ in the controller, without affecting the overall control performance. The decreasing reference signal dictates the speed of convergence and brings a smooth transition to the extremum.

Basically, we need to construct ρ such that a sliding-mode $\dot{\sigma}_j = 0$ $(j = 1, \ldots, n)$ appears in finite time on one of the manifolds $\sigma_j = kT_j$, for some integer k. From (4.24),

$$\dot{\sigma} = \dot{e} + \Lambda\,\text{sign}(e) = 0 \tag{4.27}$$

where $\text{sign}(e) = \left[\begin{array}{ccc} \text{sign}(e_1) & \ldots & \text{sign}(e_n) \end{array} \right]^\mathsf{T}$. It is possible to conclude that if e tends to zero, then $y = h(x)$ tracks y_m. In other words, each y_j is driven to the neighborhood of a minimum y_j^*, which is a function of x_i $(\forall i \neq j)$, remaining close

to it as long as y_j is kept away from a small neighborhood of y_j^*, since $\partial h_j/\partial x_j$ is away from zero, according to Assumptions 16 and 17.

On the other hand, once y_j enters the neighborhood of y_j^*, the term $\partial h_j/\partial x_j$ becomes too small and the controllability for the j-th channel is lost. Consequently, tracking would stop. However, the neighborhood of the optimum would have been achieved. Thus, the control strategy would guarantee at least that y_j would remain close to y_j^* thereafter.

The control law given in (4.22)–(4.26) can be summarized in table 4.1. The meaning of each parameter is the same as explained in section 2.2.1. The more intricate design part is that of the modulation function $\rho(t)$, which needed to assure the closed-loop stability properties, as it will be shown in sections 4.2.3 and 4.2.4. However, from a practical point of view, the design can be simplified, according to the remark below.

Table 4.1: Control variables and parameters of the sliding-mode extremum-seeking controller for static nonlinear plants, that is capable of finding a Nash equilibrium.

$u(t) = \rho(t)\,\text{sign}\left(\sin\left[\pi \mathcal{T}^{-1}\sigma(t)\right]\right)$	$\rho(t) = S(x)\varrho(t),\ \ \dot{x}(t) = u(t)$
$\sigma(t) = e(t) + \Lambda \int_0^t \text{sign}[e(\tau)]d\tau$	$e(t) = y(t) - y_m(t),\ \ x(0) = x_0$
$\Lambda = \text{diag}\,(\lambda_1, \ldots, \lambda_n)$	$\mathcal{T} = \text{diag}\,(T_1, \ldots, T_n)$
$K_m = [K_{m1}, \ldots, K_{mn}]^\mathsf{T}$	$\dot{y}_m(t) = K_m,\ \ y_m(0) = y_{m0}$

Remark 8 (Practical Implementation Aspects)**.** *It must be highlighted that the design of the pre-compensation matrix $S(x)$ and the modulation function ϱ is intricate but is needed mainly to obtain the stability result of the thesis. On the other hand, the algorithm implementation can be straightforward. In practice, for weakly coupled systems, there is usually no need to compute or estimate matrix $S(x)$, and a constant modulation gain (instead of a modulation function) can be used. By doing so, only two integrators and trivial mathematical operations are needed to compute the control law, see (4.22) to (4.26) or table 4.1. In such a case, the only drawback of not computing the full modulation function (where the main complexity of the algorithm is incorporated) is that local stability results are achieved, but these are often sufficient in practical applications.*

4.2.3 Modulation Function Design

Matched Input Disturbance

From equations (4.25), (4.26), (4.27) and the y-dynamics in (4.6), one can obtain the σ-dynamics as

$$\dot{\sigma}(t) = k_p(x)[u(t) + d_\sigma(x, e, \dot{y}_m)] \tag{4.28}$$

where the relationship $\partial h(x)/\partial x = k_p(x) + \tilde{k}_p(x)$ is used, with k_p in (4.13), and the *matched input disturbance* is defined by

$$k_p(x)\, d_\sigma(x, u, e, \dot{y}_m) = \left[\tilde{k}_p(x)u(t) - \dot{y}_m + \Lambda\, \mathrm{sign}(e)\right] \tag{4.29}$$

Furthermore, by expanding the components of (4.28), one obtains n equalities

$$\dot{\sigma}_j(t) = k_{pj}(x)\,[u_j(t) + d_{\sigma j}(x, \bar{u}_j, e_j, \dot{y}_{mj})] \tag{4.30}$$

where $\bar{u}_j$ denotes all components of u, but for the j-th component. Expanding the matched disturbance component-wise,

$$k_{pj} d_{\sigma j} = \sum_{i \neq j} \left(\frac{\partial h_j}{\partial x_i}\right) u_i - \dot{y}_{mj} + \lambda_j\, \mathrm{sign}(e_j)\,,\ \forall x \tag{4.31}$$

where the dependency of each function on its variables is omitted for simplicity. It must be highlighted that (4.31) holds $\forall x$.

Matched Input Disturbance Norm Bound

From (4.31), it follows that

$$d_{\sigma j} = \sum_{i \neq j} \left(\frac{\partial h_j/\partial x_i}{k_{pj}}\right) u_i - \frac{\dot{y}_{mj}}{k_{pj}} + \frac{\lambda_j}{k_{pj}}\, \mathrm{sign}(e_j)\,, \quad x \notin \mathcal{D}_{\Delta_j} \tag{4.32}$$

Then, by noting that $|u_i| \leq \rho_i$ $(i = 1, \ldots, n)$ and considering Assumption 18, one can obtain the following upper-bound for $|d_{\sigma j}|$:

$$|d_{\sigma j}| \leq \sum_{i \neq j} p_{ij} \frac{\rho_i s_j}{s_i} + \frac{1}{k_{pj}} \left(|\dot{y}_{mj}| + \lambda_j\right)\,, \quad x \notin \mathcal{D}_{\Delta_j} \tag{4.33}$$

where p_{ij} denotes the ij-th components of matrix P in Assumption 19. Since $\rho_i = s_i \varrho_i$, according to (4.22), then one can write the inequality

$$\frac{|d_{\sigma j}|}{s_j} \leq \sum_{i \neq j} p_{ij} \varrho_i + \frac{1}{k_{pj} s_j} \left(|\dot{y}_{mj}| + \lambda_j \right) , \quad x \notin \mathcal{D}_{\Delta_j} \tag{4.34}$$

Recall that, according to Assumption 19, for the triangular case, it is possible to set $S = I$ or use the suggestion from Remark 7. For the non-triangular case, a pre-multiplier matrix $S(x)$, function of the states, is usually needed. In this formulation where the process is a static output map, the states are available for the controller since they are the integral of the control input.

Modulation Function Design

The modulation function ρ is designed so that a sliding-mode is obtained in finite time on one of the manifolds $\sigma_j = kT_j$, for some integer k.

In the convergence analysis, it is shown that one sufficient condition to assure the realization of the σ-sliding-modes is given by

$$\rho_j = \varrho_j s_j \geq |d_{\sigma j}| + s_j \, \delta_j , \quad j = 1, \ldots, n \tag{4.35}$$

where $\delta_j > 0$ is an arbitrary constant. Below, one possible implementation of the modulation function such that (4.35) holds is presented. In contrast to (Peixoto and Oliveira 2012), the modulation functions ρ_j assume different values.

From (4.34), one can choose ϱ_j to satisfy

$$\varrho_j \geq \sum_{i \neq j} p_{ij} \varrho_i + \frac{1}{L_j} \left(|\dot{y}_{mj}| + \lambda_j + \delta_j \right) \tag{4.36}$$

with $\delta_j > 0$, which assures that (4.35) holds using the inequality $L_j \leq |k_{pj} s_j|$ from Assumption 18. The inequalities in (4.36) can be rewritten in the compact form

$$(I - P) x_\varrho \geq \mathcal{B}(\dot{y}_m) \tag{4.37}$$

which is understood element-wise and must hold for $x \notin \mathcal{D}_{\Delta_j}$, where $\mathcal{B} = \begin{bmatrix} \mathcal{B}_1 & \ldots & \mathcal{B}_n \end{bmatrix}^\mathsf{T}$, with $\mathcal{B}_j(\dot{y}_{mj}) = \frac{1}{L_j} \left(|\dot{y}_{mj}| + \lambda_j + \delta_j \right)$, and $x_\varrho = \begin{bmatrix} \varrho_1 & \ldots & \varrho_n \end{bmatrix}^\mathsf{T}$. Matrix P comes from Assumption 19.

Defining the objective function $C^\mathsf{T} x_\varrho$, where $C \in \mathbb{R}^n_+$ is a weight vector, the modulation function is obtained by solving (e.g. via linear programing) the following

optimization problem:

$$\min_{x_\varrho} \quad C^\mathsf{T} x_\varrho \,, \tag{4.38a}$$

$$\text{s.t.} \quad (I - P)x_\varrho \geq \mathcal{B}(\dot{y}_m) \,, \tag{4.38b}$$

$$x_\varrho > 0$$

One can verify that Assumption 20 is a sufficient condition for the existence of a solution. To avoid solving this problem for all possible values of $\dot{y}_m$, one might use an upper-bound for $|\dot{y}_m|$ when computing $\mathcal{B}$.

4.2.4 Stability and Convergence Analysis

In this section, the stability and convergence results of the real-time multi-variable extremum-seeking controller are carried out by showing that the existence of ideal sliding-modes guarantees the attractiveness of the Δ-neighborhoods, which are reached in finite time.

Realization of Ideal Sliding-Modes

The next proposition summarizes the results regarding the existence of ideal sliding-modes. This proposition is very similar to Proposition 1, and this similarity will be used to shorten the proof.

Proposition 2 (Occurence of Sliding-Modes)**.** *Consider the system (4.6), with control law (4.22) and modulation function ρ in (4.22) satisfying (4.38), while $x \notin \mathcal{D}_{\Delta_j}$. Let $t_0 \geq 0$ be a time instant such that σ_j is not in sliding motion at $t = t_0$, i.e., $\sigma_j(t_0) \neq k_{0j} T_j$ where $k_{0j} = \left\lfloor \frac{\sigma_j(t_0)}{T_j} \right\rfloor$. Then, if $x \notin \mathcal{D}_{\Delta_j}$, $\forall t \geq t_0$, a sliding-mode on $\sigma_j = k^* T_j$, with $k_j^* = k_{0j} - 1$, $k_j^* \in \{k_{0j} - 1, k_{0j} + 1\}$, is achieved in some finite time $t_j \geq t_0$ which satisfies $t_0 \leq t_j < t_0 + \mathcal{O}(T_j/\delta_j)$ with $\delta_j > 0$ in (4.35). Moreover, no finite-time escape occurs in the system signals.*

Proof. First, note that the inequality

$$k_j \pi \leq \frac{\pi}{T_j} \sigma_j < (k_j + 1)\pi \,, \qquad k_j = \left\lfloor \frac{\sigma_j}{T_j} \right\rfloor \tag{4.39}$$

holds $\forall t \geq 0$, where k is time-dependent, and $j = 1, \ldots, n$. For any real s, $\lfloor s \rfloor$ denotes the greatest integer lower or equal to s. Moreover, define

$$k_{0j} = \lfloor \sigma_j(t_0)/T_j \rfloor \tag{4.40a}$$

$$k_j^* = k_{0j} + \frac{\operatorname{sign}(k_{pj}) + (-1)^{k_{0j}}}{2} \tag{4.40b}$$

$$\alpha_j = \left(\frac{\sigma_j(t)}{T_j} - k_j^*\right)\pi \tag{4.40c}$$

and, since $\sin\left(\frac{\pi}{T_j}\sigma_j\right) \geq 0$ for an even k and $\sin\left(\frac{\pi}{T_j}\sigma_j\right) \leq 0$ for an odd k, rewrite the control signal as

$$u_j(t) = \rho_j(t)\,\text{sign}\left(\sin\left[\frac{\pi}{T_j}\sigma_j(t)\right]\right) = \rho_j(-1)^{k_j} \tag{4.41}$$

Comparing (4.40) and (4.72) to (2.21) and (2.18) from the proof of Proposition 1, we define the candidate Lyapunov function

$$V(\alpha_j) = \frac{T_j}{\pi}\left|\alpha_j\right| \quad , \quad \dot{V}(\alpha_j) = \frac{T_j}{\pi}\,\text{sign}(\alpha_j)\dot{\alpha}_j \tag{4.42}$$

and conclude that

$$\begin{aligned}\dot{V}(\alpha_j) &= |k_{pj}|\left[-\rho_j + d_{\sigma j}\,\text{sign}(k_{pj})\,\text{sign}(\alpha_j)\right] \\ &\leq |k_{pj}|\left(-\rho_j + |d_{\sigma j}|\right)\end{aligned}$$

Considering the modulation function design from the solution of (4.38) that guarantees that (4.35) holds, it follows that

$$\dot{V}(\alpha_j) \leq |k_{pj}|\left(-\frac{s_j\delta_j}{L_j} - |d_{\sigma j}| + |d_{\sigma j}|\right) = -\left|k_{pj}\right|\frac{s_j\delta_j}{L_j}$$

and, using the inequality $L_j \leq |k_{pj}s_j|$ from Assumption 18,

$$\dot{V}(\alpha_j) \leq -\delta_j < 0 \tag{4.43}$$

Equation (4.42) confirms that $V(\alpha_j)$ from (4.42) is a Lyapunov function. Nonetheless, before stating that a sliding-mode occurs, we first prove that no finite-time escape occurs in the closed-loop signals.

If $x_j \notin \mathcal{D}_{\Delta_j}$, $\forall t \geq t_0$, then one can conclude that α_j is uniformly bounded in $\forall t \geq t_0$ and cannot escape in finite time. Subsequently, σ_j and the objective function output y_j do not escape in finite time. Otherwise, if $x_j \in \mathcal{D}_{\Delta_j}$ for some period of time, one has that x_j is uniformly bounded, leading to the same conclusion that y_j does not escape in finite time due to the continuity regarding the output function $h(\cdot)$. Since the above reasoning holds for $j = 1, \ldots, n$, then the full plant output y does not escape in finite time. Hence, from the unbounded observability assumption (Assumption 15), one can assure that all closed-loop signals cannot escape in finite time.

Finally, from (4.43), one can conclude that an ideal sliding-mode occurs on the

manifold $\alpha_j = 0$, or equivalently, on the manifold $\sigma_j = k_j^* T_j$. From the Lyapunov function definition, this sliding-mode occurs after a finite time t_j, which belongs to the interval

$$t_0 \leq t_j \leq t_0 + \frac{T_j \left|\alpha_j(t_0)\right|}{\pi \delta_j} \leq t_0 + \frac{T_j}{\delta_j} \tag{4.44}$$

■

Convergence Analysis

Theorem 6 (Globally Convergent ESC)**.** *Consider the system* (4.6)*, with control law* (4.22) *and modulation function* ρ *in* (4.22) *satisfying* (4.38)*, while* $x \notin \mathcal{D}_{\Delta_j}$*. Assume that all assumptions regarding the objective function hold, i.e. Assumptions 16 to 19. Then: (i) all* Δ_j*-neighborhoods* $\mathcal{D}_{\Delta_j}$ *are globally attractive and are achieved in finite time, and (ii) for a sufficiently small design constant* L_j *from Assumption 18, the oscillations around the minimum* y^* *of* y *can be made of order* $\mathcal{O}(\max_j T_j)$*, with* $\mathcal{T}$ *given in* (4.22)*. In addition, all closed-loop signals remain uniformly bounded, except for* σ *which is only the argument of a sine function in* (4.22)*.*

Proof. In what follows, consider $j \in \{1, \ldots, n\}$. The general idea of the proof is based on two properties. First, no finite-time escape occurs in the system signals according to Proposition 2. Second, outside the neighborhood $\mathcal{D}_{\Delta_j}$, the partial derivative of the output function h_j w.r.t. x_j do not vanish, that is $k_{pj} = \partial h_j / \partial x_j \neq 0\,, \forall x \notin \mathcal{D}_{\Delta_j}$. While x_j stays outside the neighborhood $\mathcal{D}_{\Delta_j}$ sliding-mode on $\sigma_j = k^* T_j$ assures that x is driven back to the neighborhood $\mathcal{D}_{\Delta_j}$. Now, we proceed to the detailed proof of the properties **(i)** and **(ii)** of Theorem 6.

(i) Attractiveness of $\mathcal{D}_{\Delta_j}$: First, recall that finite-time escape is avoided for the closed-loop system signals. Then, from Proposition 2, there exists a finite time t_j such that $\sigma_j = k_j^* T_j$, $\forall t \in [t_j, t_j^*)$, where $t_j^* > t_j$ is a time for which $x \in \mathcal{D}_{\Delta_j}$. Then, $\dot{\sigma}_j = 0$ and, from (4.24), one has that $\dot{e}_j = -\lambda_j \operatorname{sign}(e_j)$ or $e_j \dot{e}_j = -\lambda_j |e_j| \leq 0$, $\forall t \in [t_j, t_j^*)$. Thus, the error $e_j = y_j - y_{mj}$ tends to zero. Moreover, one of the following properties hold at any time $t > t_0$:

- $y_j < y_{mj} \iff \operatorname{sign}(e_j) < 0$

  Meaning that y_j is below a function y_{mj} of time which decreases and, eventually, crosses the minimum value y_j^*.

- $y_j > y_{mj} \iff \operatorname{sign}(e_j) > 0$

  Meaning that $\dot{y}_j(t) = -K_{mj} - \lambda_j \operatorname{sign}(e_j) = -(K_{mj} + \lambda_j) < 0$, thus, the objective function value y_j decreases at a rate $K_{mj} + \lambda_j$ when it is above the model reference y_{mj}.

Therefore, y_j must decrease, reaching the neighborhood $h_j\left(\mathcal{D}_{\Delta_j}\right)$ in a finite time t_j^*, which satisfies

$$t_j^* \leq t_j + \frac{\left|y_{mj}(t_0) - y_j^*\right|}{K_{mj} + \lambda_j} \tag{4.45}$$

where y_j^* denotes the j-th component of the objective function global minimum y^*, which exists and is unique from Assumption 16. Therefore, from the definition of $\mathcal{D}_{\Delta_j}$ (Definition 9) and the continuity assumption on the output function, x_j reaches $\mathcal{D}_{\Delta_j}$ no longer than $t = t_j^*$.

(ii) Oscillations of Order $\mathcal{O}(\max_j T_j)$:

It is proved that x_j reaches $\mathcal{D}_{\Delta_j}$, nonetheless, it might remain inside this region or oscillate above it, and, similarly, y_j with respect to the corresponding neighborhood $h_j(\mathcal{D}_{\Delta_j})$, for $t \geq t_j^*$.

These oscillations come from the loss of control strength as $k_{pj} \to 0$ whenever the relation $L_j \leq |s_j(x)k_{pj}(x)|$ is violated or are due to the recurrent changes in the sign of k_j at the extremum point. During these oscillations, σ_j can go from one sliding manifold $\sigma_j = k_j^* T_j$ (k_j^* even) to another (k_j^* odd). In such case, y_j could start oscillating around y_j^* with increasing amplitude. We show now that this does not happen.

Assume that x_j reaches the frontier of $\mathcal{D}_{\Delta_j}$ (from the inside) at some time $t_j' > t_j^*$ and σ_j is not in sliding-mode at $t = t_j'$. The time instant t_j^* is defined in the first part of the proof. From Proposition 2, σ_j reaches a sliding-mode again after a finite time $t_j'' \leq t_j' + T_j \left|\alpha_j(t_j')\right| / \pi\delta_j$. Following the same procedure as in Theorem 3, we show that, during this time, the output y_j distances itself from $h_j\left(\mathcal{D}_{\Delta_j}\right)$ by an amount of order $\mathcal{O}(\max_j T_j)$.

Recall the sliding-variable σ definition (4.24). From this equation and the inequality for $\Delta t_j = t_j'' - t_j' \leq T_j \left|\alpha_j(t_j')\right| / \pi\delta_j$,

$$\begin{aligned}
\sigma_j(t_j'') - \sigma_j(t_j') &= e_j(t_j'') - e_j(t_j') + \int_{t_j'}^{t_j''} \lambda_j \operatorname{sign}(e_j)\, d\tau \\
\left|e_j(t_j'') - e_j(t_j')\right| &\leq \left|\sigma_j(t_j'') - \sigma_j(t_j')\right| + \int_{t_j'}^{t_j''} \left|\lambda_j \operatorname{sign}(e_j)\right| d\tau \\
\left|e_j(t_j'') - e_j(t_j')\right| &\leq \left|\sigma_j(t_j'') - \sigma_j(t_j')\right| + \lambda_j \int_{t_j'}^{t_j''} d\tau \\
\left|e_j(t_j'') - e_j(t_j')\right| &\leq T_j + \lambda_j \Delta t_j \leq \left(1 + \frac{\lambda_j \left|\alpha_j(t_j')\right|}{\pi\delta_j}\right) T_j \\
\left|e_j(t_j'') - e_j(t_j')\right| &\leq (1 + \lambda_j/\delta_j)\, T_j \leq \mathcal{O}(T_j)
\end{aligned} \tag{4.46}$$

where $\left|\sigma(t_j'') - \sigma(t_j')\right| \leq T_j$ and $\left|\alpha_j(t_j')\right| \leq \pi$ were used. From the output error

definition (4.25),

$$\begin{aligned}\left|y(t_j'') - y(t_j')\right| &\leq \left|e_j(t_j'') - e_j(t_j')\right| + \left|y_{mj}(t_j'') - y_{mj}(t_j')\right| \\ &\in \left\{ (1 + \lambda_j/\delta_j)\, T_j\,,\ [1 + (1 + K_{mj})\, \lambda_j/\delta_j]\, T_j \right\} \end{aligned} \tag{4.47}$$

where the term K_{mj} is considered if the reference model is not saturated at some lower-bound below y_j^*. Otherwise, after a time $t_j''' : y_{mj}(t) < y_j^*(t)$ and $\dot{y}_{mj}(t) = 0$, for $t \geq t_j'''$, the lower-bound $(1 + \lambda_j/\delta_j)\, T_j$ is considered. Either way, it follows that the oscillations above the region $h_j\left(\mathcal{D}_{\Delta_j}\right)$ are of order $\mathcal{O}(T_j)$, for every $j \in \{1, \ldots, n\}$.

Thus, for each channel, the oscillations above the $h_j\left(\mathcal{D}_{\Delta_j}\right)$ regions are of order $\mathcal{O}(T_j)$, and, in the worst case, the oscillations of y above y^* are of order $\mathcal{O}(\max_j T_j)$. ■

4.2.5 Numerical Simulation Example

As an example, consider

$$y = h(x) = \frac{1}{2}\begin{bmatrix}(x_1 - x_2 + 1)^2 \\ (x_2 - 1)^2\end{bmatrix} \implies x^*(x_2) = \begin{bmatrix}x_2 - 1 \\ 1\end{bmatrix} \text{ and } y^*(x^*) = \begin{bmatrix}0 \\ 0\end{bmatrix} \tag{4.48}$$

With the proposed control strategy, we set $\dot{x}(t) = u(t)$, such that the high-frequency gain (with respect to u) is

$$\frac{\partial h}{\partial x}(x) = \begin{bmatrix}(x_1 - x_2 + 1) & -(x_1 - x_2 + 1) \\ 0 & (x_2 - 1)\end{bmatrix} \tag{4.49}$$

Furthermore, note that Assumption 16 is satisfied and $\Theta^* = \{(0, 1)\}$. In this example, the HFG matrix is triangular, so we set $s_1(x) = s_2(x) = 1$. Another possibility would be to choose $s_j(x)$ according to Remark 7.

To achieve minimization at a rate of 4 units per second, let $\lambda_1 = \lambda_2 = 4$ and, for simplicity, $\delta_j = \lambda_j$, $K_{mj} = 0$, and $y_{mj}(t) = 0$. To keep oscillations of y_j above the regions $h_j(\mathcal{D}_{\Delta_j})$ bounded by (at most) 0.7 for y_1 and 0.5 for y_2, from (4.47), let $T_1 = 0.35$ and $T_2 = 0.25$. The only control parameters missing in the design are ρ_1 and ρ_2. Instead of assuming we now the bounds p_{ij} and L_j from Assumptions 18 and 19, we fix

$$\rho = \begin{bmatrix}20 \\ 11\end{bmatrix} \tag{4.50}$$

which, from (4.38), should guarantee convergence as long as

$$\begin{bmatrix} 20 - 11p_{12} \\ -20p_{21} + 11 \end{bmatrix} \geq 8 \begin{bmatrix} 1/L1 \\ 1/L2 \end{bmatrix} \tag{4.51}$$

Note that, as long as $p_{12} < 20/11 \approx 1.8$ and $p_{21} < 11/20 = 0.55$, there will always be a solution for L_1 and L_2. In fact, since we know the objective function, we also know that $p_{12} = 1$, $p_{21} = 0$, $L_1 = 8/9 \approx 0.9$, and $L_2 = 8/11 \approx 0.7$. Thus, we also know in advance that $\Delta_1 \approx 1.8$ and $\Delta_2 \approx 1.4$. These values for Δ_j imply that controllability is lost when $(y_1 - y_1^*) \leq 0.5(0.9)^2 \approx 0.4$ or $(y_2 - y_2^*) \leq 0.5(0.7)^2 \approx 0.25$. Therefore, based on the values the tuples $(\lambda_j, \delta_j, K_{mj}, T_j)$, after the ESC algorithm converges, the oscillations above the minimum must stay below $(y_1 - y_1^*) \leq 0.4 + 0.7 = 1.1$ and $(y_2 - y_2^*) \leq 0.25 + 0.5 = 0.75$.

The simulation outputs are shown in fig. 4.2, where the controllers are implemented in discrete-time with sampling periods of 1 ms. To the right, we show the results with the sliding-mode controller exactly as proposed in this section. To the left, we change the sign functions by continuous approximations, just as it was done in section 2.2.3, using $\epsilon = 0.2$ for the sigmoid in the definition of $u(t)$ and $\bar{\epsilon} = 1$ for the sigmoid in the definition of $\sigma(t)$. In this example we see that, with the discrete-time implementation, the controller using a continuous approximation outperforms the one with discontinuous control action in terms of chattering, but presents a slightly bigger maximum error after convergence.

To demonstrate the robustness of the proposed strategy, first-order input and output filters,

$$H_{\text{in}} = \frac{s}{0.03s + 1} \quad \text{and} \quad H_{\text{out}} = \frac{s}{0.01s + 1} \tag{4.52}$$

are added and another simulation is conducted (fig. 4.3). Using the continuous switching function via saturated gain, there is a slight increase in control effort, but the final performance is similar to the one observed in the static case. For the discontinuous control implementation and the same control parameters, convergence was not achieved after adding the filters.

4.2.6 Application to Raman Optical Amplifiers

As a more interesting and practical example, as published in (Peixoto et al. 2020), the controller is applied to the optimization of the output signal power spectrum of Raman optical amplifiers. A Raman amplifier possesses a set of n_p pump lasers (actuators), which emit electromagnetic waves inside an optical fiber (of length L), centered at a collection of discrete, non-overlapping wavelengths

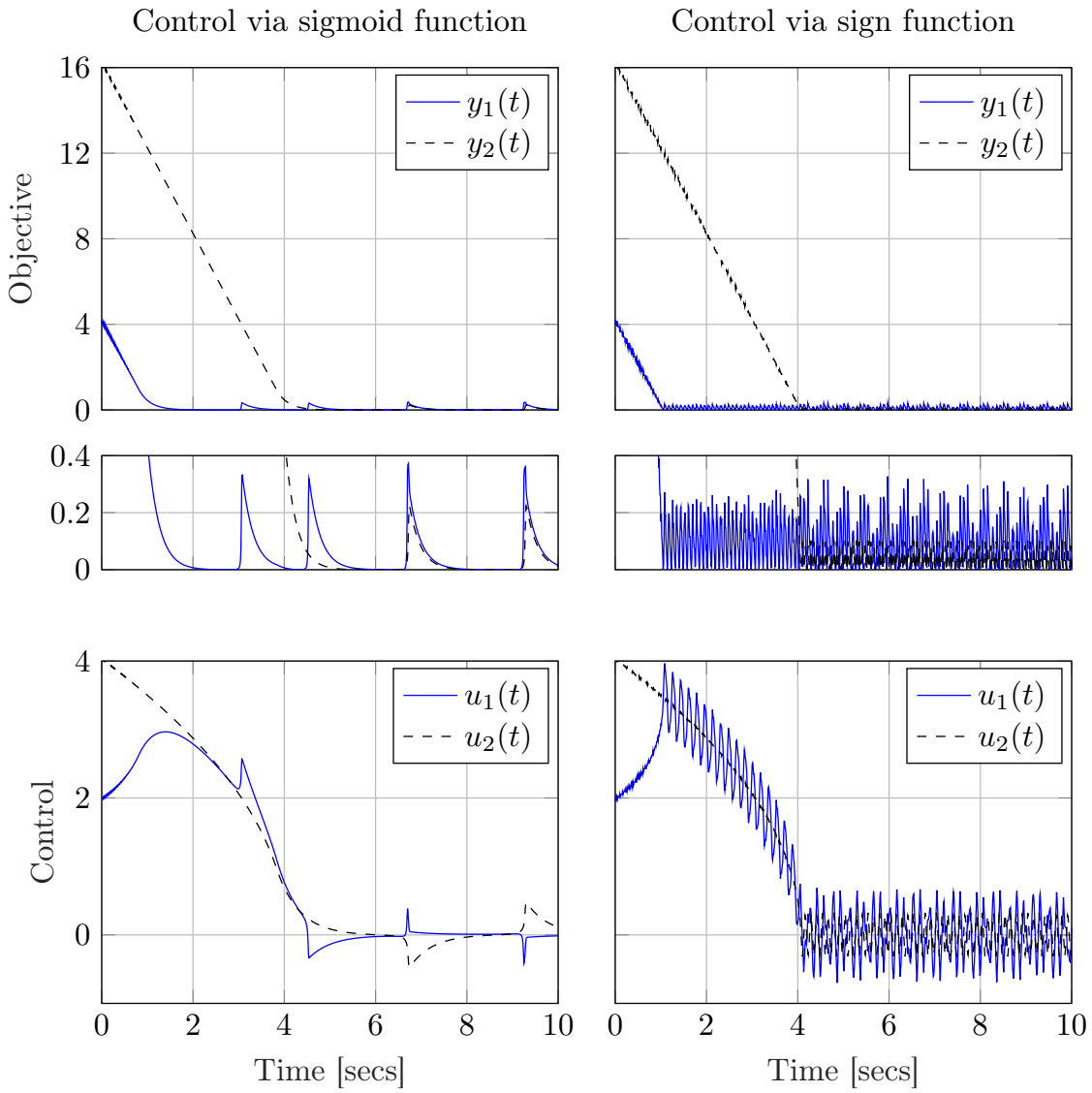


Figure 4.2: Simulation results for the MIMO ESC example of section 4.2.5. To the right, the implementation with discontinuous control action. To the left, the implementation using a continuous approximation of the sign function.

$\Lambda_p = \{\lambda_i : \ i = 1, 2, \ldots, n_p\}$. Through the optical fiber, data propagates on n_s discrete and non-overlapping wavelengths, from the fiber upstream to the fiber downstream. Each data signal is subject to wavelength-dependent loss, crosstalk and propagation delay (Headley and Agrawal 2005; Palais 1998; Kidorf et al. 1999).

Let $P_s(t, z) \in \mathbb{R}^{n_s}$ and $P_p(t, z) \in \mathbb{R}^{n_p}$ be vectors corresponding to the power (in W) of a data signal or a pump signal, respectively, propagating at time t and distance z (measured along the fiber) from the upstream end of the span ($z = 0$) to the downstream ($z = L$). At the upstream, it is desired to regulate the data signals power, which travel across the fiber with discrete, and non-overlapping wavelengths $\Lambda_s = \{\lambda_i : \ i = n_p + 1, n_p + 2, \ldots, N\}$, with $N = n_p + n_s$. In general, such a regulation process is implemented in open-loop. The pump powers at the end of the span

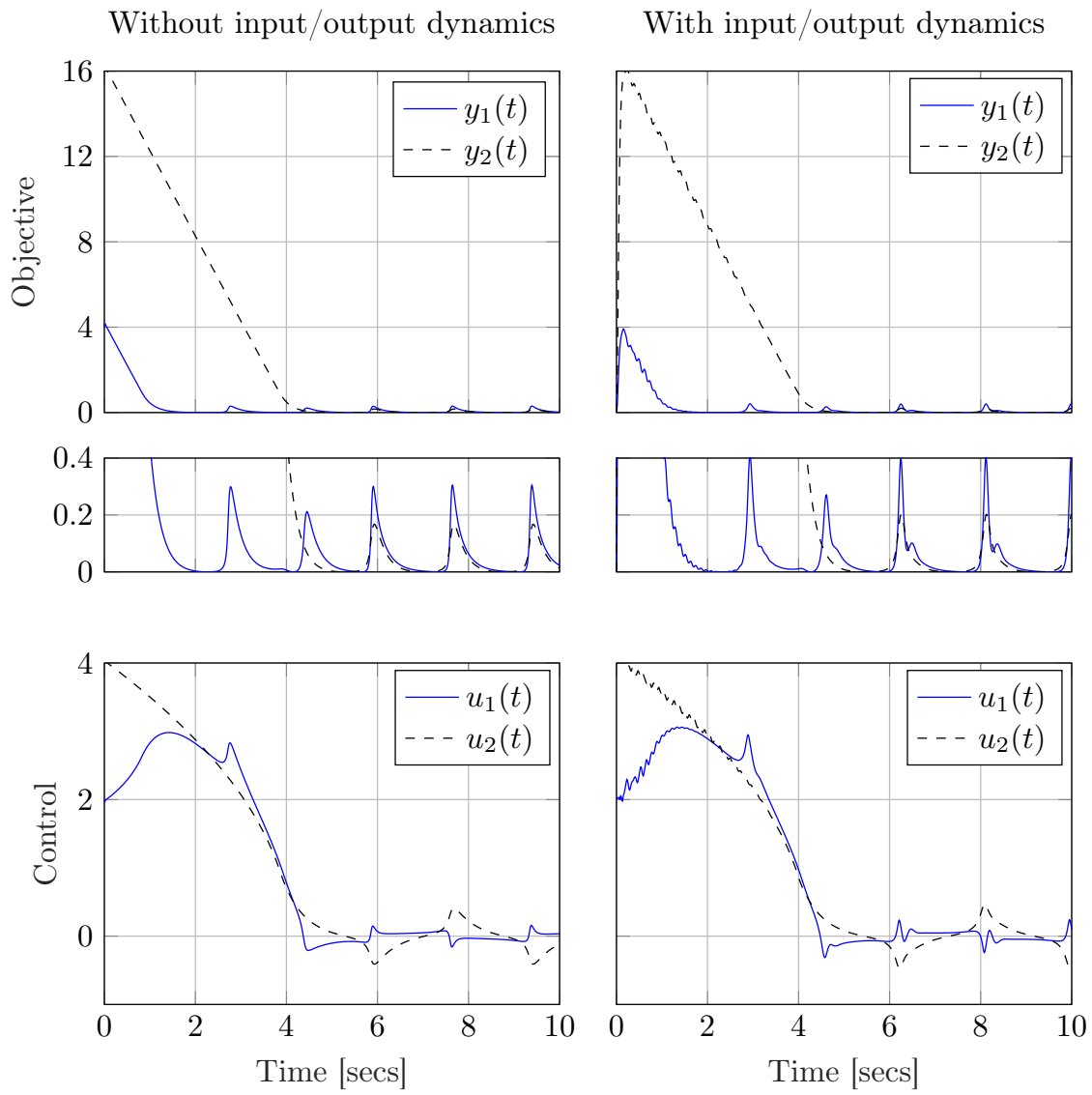


Figure 4.3: Simulation results for the MIMO ESC example of section 4.2.5 with input and output filters (4.52) and the continuous control implementation. To the left, the nominal results without the filters. To the right, the results after adding the filters.

(downstream) are considered as the control efforts $U(t) = P_p(t, L)$, the measured output is the data signals also at the end of the span $Y(t) = P_s(t, L)$, and the upstream data signals power $d(t) = P_s(t, 0)$ are faced as disturbances. The idea in considering feedback control of such amplifiers is to obtain robust signal power regulation in the presence of upstream signal power disturbances $d(t)$ and other uncertainties (Palais 1998; Dower et al. 2008; Peixoto and Oliveira 2012). The aim is also to minimize the variation of the downstream data channels power.

It is proposed a static model to describe the signal power dynamics in the fiber, which satisfactorily approximates the well-known model based on partial differential equations (PDEs). We show a more realistic numerical application example for

the optical fiber using 32 signal channels, while only four channels were considered in (Peixoto and Oliveira 2012), that highlights the advantages of the proposed global and multi-variable extremum-seeking controller.

In what follows, we give the detailed steps required to represent the simulated Raman optical amplifier in the form (4.6). First, we find the static MIMO model by numerically solving the corresponding PDEs representing the propagation of the signals along the fiber. Then, we select the pump wavelengths so that each pump strongly affects a set of data signals. Finally, we apply the proposed objective function to obtain the final output according to (4.5), and include the input integrator to fit equation (4.6).

Static MIMO Model in Raman Problem Propagation

Consider the simplified average field power model given in (Bromage 2004) for Raman amplifier fiber spans. This mathematical model can be described by N first-order nonlinear transport partial differential equations as in (Dower et al. 2008; Peixoto and Oliveira 2012), see also appendix B. By numerically solving the PDEs and considering the steady-state solution

$$\lim_{t\to\infty} P_s(t,L) = H(\theta_p) \tag{4.53}$$

which relates a constant pump power $\theta_p = P_p(t,L) \in \mathbb{R}^{n_p}$ to the corresponding limit of the output $P_s(t,L)$, the smooth and static map $H : \mathbb{R}^{n_p} \mapsto \mathbb{R}^{n_s}$ is obtained, with $H = \left[\begin{array}{ccc} H_1 & \dots & H_{n_s} \end{array} \right]^{\mathsf{T}}$ and $H_i : \mathbb{R}^{n_p} \mapsto \mathbb{R}_+$, provided $D(t) = P_s(t,0) = \theta_s$ is a constant vector.

Besides the fast transient behavior of the power dynamics, which is explained in details in (Peixoto and Oliveira 2012), chattering is reduced by adding integrators at the input side. Then, the downstream signal power $Y(t) = P_s(t,L)$ can be approximated by the following smooth and static map $Y(t) = H(U(t))$. In a nutshell, the following first-order MIMO non-linear system represents the process:

$$\dot{x} = u\,, \quad Y = H(x)$$

with $u \in \mathbb{R}^{n_p}$ being the control input, $x = U \in \mathbb{R}^{n_p}$, the state vector (the actual plant input), and Y the measured output.

Selection of the Pump Wavelength

As shown in (Kidorf et al. 1999), a pump with wavelength λ_j strongly affects data signals centered 100 nm above λ_j. Thus, let g_j be disjoint subsets of $\{1, 2, \dots, n_s\}$ constituted by the index of the data signal power vector Y related to the signals

strongly affected by the j-th pump signal power x_j, with x_j being the j-th element of the pump signal power vector x (Peixoto and Oliveira 2012). Let $H_{g_j} : \mathbb{R}^{n_p} \mapsto \mathbb{R}^{p_j}$ be the map obtained by collecting all H_i, with $i \in g_j$ and H_i from (4.53), where p_j is the number of elements of the set g_j, i.e., the number of data signals belonging to group g_j.

The output of the group g_j is given by the performance index of interest:

$$y_j(t) = J_j(H_{g_j}(x))$$

where the *objective function* (or output static mapping) $J_j : \mathbb{R}^{p_j} \mapsto \mathbb{R}$ is defined by (Dower et al. 2008; Peixoto and Oliveira 2012):

$$J_j(\zeta) = -(J_{aj}(\zeta) + R_j J_{bj}(\zeta) \tag{4.54a}$$

$$J_{aj}(\zeta) = \sqrt{(\zeta - \zeta_d)^{\mathsf{T}} Q_j (\zeta - \zeta_d)} \tag{4.54b}$$

$$J_{bj}(\zeta) = \sqrt{\frac{1}{p_j - 1} \sum_{i=1}^{p_j} [\zeta_i - E_j(\zeta)]^2} \tag{4.54c}$$

$$E_j(\zeta) = \frac{1}{p_j} \sum_{i=1}^{p_j} \zeta_j \tag{4.54d}$$

where $Q_j \in \mathbb{R}^{p_j \times p_j}$ is a positive semi-definite design matrix and $R_j \geq 0$ is a positive design constant.

The first term J_{aj} penalizes deviations of the output signal power (for group g_j) from the desired power level $\zeta_d \in \mathbb{R}^{p_j}$, and the second term J_{bj} (proportional to the standard deviation) penalizes ripple in the corresponding signal spectrum. An alternative (Dower et al. 2008) for penalizing ripple in the signal spectrum is given by $J_{bj}(\zeta) = \sum_{i,k=1,\ i \neq k}^{p_j} R_{ik}^j (\zeta_i - \zeta_k)^2$, where the constants R_{ik}^j can be chosen to satisfy $R_{ik}^j \propto \frac{1}{(\lambda_i - \lambda_k)^2}$, with λ_i and λ_k being the wavelength of the signals belonging to the same group g_j.

Finally, the process can be modeled as in (4.6):

$$\dot{x}(t) = u(t) \tag{4.55a}$$

$$y = h(x) = \begin{bmatrix} h_1(x) & \dots & h_n(x) \end{bmatrix}^{\mathsf{T}} \tag{4.55b}$$

where $h_j(x) = J_j(H_{g_j}(x))$.

A Particular Optical Fiber with 32 Data Signals

An optical fiber *TrueWave® Reach - Low Water Peak* with 32 input signals ($n_s = 32$) and two backward pumps ($n_p = 2$), spreading a length of $L = 100$ km is considered. The data signals are included at $z = 0$ and the pumps at $z = L$. In contrast to (Peixoto and Oliveira 2012), here we extend the number of data signals from 4 to 32, with 32 wavelengths (given in nm), split into two groups, $\Lambda_s = \begin{bmatrix} 1530 \ldots 1541.8 & 1570 \ldots 1582.4 \end{bmatrix}^\mathsf{T}$. In each group, signals are equally spaced by 100 GHz. As in (Peixoto and Oliveira 2012), we use the same number of pumps with wavelengths $\Lambda_p = \begin{bmatrix} 1442 & 1490 \end{bmatrix}^\mathsf{T}$.

As a nominal scenario, we consider that all data signals enter the fiber with -10 dBm power.

Model Parameters: Experimental Characterization

Comparing our simulation results with experimental data and the overall qualitative behavior of the Raman optical amplifier system, it is possible to verify that our simulations are representative of the expected steady-state behavior of such systems. Specifically, our simulations reflect a Raman optical amplifier based on the *TrueWave® Reach - Low Water Peak* (RFLWP) optical fiber, with model parameters experimentally characterized by the company OFS Fitel Denmark ApS. The results are consistent with those in (Dower et al. 2008) and (Bromage 2004), which indicates that our proposed control scheme has the potential to provide consistent practical results.

The Raman gain spectra (Agrawal 2001) of the optical fiber considered here fits the experimental spectra provided in (Bromage et al. 2002), see fig. 4.4, differing only by a scale factor of 1.2.

Objective Function

The objective functions in (4.54a) are implemented with $p_1 = p_2 = 16$, $Q_1 = Q_2 = I_{16\times16}$, $\zeta_d = [-10 \ \ldots \ -10]^\mathsf{T}$ (dBm), $R_1 = 10$, and $R_2 = 5$. Note that deviations from the desired level of -10 dBm are penalized for both groups as well as the standard deviation.

Approximations to Speed Up the Numerical Simulations

Consider that both pump powers x_1 and x_2 are fixed and belong to the grid $\mathcal{P} = \{15, 15.1, 15.2, \ldots, 30\}$ (dBm), i.e., $x(t) = P_p(t, L)$ is constant ($\forall t$), where $x = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^\mathsf{T} \in \mathbb{R}^2$ and $x_1, x_2 \in \mathcal{P}$. Then, by numerically solving the PDEs it is possible to obtain the static mapping $H(\cdot)$ from (4.53), where $H(x) = \lim_{t\to\infty} P_s(t, L) =$,

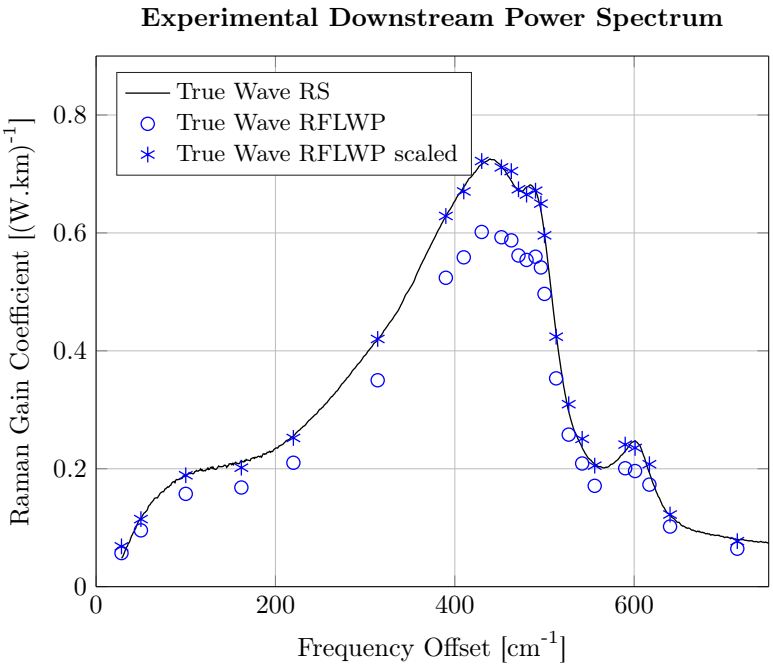


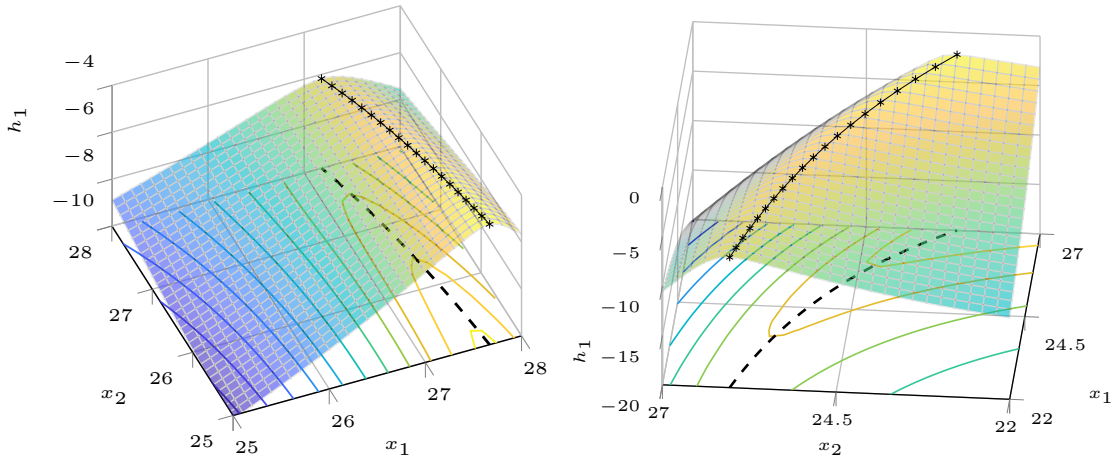
Figure 4.4: In solid black, the experimental data extracted from (Bromage et al. 2002), for a *TrueWave*® *Reduced Slope* (RS) optical fiber. In blue circles our experimental data characterized by OFS Fitel Denmark ApS for a *TrueWave*® *Reach - Low Water Peak* (RFLWP) optical fiber. The "x" markers correspond to our experimental data multiplied by a factor of 1.2, showing that the *TrueWave*® (RS) and the *TrueWave*® (RFLWP) data differ only by a scale factor.

for each fixed $x_1, x_2 \in \mathcal{P}$. To further refine this grid, the output functions $H_i(x)$ $(i = 1, \ldots, 32)$ and objectives functions $h_1(x) = J_1(H_{g_1}(x))$ and $h_2(x) = J_2(H_{g_2}(x))$ are approximated via spline interpolation.

Figures 4.5a and 4.5b illustrate the input-output static mapping $y_1 = h_1(x) = J_1(H_{g_1}(x))$ and $y_2 = h_2(x) = J_2(H_{g_2}(x))$, for the RFLWP optical fiber with downstream data signals power at -10 dBm and a desired upstream power of -10 dBm. The maximum values in the region of interest $x_1, x_2 \in [20, 28]$ (dBm) are also displayed.

Controller Parameters and Modulation Function

The control strategy (4.22), (4.24) was implemented with $\lambda_1 = \lambda_2 = 6$ (dBm/μs). The tracking error (4.25) is such that the reference signal (4.26) is a ramp function generated with $K_m = [6 \;\; 17.36]^\mathsf{T}$ (dBm/μs), and $y_m(0) = [-7.46 \;\; -13.02]^\mathsf{T}$ (dBm). In order to reduce the oscillations around the set of maximizers, we have considered sufficiently small values of $T_1 = T_2 = 0.01$ (dBm).



(a) Nominal static map $h_1(.)$.

(b) Nominal static map $h_2(.)$.

Figure 4.5: IO nominal maps $y_1 = h_1(x)$ and $y_2 = h_2(x)$ and the maximum values (in black) corresponding to the set of maximizers Θ_1^* and Θ_2^*, respectively.

Simple Modulation Function Implementation

Since there is an inherent decoupling between each pump and its most affected data channels, it is not necessary to compute the modulation function as a time-varying function. For instance, the same results displayed in this section can be obtained with constant modulation functions $\rho_1 = \rho_2 = 1.6$ (dBm/μs).

Full Modulation Function Implementation

For completeness regarding the global convergence results, we show how to obtain the time-varying modulation function, as described in section 4.2.3. To design the modulation function we perform all computations using the nominal plant parameters inside the region of intereset. Initially, we adjust the size of the Δ-neighborhoods to $\Delta = 0.01$ mW so that the desired final error is smaller than 0.01 mW for both channels. Then, we compute the norm bound $g_{21}(x)$ for $(\partial h_2/\partial x_1)/(\partial h_2/\partial x_2)$.

We set $s_2(x) = (1 + r)\, g_{21}(x)$, with $r = 0.5$, and $s_1(x) = 1$, and select the constants $p_{12} = 1/(1 + r) = 2/3$ and $p_{21} = 0.9$ in order to assure that

$$\left|\frac{\partial h_2}{\partial x_1}\right| < p_{12} \left|\frac{\partial h_2}{\partial x_2}\right| \frac{s_2}{s_1}, \quad x \notin \mathcal{D}_{\Delta_2}$$

and

$$\left|\frac{\partial h_1}{\partial x_2}\right| < p_{21} \left|\frac{\partial h_1}{\partial x_1}\right| \frac{s_1}{s_2}, \quad x \notin \mathcal{D}_{\Delta_1}$$

Hence, the matrix P in Assumption 19 is given by

$$P = \begin{bmatrix} 0 & 2/3 \\ 0.9 & 0 \end{bmatrix}$$

Finally, we compute the partial derivatives $\partial h_j/\partial x_j$ and use the values for Δ and s_j to find the lower-bounds $L_1 = 1$ and $L_2 = 3.7$, which ensure that Assumption 18 holds, i.e., $L_j \leq |k_{pj} s_j|$ holds in the region of interest.

The modulation functions $\varrho_1 = 44.7$ (dBm/μs) and $\varrho_2 = 47.5$ (dBm/μs) are obtained via linear programming by solving (4.38). In this example, since both $p_{12} < 1$ and $p_{21} < 1$, this solution is equivalent to obtaining $x_\varrho = [\varrho_1 \ \varrho_2]^\mathsf{T}$ from $x_\varrho = (I - P)^{-1}\mathcal{B}$, where $\mathcal{B} = \begin{bmatrix} \mathcal{B}_1 & \mathcal{B}_2 \end{bmatrix}^\mathsf{T}$, with $\mathcal{B}_j = \frac{1}{L_j}(K_{mj} + \lambda_j + \delta)$, for $j = 1, 2$, and $\delta = 1$ (dBm/μs).

Numerical Simulations

The following numerical simulations illustrate the applicability of the proposed extremum-seeking controller, where a parametric perturbation in the fiber is considered in order to point out the recovery capability of the proposed scheme.

It must be stressed that our control algorithm is easy to implement in a real experiment due to its simplicity. Only two integrators are needed, the sine function can be implemented via lookup tables, and the sign function can be approximated by any sigmoid function if a smooth control effort is required.

The plant initial conditions (pumps powers) are given by $x_1(0) = 25$ dBm and $x_2(0) = 27$ dBm. The Euler integration method was used with step size equal to 2^{-32} s ≈ 0.2 ns. Although apparently very small, this simulation step size is around one thousand times smaller than the plant dynamics (which are in the order of microseconds), thus consistent with the overall problem statement.

Full Pump Effort $\times$ Online Optimized Pump Effort

Here, it is illustrated that if one simply drives the Raman amplifier with full pump power, a signal power spectrum far from a flat shape is produced. Figure 4.6 shows the data and pump signals power spectrum. The pink line corresponds to the initial power distribution, while the red line corresponds to the final power spectrum when full power is employed in the pump signals $(x_1, x_2) = (28, 28)$ dBm. It is clear that the power of all data signals increase but without assuring any optimum criterion.

On the other hand, when ESC is applied, a flat spectrum is obtained (blue line). Moreover, since the objective functions of both groups are penalized to achieve the desired downstream signal power level (black line) and to reduce ripple in the power spectrum, both groups achieve a power spectrum around the desired level with a root mean squared error and ripple smaller than 0.5 dBm. This small error is also evident in fig. 4.7, which shows the evolution of the downstream data signals powers through time.

The corresponding time behavior of the control pump signals are illustrated

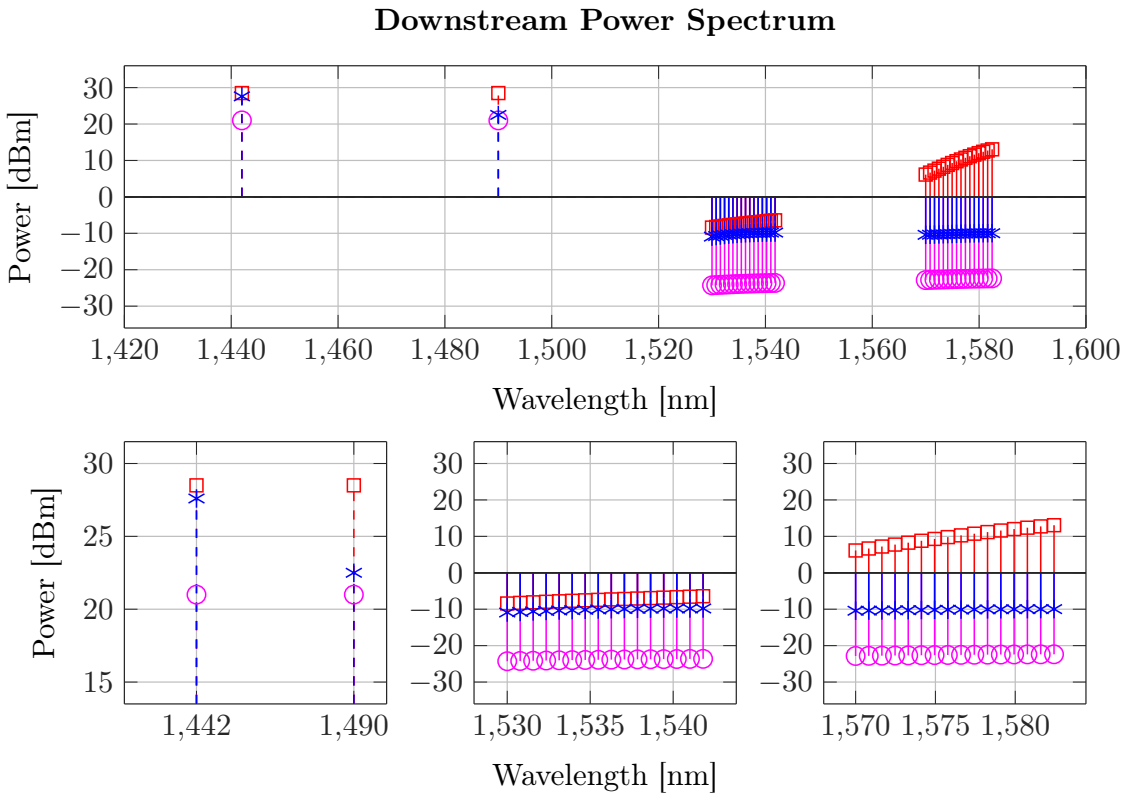


Figure 4.6: Power spectrum of all 34 signals transmitted in the optical fiber (top): power in dBm and wavelength in nm. Zoomed-in views are displayed in the bottom figures: 2 pumps control signals (bottom left), 16 data signals belonging to group 1 (bottom center) and 16 data signals belonging to group 2 (bottom right). Initial power in pink (circle), optimal power in blue (star), and the results of applying maximum pump power in red (square). The black line corresponds to the desired power level (−10 dBm).

in fig. 4.8, where the convergence to the Δ-neighborhoods is apparent. The time evolution of the plant outputs is displayed in fig. 4.9. This can also be observed in fig. 4.8, where $x_2(t)$ enters the Δ-neighborhood at $t = 75$ μs while $x_1(t)$ enters in the corresponding Δ-neighborhood at $t = 50$ μs.

Downstream Power Regulation in the Presence of Upstream Power Variations

The coefficients of $h_1(\cdot)$ and $h_2(\cdot)$ are assumed uncertain for control design, belonging to intervals around the nominal values. Step changes in the upstream signal powers or in the downstream desired power level correspond to some change in the coefficients of $h_1(\cdot)$ and $h_2(\cdot)$ and, consequently, in the optimal power of the pumps.

For robustness assessment, before $t = 150$ μs, the nominal plant is considered with upstream (desired downstream) signal power at -10 dBm (-10 dBm) and, after that, the static output maps are abruptly replaced by their perturbed versions

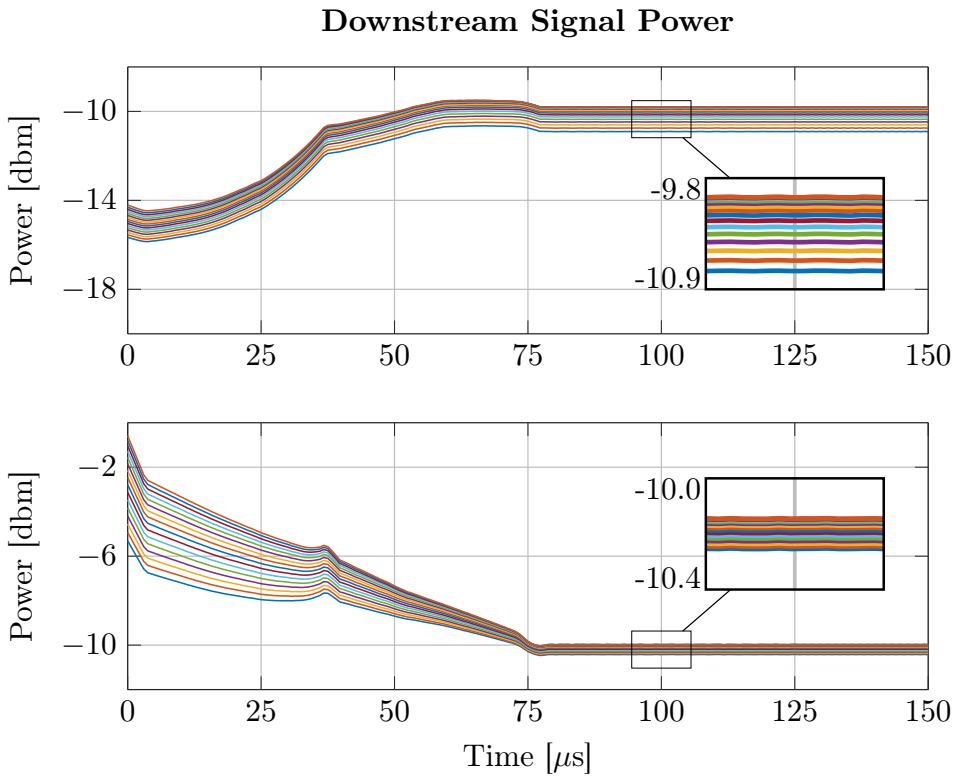


Figure 4.7: Time behavior of the power of all 32 data signals in dBm: 16 data signals belonging to group 1 (top) and 16 data signals belonging to group 2 (bottom).

corresponding to a step change in the upstream signal power from -10 dBm to -3 dBm. In the sequel, at $t = 200$ μs, the upstream signal power returns to the nominal value (-10 dBm) while the desired downstream signal power drops to -20 dBm.

Figures 4.10 and 4.11 show the convergence of the pumps power (x_1, x_2), initialized at $(25, 27)$ dBm, to the Δ-neighborhood of the final maximizer $(23.8, 21.2)$ dBm, after two abrupt perturbations which modify the optimal operating condition.

Comparison: Downstream Power Spectrum Optimization

In (Dower and Farrell 2006), the authors propose a regulatory control to maintain the downstream signals power at some fixed desired value, corresponding to fixed pump powers, which should be provided by an external set-point selection algorithm. Thus, optimization is not incorporated in the closed-loop control. Using this fixed set-point and a known model of the fiber dynamics, a linearized model is obtained around the operational point. Through simulations, the authors indicate the robustness of this technique with respect to small changes around the operational point (upstream and downstream pumps and signals powers), see figure 3 in (Dower and Farrell 2006).

In contrast to (Dower and Farrell 2006), we incorporate online optimization in the control-loop, so that the proposed control scheme can find the optimal pump

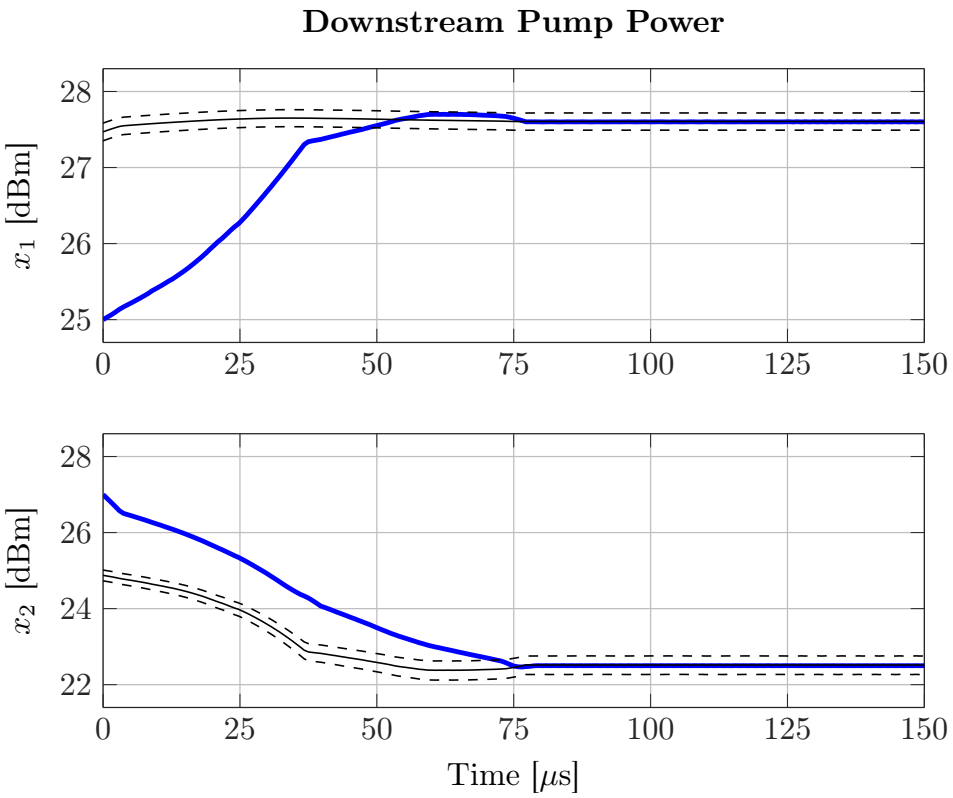


Figure 4.8: Time behavior of the pump signals (in blue). Signal $x_1(t)$ (top) and signal $x_2(t)$ (bottom). All signals are expressed in dBm. The black line shows the maximizers values, and the dashed lines are the corresponding *Delta*-neighborhoods.

powers required for regulating the downstream signals powers at some desired level. Our control strategy can easily cope with more significant changes to the upstream signals power and the desired downstream signals power. In comparison to (Dower and Farrell 2006), where small variations of 1 dBm are applied to the upstream signals and the desired downstream signals power, we show robustness to changes of 10 dBm. Only small variations are possible in (Dower and Farrell 2006) because the authors rely on linearization.

In (Dower et al. 2008), a perturbation based ESC is employed to shape the downstream signal power spectrum according to some flat function, i.e., the same control objective considered in this Thesis and published in (Peixoto et al. 2020). However, in (Dower et al. 2008), the system outputs are evaluated many times per iteration to obtain approximated gradient calculations and to conduct line searches.

In comparison to (Dower et al. 2008), our proposed controller is more straightforward because it does not require probing the system to collect evaluations of the objective function output, and no gradient estimation is needed. Our ESC scheme drives the pump powers toward the optimum point by using only two integrators and the periodic switching function control law. Moreover, we achieve similar results

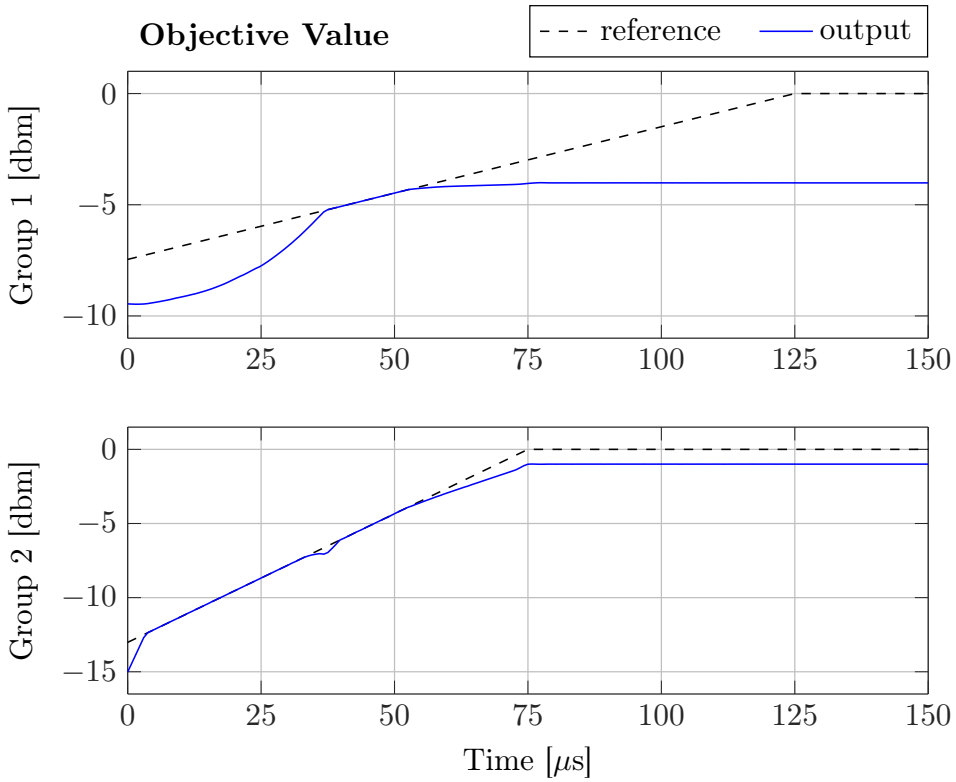


Figure 4.9: Time behavior of the objective functions corresponding to the static maps. Top: plant output y_1 and reference trajectory y_{m1}. Bottom: plant output y_2 and reference trajectory y_{m2}.

but with smaller relative error and ripple (both lower than 0.5 dBm, cf. fig. 4.7).

4.3 Continuous SM-ESC for Nash Equilibrium Seeking in Input-Affine Processes

The results published by Peixoto et al. (2020) and described in the last section, apply to nonlinear static processes. Such processes are usually steady-state approximations of some dynamic system. In this section, we proceed in the direction of relaxing the assumption of a static process and generalize the results for input-affine nonlinear systems. Furthermore, following the control design proposed in chapter 2, we also generalize the control algorithm to provide for continuous control actions, and control laws based on techniques other than sliding-modes.

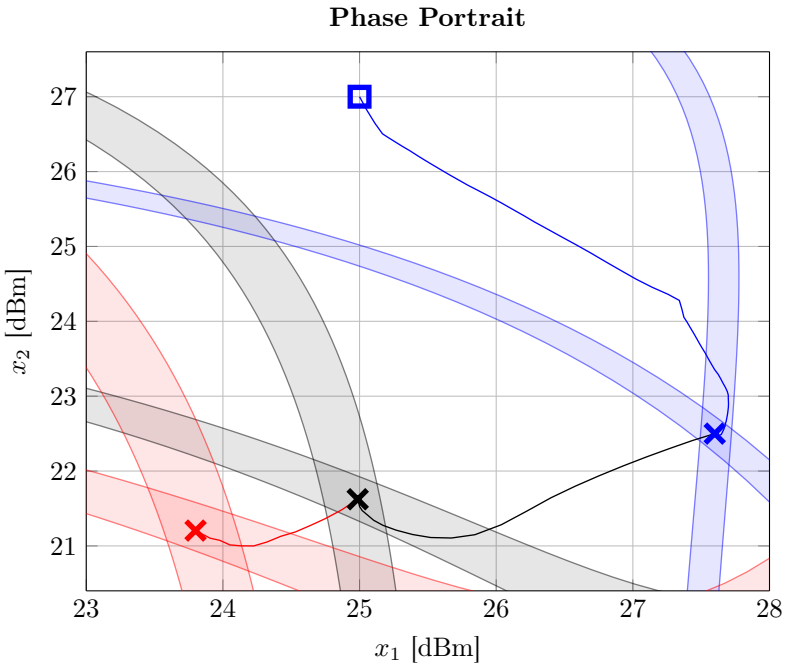


Figure 4.10: Phase portrait starting from the pair $(25, 27)$ dBm and converging to the pair $(27.6, 22.5)$ dBm, when the nominal plant is considered (blue); converging subsequently to the pairs $(25, 21.6)$ and $(23.8, 21.2)$ when the perturbed plants are considered (black and red lines). The Δ-regions for each case are the shaded areas.

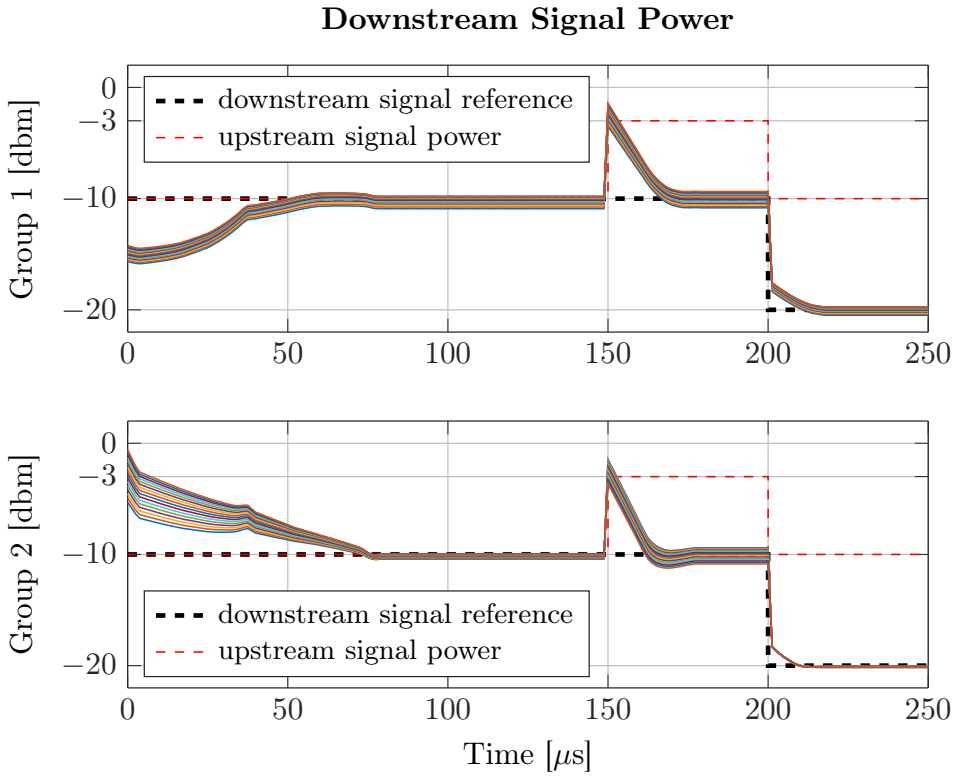


Figure 4.11: Robust downstream signal power regulation in the presence of upstream signal power fluctuations (from -10 dBm to -3 dBm) and changes to the reference downstream power level (from -10 dBm to -20 dBm).

4.3.1 The Class of Input-Output Mappings

Regarding the process dynamics, instead of considering a static nonlinear map, consider the optimization problem (4.1), and the process dynamics

$$\dot{x}(t) = f(x) + g(x)u(t) \tag{4.56a}$$

$$y(t) = h(x) \tag{4.56b}$$

with internal states $x : \bar{\mathbb{R}}_+ \mapsto \mathbb{R}^n$, measured outputs $y : \bar{\mathbb{R}}_+ \mapsto \mathbb{R}^n$, piecewise continuous control input $u : \bar{\mathbb{R}}_+ \mapsto \mathbb{R}^n$, and unknown functions $f : \mathbb{R}^n \mapsto \mathbb{R}^n$ and $g : \mathbb{R}^n \mapsto \mathbb{R}^{n\times n}$. Discontinuities are only allowed in the control input, such that functions f, g, and h are at least C^0.

As in section 4.2, for simplicity, we assume that $y(t) = h(x)$ represents both the process output and the objective function value.

Following the usual procedure of differentiating the output with respect to time, the first derivative of y is written as,

$$\dot{y}(t) = L_f h(x) + L_g h(x)u(t) \tag{4.57a}$$

$$L_g h(x) = \left(\frac{\partial h}{\partial x}\right)^{\mathsf{T}} g(x) = k_p(x) + \tilde{k}_p(x) \tag{4.57b}$$

with $k_p(x)$ holding the elements in the diagonal of $L_g h(x)$, and $\tilde{k}_p(x)$ the off-diagonal elements, just as it was done in (4.13). Regarding the diagonal entries, we make the following assumption.

Assumption 21 (Unit Relative Degree)**.** *Let the Δ-neighborhood be defined as in Definition 9. Outside the Δ-neighborhood, the elements on the main diagonal of matrix k_p, denoted k_{pj}, are bounded away from zero,*

$$0 < |k_{pj}| \quad \forall x \notin \mathcal{D}_{\Delta_j} \tag{4.58}$$

with $j \in \{1, \ldots, n\}$. Thus, for $x \notin \mathcal{D}_{\Delta_j}$, the process has unit relative degree from input u_j to output y_j.

Considering the change of variables $u(t) = S(x)v(t)$, with matrix $S(x)$ such that Assumptions 18 and 19 are satisfied with $k_p(x)$ from (4.57b), we rewrite the output dynamics as

$$\dot{y}(t) = k_p S(x)\left[v(t) + S^{-1}k_p\tilde{k}_p S(x)v(t) + S^{-1}k_p^{-1}(x)L_f h(x)\right] \tag{4.59}$$

Two aspects must be taken into account. First, we whish to develop an output-feedback control law. Therefore, the matrix $S(x)$ should be replaced by an estimate

which does not depend on the states. Second, to make for a straightforward proof, we merge the control laws presented in sections 2.2 and 4.2. Thus, considering the following assumption, which allows us to replace matrix $S(x)$.

Assumption 22 (Norm Observability)**.** *Let $\xi(t) = x(t) - x_d$, with $x_d \in \mathbb{R}^n$ some estimate of the optimizer $x^* : y^* = h(x^*)$, and assume that (4.56) is uniformly input-output-to-state stable (UIOSS), such that it admits a norm observer (see Assumption 6). Hence, there is η such that*

$$\|\xi\| \leq \bar{\alpha}_3\left(\|\eta\|\right) + \beta_2(\|\xi\|, t) \tag{4.60a}$$

$$\dot{\eta} = -\eta + \alpha_3(\|u\|) + \alpha_4(\|y\|), \ \eta(0) = 0 \tag{4.60b}$$

with known functions $\bar{\alpha}_3(s) = \bar{\alpha}_1^{-1}(2s) \in \mathcal{K}_\infty$, $\beta_2(s,t) = \bar{\alpha}_1(2\bar{\alpha}_2(s)e^{-t}) \in \mathcal{KL}$, and $\alpha_3, \alpha_4 \in \mathcal{K}$, from Assumption 6.

Matrix $S(x)$ is chosen such that the controllability loss near the Δ_j-neighborhood is kept above a lower-bound (Assumption 18) and that the j-th terms in the diagonal of the HFG matrix dominate the off-diagonal terms of the j-th row (equation (4.18) and Assumptions 19 and 20).

Assumption 23. *Let $\xi(t) = x(t) - x_d$, with $x_d \in \mathbb{R}^n$ some estimate of the optimizer $x^* : y^* = h(x^*)$. There exists $S(x) = \text{diag}\left(\ s_1(\|\xi\|), \ \ldots, \ s_n(\|\xi\|) \ \right)$, with $s_j(.) \in \mathcal{K}$, such that after some finite time Assumptions 18 and 19 hold replacing $s_j(\|x\|)$ by $s_j \circ \bar{\alpha}_3(\|\eta\|)$, with η and $\bar{\alpha}_3$ from Assumption 22.*

Since we assume the existence of a state norm observer, the previous assumption is enough to write the ESC algorithm in an output-feedback framework. It restricts the classes of objective function for which the control law is valid. Nonetheless, as before, objective functions that are (at least locally) polynomial in x, and with a "dominant triangular" Jacobian matrix are still covered.

Remark 9. *Additionally, if the functions f and g are such that $\dot{x}_j$ depends only on the states x_j up to x_n (x_0 up to x_j) and the objective function Jacobian is "dominant upper-triangular", then it is possible to develop n norm observers, each with state $\eta_j \in \mathbb{R}^j$, one for each $|x_j - x_{dj}|$. With these estimates, based on Remark 7, one might use a diagonal matrix $S : \mathbb{R}^n_+ \mapsto \mathbb{R}^{n\times n}_+$, such that*

$$s_j(\eta_j, \ldots, \eta_n) = s_{j+1}(\eta_{j+1}, \ldots, \eta_n) \max\left(\frac{a_j}{\bar{\alpha}_{3,j}^{r_j - 1}(\|\eta_j\|) + b_j}, \ 1 \right) \tag{4.61a}$$

$$s_n(\eta_n) = 1 \tag{4.61b}$$

with $a_j, b_j > 0$, and $\bar{\alpha}_{3,j}(.) \in \mathcal{K}_\infty$ obtained from each norm observer.

To obtain a control law that does not depend on the states, the disturbances due to $L_f h(x)$ must also be bounded. Thus, consider also the following assumption.

Assumption 24 (Disturbance Boundedness). *Let $(L_f h)_j$ denote the j-th element of $L_f h$, and $\xi(t) = x(t) - x_d$, as in Assumption 22. There exist known functions $\alpha_{1,j} \in \mathcal{K}$ and $\beta_{1,j} \in \mathcal{KL}$ such that $\left|(L_f h)_j\right| \leq \alpha_{1,j}(\|\xi\|) + \beta_{1,j}(\|\xi(0)\|, t)$.*

4.3.2 Extremum-Seeking Controller

The control law presented in this section is an extension of that of section 4.2.2. Note, however, that although the controller in section 4.2.2 is not explicitly developed for input-affine nonlinear systems, it is capable of dealing with such systems given that the modulation function overcomes the disturbances. In this section, we not only make sure this happens, but we also extend the controller so that continuous and non-sliding-mode based algorithms can be implemented.

Combining the controllers (4.22) and (2.16), the ESC with periodic switching function is written as

$$u(t) = \rho(t)\,\text{sigmoid}_\epsilon\big(\sin\left[\pi \mathcal{T}^{-1}\sigma(t)\right]\big) \tag{4.62a}$$

$$\rho(t) = S(\eta)\varrho(t) \tag{4.62b}$$

$$\sigma(t) = e(t) + \int_0^t f_e(e)d\tau \tag{4.62c}$$

$$e(t) = y(t) - y_m(t) \tag{4.62d}$$

with $\rho(t) = \text{diag}\left(\ \rho_1(t)\,,\ \ \ldots\ ,\ \ \rho_n(t)\ \right)$ being the *modulation function* (continuous in t) to be defined later on by designing $\varrho(t)$, $S(\eta)$ satisfying Assumption 23, $\mathcal{T} = \text{diag}\left(\ T_1\,,\ \ \ldots\ ,\ \ T_n\ \right) > 0$. The function $y_m \in C^1 : \bar{\mathbb{R}}_+ \mapsto \mathbb{R}^n$ is such that $\exists\, t_m : y_m(t) \leq y^*$, $\forall t \geq t_m$, and the function $f_e : \mathbb{R}^n \mapsto \mathbb{R}^n$ is such that the following assumption holds.

Assumption 25 (Stable Error Dynamics Design). *Function $f_e : \mathbb{R}^n \mapsto \mathbb{R}^n$ is chosen such that*

- *the j-th element of f_e, denoted f_{ej} is a function of e_j only;*
- *the origin of $\dot{e}_j + f_{e_j}(e_j) = 0$ is globally asymptotically stable (GAS);*
- *solutions to $\dot{e}_j + f_{ej}(e_j + \mathcal{O}(\epsilon)) = 0$ with small disturbances of order $\mathcal{O}(\epsilon)$ in f_{ej} are globally ultimately bounded, with ultimate bound also of order $\mathcal{O}(\epsilon)$;*
- *the integral $\int_{t_a}^{t_b} |f_{ej}(e_j)|\, d\tau \leq \mathcal{O}(t_b - t_a)$;*

- *given an initial state* $e_j(t_0)$*, the dynamics* $\dot{e}_j + f_{ej}(e_j + \mathcal{O}(\epsilon)) = 0 \implies |e_j(t)| \leq |e_j(t_0)|$, $t \geq t_0$.

4.3.3 Modulation Function Design

Matched Input Disturbance Norm Bound

To develop the modulation function, once again, we must write the σ-dynamics and find an expression for the matched disturbances. Thus, from (4.56) and (4.62),

$$\dot{\sigma}(t) = k_p(x)[u(t) + d_\sigma(x, e, \dot{y}_m)] \tag{4.63a}$$

$$k_p(x)\, d_\sigma(x, u, e, \dot{y}_m) = \left[\tilde{k}_p(x)u(t) - \dot{y}_m(t) + f_e(e) + L_f h(x)\right] \tag{4.63b}$$

Expanding the σ-dynamics element-wise,

$$\dot{\sigma}_j(t) = k_{pj}(x)\left[u_j(t) + d_{\sigma j}(x, \bar{u}_j, e_j, \dot{y}_{mj})\right] \tag{4.63c}$$

$$k_{pj}(x) d_{\sigma j} = \sum_{i \neq j} \left(\frac{\partial h_j}{\partial x_i}\right) u_i - \dot{y}_{mj} + f_{ej}(e_j) + \left(L_f h\right)_j \tag{4.63d}$$

where $\left(L_f h\right)_j$ denotes the j-th element of $L_f h$.

The above equalities hold $\forall x$. To design the modulation function, however, we must find a bound for d_σ, which must hold $\forall x \notin \mathcal{D}_j$. Thus, from Assumptions 18, 19 and 23, and comparing (4.63) with (4.34),

$$\frac{|d_{\sigma j}|}{s_j} \leq \sum_{i \neq j} p_{ij}\varrho_i + \frac{1}{k_{pj}s_j}\left[|f_{ej}(e_j) - \dot{y}_{mj}| + \left|\left(L_f h\right)_j\right|\right], \quad x \notin \mathcal{D}_{\Delta_j} \tag{4.64}$$

Furthermore, since we assume (i) the existence of a state norm observer (Assumption 22), (ii) that $S(x)$ satisfies Assumption 23, and (iii) that the state-dependent disturbances are bounded by class $\mathcal{K}$ and class $\mathcal{KL}$ functions (Assumption 24), there are known functions $\alpha_{\sigma_j}(.) \in \mathcal{K}$ and $\beta_{\sigma_j}(., t) \in \mathcal{KL}$, such that, after a finite time t_{σ_j},

$$\frac{|d_{\sigma j}|}{s_j} \leq \sum_{i \neq j} p_{ij}\varrho_i + \frac{1}{k_{pj}s_j}\left[|f_{ej}(e_j) - \dot{y}_{mj}| + \alpha_{\sigma_j}(\|\eta\|) + \beta_{\sigma_j}(\|\eta\|, t) + \delta_j\right], \quad \begin{matrix} x \notin \mathcal{D}_{\Delta_j} \\ t \geq t_{\sigma_j} \end{matrix} \tag{4.65}$$

where, based on Assumptions 22 and 24, α_{σ_j} and β_{σ_j} are obtained by the following expansion

$$\begin{aligned}\left|\left(L_f h\right)_j\right| &\leq \alpha_{1,j}(\|\xi\|) + \beta_{1,j}(\|\xi(0)\|, t) \\ &\leq \alpha_{1,j}\big(\bar{\alpha}_3(\|\eta\|) + \beta_2(\|\xi\|, t)\big) + \beta_{1,j}(\|\xi(0)\|, t)\end{aligned}$$

$$
\begin{aligned}
&\leq \alpha_{1,j}\left(2\bar{\alpha}_3(\|\eta\|)\right) + \alpha_{1,j}(2\beta_2(\|\xi\|,t)) + \beta_{1,j}(\|\xi(0)\|,t) \\
&= \alpha_{\sigma_j}\left(\|\eta\|\right) + \beta_{\sigma_j}(\|\xi\|/2,t) + \beta_{1,j}(\|\xi(0)\|,t) \\
&\leq \alpha_{\sigma_j}\left(\|\eta\|\right) + \beta_{\sigma_j}(\|\eta\|,t) + \beta_{\sigma_j}(\|\xi\| - \|\eta\|,t) + \beta_{1,j}(\|\xi(0)\|,t) \\
&= \alpha_{\sigma_j}\left(\|\eta\|\right) + \beta_{\sigma_j}(\|\eta\|,t) + \beta_{\sigma_j}(|e_\eta(0)|,t) + \beta_{1,j}(\|\xi(0)\|,t) \qquad (4.66)
\end{aligned}
$$

with $\alpha_{\sigma_j}(.) = \alpha_{1,j}\left(2\bar{\alpha}_3(.)\right) \in \mathcal{K}$ and $\beta_{\sigma_j}(.,t) = \alpha_{1,j}(\beta_2(.,t)) \in \mathcal{KL}$.

Modulation Function Design

From the previous section, it is known that a sufficient condition for producing sliding-modes on the manifolds $\sigma_j = k_j T_j$, for some integers k_j, is that (4.35) is satisfied, that is

$$
\rho_j = \varrho_j s_j \geq |d_{\sigma j}| + s_j\,\delta_j\,, \qquad j = 1,\ldots,n \qquad (4.67)
$$

Applying the new upper-bound (4.65) for $\left|d_{\sigma_j}\right|$, it follows that

$$
\varrho_j \geq \sum_{i\neq j} p_{ij}\varrho_i + \frac{1}{L_j}\left[|f_{ej}(e_j) - \dot{y}_{mj}| + \alpha_{\sigma_j}(\|\eta\|) + \delta_j\right] \qquad (4.68a)
$$

$$
(I-P)x_\varrho \geq \mathcal{B}(\dot{y}_m, e, \eta) \qquad (4.68b)
$$

$$
\mathcal{B}_j(\dot{y}_{mj}, e_j, \eta) = \frac{1}{L_j}\left[|f_{ej}(e_j) - \dot{y}_{mj}| + \alpha_{\sigma_j}(\|\eta\|) + \delta_j\right] \qquad (4.68c)
$$

where (4.68b) is understood element-wise and must hold for $x \notin \mathcal{D}_{\Delta_j}$, $\mathcal{B} = \begin{bmatrix} \mathcal{B}_1 & \ldots & \mathcal{B}_n \end{bmatrix}^\mathsf{T}$, and $x_\varrho = \begin{bmatrix} \varrho_1 & \ldots & \varrho_n \end{bmatrix}^\mathsf{T}$.

Thus, defining the objective function $C^\mathsf{T} x_\varrho$, with weight vector $C \in \mathbb{R}^n_+$, each ρ_j can be obtained solving

$$
\begin{aligned}
\min_{x_\varrho} \quad & C^\mathsf{T} x_\varrho\,, && (4.69a)\\
\text{s.t.} \quad & (I-P)x_\varrho \geq \mathcal{B}(\dot{y}_m, e, \eta)\,, && (4.69b)\\
& x_\varrho > 0
\end{aligned}
$$

where the only difference from (4.38) is in the matrix $\mathcal{B}$ formula, i.e. equation (4.68c).

4.3.4 Stability and Convergence Analysis

Having adapted the necessary parts from sections 2.2 and 4.2, we are now ready to state the proposition which summarizes the results regarding the existence of sliding-models (real or ideal).

Proposition 3 (Occurence of Sliding-Modes)**.** *Consider the system* (4.56)*, with control law* (4.62) *and modulation function* ρ *in* (4.62b) *satisfying* (4.69)*, while* $x \notin \mathcal{D}_{\Delta_j}$*. Let* $t_0 \geq t_{\sigma_j}$ *be a time instant such that* σ_j *is not in sliding motion at* $t = t_0$*, i.e.,* $\sigma_j(t_0) \neq k_{0j} T_j$ *where* $k_{0j} = \left\lfloor \frac{\sigma_j(t_0)}{T_j} \right\rceil$*, and such that* (4.65) *holds. Then, if* $x \notin \mathcal{D}_{\Delta_j}$*,* $\forall t \geq t_0$*, an* $\mathcal{O}(\epsilon_j)$ *real sliding-mode on* $\sigma_j = k^* T_j$*, with* $k_j^* = k_{0j} - 1$*,* $k_j^* \in \{k_{0j} - 1, k_{0j} + 1\}$*, is achieved in some finite time* $t_j \geq t_0$ *which satisfies* $t_0 \leq t_j < t_0 + \mathcal{O}(T_j/\delta_j)$ *with* $\delta_j > 0$*. Moreover, no finite-time escape occurs in the system signals.*

Proof. Following the same steps as in the proof of Proposition 2, let

$$k_j \pi \leq \frac{\pi}{T_j} \sigma_j < (k_j + 1)\pi \,, \qquad k_j = \left\lfloor \frac{\sigma_j}{T_j} \right\rfloor \tag{4.70}$$

and

$$k_{0j} = \lfloor \sigma_j(t_0)/T_j \rfloor \tag{4.71a}$$

$$k_j^* = k_{0j} + \frac{\operatorname{sign}(k_{pj}) + (-1)^{k_{0j}}}{2} \tag{4.71b}$$

$$\alpha_j = \left(\frac{\sigma_j(t)}{T_j} - k_j^* \right) \pi \tag{4.71c}$$

Now, however, we rewrite the control signal as

$$u_j(t) = \kappa_j \rho_j(t) \operatorname{sigmoid}_\epsilon \left(\sin \left[\frac{\pi}{T_j} \sigma_j(t) \right] \right) = \kappa_j \rho_j \beta_j(\sigma_j) (-1)^{k_j} \tag{4.72}$$

for nonnegative functions $\beta_j(\sigma_j)$ such that

$$\beta_j(\sigma_j) \geq 1/\kappa_j \quad \Longleftrightarrow \quad \left| \sin \left[\frac{\pi}{T_j} \sigma_j(t) \right] \right| > \epsilon_j \tag{4.73}$$

Without loss of generality, let $\epsilon_j \ll 1$, such that

$$\beta_j(\sigma_j) \geq 1/\kappa_j \quad \Longleftrightarrow \quad \left| \frac{\sigma_j(t)}{T_j} - \tilde{k}_j(t) \right| \pi > \epsilon_j \tag{4.74}$$

where $\tilde{k}_j(t) = \operatorname{round}(\sigma_j(t)/T_j)$ denotes the rounding of $\sigma_j(t)/T_j$ to the nearest integer value. Thus, defining $V(\alpha_j) = \frac{T_j}{\pi} |\alpha_j|$,

$$\begin{aligned} \dot{V}(\alpha_j) &= |k_{pj}| \left[-\kappa_j \rho_j \beta_j(\sigma_j) + d_{\sigma j} \operatorname{sign}(k_{pj}) \operatorname{sign}(\alpha_j) \right] \\ &\leq |k_{pj}| \left(-\kappa_j \rho_j \beta_j(\sigma_j) + |d_{\sigma j}| \right) \end{aligned}$$

and applying (4.74),

$$\dot{V}(\alpha_j) \leq |k_{pj}| \left(-\rho_j + |d_{\sigma j}|\right) , \quad |\alpha_j| > \epsilon_j$$

Considering the modulation function design from the solution of (4.69) that guarantees that (4.67) holds, and, once again, the proof of Proposition 2

$$\dot{V}(\alpha_j) \leq -\delta_j < 0 , \quad |\alpha_j| > \epsilon_j \tag{4.75}$$

and no finite-time escape occurs in the closed-loop signals.

Then, from (4.75), one can conclude that an $\mathcal{O}(\epsilon_j)$ sliding-mode occurs on the manifold $\alpha_j = 0$, or equivalently, on the manifold $\sigma_j = k_j^* T_j$. From the Lyapunov function definition, this sliding-mode occurs after a finite time t_j, which belongs to the interval

$$t_0 \leq t_j \leq t_0 + \frac{T_j \, |\alpha_j(t_0)|}{\pi \delta_j} \leq t_0 + \frac{T_j}{\delta_j} \tag{4.76}$$

■

Theorem 7. *Consider the system (4.56), with control law (4.62) and modulation function ρ satisfying (4.69), while $x \notin \mathcal{D}_{\Delta_j}$. Assume that all assumptions regarding the objective function hold, i.e. Assumptions 16 to 19 and 23. Then: (i) all Δ_j-neighborhoods $\mathcal{D}_{\Delta_j}$ are globally attractive and are achieved in finite time, and (ii) for a sufficiently small design constant L_j from Assumption 18, the oscillations around the minimum y^* of y can be made of order $\mathcal{O}(\max_j T_j)$, with T_j given in (4.62). In addition, all closed-loop signals remain uniformly bounded, except for σ which is only the argument of a sine function in (4.62a).*

Proof. (i) Attractiveness of $\mathcal{D}_{\Delta_j}$: Let $j \in \{1, \ldots, n\}$. To prove that all Δ_j-neighborhoods are attractive, let $t \geq t_j$, with t_j from Proposition 3, such that $\sigma_j = k_j^* T_j + \mathcal{O}(\epsilon_j)$, $\forall t \in [t_j, t_j^*)$, and consider the change of variables $\bar{e}_j = e_j - \sigma_j + k_j^* T_j$, such that, for $t \in [t_j, t_j^*)$, the σ_j-dynamics from (4.62c) is rewritten as

$$k_j^* T_j = \bar{e}_j + \int_0^t f_{ej}(\bar{e}_j + \mathcal{O}(\epsilon_j)) d\tau \tag{4.77a}$$

$$0 = \dot{\bar{e}}_j + f_{ej}(\bar{e}_j + \mathcal{O}(\epsilon_j)) \tag{4.77b}$$

Thus, from Assumption 25, the origin of $\bar{e}_j$ is UGPAS, with ultimate bound $\bar{e}_j \leq \mathcal{O}(\epsilon_j)$. Since $|e_j| \leq |\bar{e}_j| + |\sigma - k^* \, T|$, the origin of e_j is also UGPAS with ultimate bound $|e_j| \leq \mathcal{O}(\epsilon_j)$, and since y_m is designed such that $\exists \, t_m : y_m(t) \leq y^*$, $\forall t \geq t_m$, it follows that the regions Δ_j are reached in finite time.

(ii) Oscillations of Order $\mathcal{O}(\max_j T_j)$: Assume that x_j reaches the frontier of

$\mathcal{D}_{\Delta_j}$ (from the inside) at some finite time t'_j, such that σ_j is not in sliding-mode at $t = t'_j$. From Proposition 3, σ_j reaches a sliding-mode again after a finite time $t''_j \leq t'_j + T_j/\delta_j$. Following the same procedure as in Proposition 2, we show that, during this time, the output y_j distances itself from $h_j\left(\mathcal{D}_{\Delta_j}\right)$ by an amount of order $\mathcal{O}(\max_j T_j)$.

Recall the assumptions on $f_e(.)$ from Assumption 25. Then, from the sliding-variable σ definition (4.62c) and the inequality for $\Delta t_j = t''_j - t'_j \leq T_j/\delta_j$,

$$\begin{aligned}
\sigma_j(t''_j) - \sigma_j(t'_j) &= e_j(t''_j) - e_j(t'_j) + \int_{t'_j}^{t''_j} f_{ej}(e_j) d\tau \\
\left|e_j(t''_j) - e_j(t'_j)\right| &\leq \left|\sigma_j(t''_j) - \sigma_j(t'_j)\right| + \int_{t'_j}^{t''_j} \left|f_{ej}(e_j)\right| d\tau \\
\left|e_j(t''_j) - e_j(t'_j)\right| &\leq \left|\sigma_j(t''_j) - \sigma_j(t'_j)\right| + \mathcal{O}(\Delta t_j) \\
\left|e_j(t''_j) - e_j(t'_j)\right| &\leq T_j + \mathcal{O}(T_j/\delta_j) \leq \mathcal{O}(T_j)
\end{aligned} \tag{4.78}$$

where $\left|\sigma(t''_j) - \sigma(t'_j)\right| \leq T_j$ was used. From the output error definition (4.62d),

$$\begin{aligned}
\left|y(t''_j) - y(t'_j)\right| &\leq \left|e_j(t''_j) - e_j(t'_j)\right| + \left|y_{mj}(t''_j) - y_{mj}(t'_j)\right| \\
&\in \left\{ \mathcal{O}(T_j)\,,\ \mathcal{O}(T_j) + \max_{t \in [t'_j, t''_j]} \left|\dot{y}_{mj}(t)\right| \frac{T_j}{\delta_j} \right\}
\end{aligned} \tag{4.79}$$

where the last term is considered if the reference model is not saturated at some lower-bound below y^*_j. Otherwise, after a time $t'''_j : y_{mj}(t) < y^*_j(t)$ and $\dot{y}_{mj}(t) = 0$, for $t \geq t'''_j$, the same lower-bound as in (4.78) is considered. Either way, it follows that the oscillations above the region $h_j\left(\mathcal{D}_{\Delta_j}\right)$ are of order $\mathcal{O}(T_j)$, for every $j \in \{1, \ldots, n\}$.

Thus, for each channel, the oscillations above the $h_j\left(\mathcal{D}_{\Delta_j}\right)$ regions are of order $\mathcal{O}(T_j)$, and, in the worst case, the oscillations of y above y^* are of order $\mathcal{O}(\max_j T_j)$. ■

Remark 10 (Arbitrary Relative Degree)**.** *If one considers the results discussed in section 2.3, it becomes clear that the algorithm proposed in this section can be applied to nonlinear systems with arbitrary relative degree, given that both the assumptions made in this section and those necessary to apply Tikhonov's theorem (Khalil 2002, Theorem 11.2) as described in section 2.3 are satisfied.*

Chapter 5

Real-Time Optimization with Multiple Decision Variables

In this section, we develop a control algorithm capable of solving real-time optimization with multiple decision variables and only one objective function. This development is expected, considering that the most common way of solving multi-objective problems is by combining all objectives in a single one, a process known as scalarization. We will not, however, delve into the details of scalarization, and, for the reminder of this section, we will consider that the single-objective at hand is already the result of some previous scalarization step. Additionally, if we consider that the previous objectives are convex, it is guaranteed that the solution of the single-objective problem is Pareto efficient (Wendell and Lee 1977).

Therefore, for the reminder of this section, we are interested in solving the optimization problem described by (4.2), rewritten as

$$\min_{x} : \nu = J(y) \tag{5.1a}$$

$$\text{s.t.} : y(t) = h(x) \tag{5.1b}$$

with objective function $J : \mathbb{R}^{n_y} \mapsto \mathbb{R}$ and output map $h : \mathbb{R}^n \mapsto \mathbb{R}^y$. Once again, to ease the notation and without loss of generality, we combine the objective function with the output map, such that $y(t) = h(x) \in \mathbb{R}$ represents both.

By differentiating the input-output map with respect to time, and setting $\dot{x} = u$, one obtains

$$\dot{x}(t) = u(t) \tag{5.2a}$$

$$\dot{y}(t) = \frac{\partial h(x)}{\partial x}^{\mathsf{T}} u(t) \tag{5.2b}$$

where $u(t) \in \mathbb{R}^n$ is a control signal yet to be specified. To ensure that output-

feedback controllers can be used without risking finite-time escape of any closed-loop signal, we assume that Assumption 15 holds. Besides, to obtain global results, we assume the existence of a unique global minimum.

Assumption 26 (Unique Optimizer)**.** *Let $h \in C^2$. Then, the output $h(\cdot)$ is unimodal w.r.t. x_j, for every $j \in \{1, \ldots, n\}$, and there exists a unique minimizer x^*, with $y^* = h(x^*)$, such that*

$$\left.\frac{\partial h}{\partial x_j}\right|_{x=x^*} = 0 \quad \text{and} \quad \left.\frac{\partial^2 h}{\partial x_j^2}\right|_{x=x^*} > 0 \tag{5.3}$$

The objective of the controller we propose in this section is to remain as close as possible to the minimizer, as it was done for the controllers described in the previous sections. Once more, we describe this behavior using the concept of Δ-neighborhoods.

Definition 10 (Δ-neighborhood)**.** *The Δ-neighborhood of the minimizer is the region*

$$\mathcal{D}_\Delta = \left\{ x \in \mathbb{R}^n : \|x - x^*\| \leq \frac{\Delta}{2} \right\}. \tag{5.4}$$

and the Δ_j-neighborhoods are the regions

$$\mathcal{D}_{\Delta_j} = \left\{ x \in \mathbb{R}^n : \left|\frac{\partial h}{\partial x_j}\right|_{x \in \mathcal{D}_\Delta} \leq \left|\frac{\partial h}{\partial x_j}\right| \right\}. \tag{5.5}$$

Therefore, the Δ-neighborhood is the ball of diameter Δ around the minimizer x^*, while the Δ_j-neighborhoods are volumes along the curves $\partial h/\partial x_j = 0$. Furthermore, it also follows that $\mathcal{D}_\Delta = \bigcap_{j=1}^n \mathcal{D}_{\Delta_j}$.

Assumption 27 (Bounded Gradient)**.** *Outside the Δ-neighborhood, all elements of the HFG vector are bounded away from zero,*

$$L_j < \left|\frac{\partial h}{\partial x_j}\right|, \quad \forall x \notin \mathcal{D}_\Delta \tag{5.6}$$

where L_j are known arbitrarily small positive constants, with $j \in \{1, \ldots, n\}$. Furthermore, $L = \max_j L_j$.

Note that, Definition 10 with Assumption 27 ensure that the partial derivatives $\partial h/\partial x_j$ do not change sign when $x \notin \mathcal{D}_{\Delta_j}$, i.e.,

$$\operatorname{sign}\left(\frac{\partial h}{\partial x_j}\right) = -1 \quad \text{or} \quad \operatorname{sign}\left(\frac{\partial h}{\partial x_j}\right) = 1\,, \ \forall x \notin \mathcal{D}_{\Delta_j},\ j \in \{1, \ldots, n\} \tag{5.7}$$

Finally, to ensure that output-feedback controllers can be used and no signal escapes in finite time, we make the following assumption.

Assumption 28 (Unboundedness Observability (Angeli and Sontag 1999))**.** *The closed-loop system* (5.2) *possesses an unboundedness observability property, such that if any internal signal escapes in some finite time, then all other signals escape at the same time.*

5.1 The Class of Input-Output Mapping

The first derivative of the output y with respect to time can be written as in (4.6), where the high-frequency gain vector is written as

$$\frac{\partial h}{\partial x}(x) = \left[\frac{\partial h}{\partial x_1} \quad \dots \quad \frac{\partial h}{\partial x_n}\right]^{\mathsf{T}} = k_p(x) \tag{5.8}$$

Therefore, equation (5.2) is rewritten as

$$\dot{x}(t) = u(t) \tag{5.9a}$$

$$\dot{y}(t) = \sum_{j=1}^{n} \frac{\partial h(x)}{\partial x_j} u_j(t) \tag{5.9b}$$

5.2 Extremum-Seeking Controller

In this section, we describe our sliding-mode extremum-seeking-controller for static nonlinear plants with multiple decision variables and one objective function, which we later illustrate by considering a numerical simulation example.

The developed output-feedback based ESC law with *periodic switching function* is written as

$$u(t) = \rho(t)\ \text{sign}\left(\sin[\pi \mathcal{T} \sigma(t)]\right) \tag{5.10a}$$

$$\sigma(t) = e(t) + \lambda \int_0^t \text{sign}(e(\tau)) d\tau \tag{5.10b}$$

where $\rho(t) = \text{diag}\left(\ \rho_1(t)\,,\ \ \dots\ ,\ \ \rho_n(t)\ \right)$ is the *modulation function* (continuous in t) to be defined later on, $\mathcal{T} = \left[\ 1/T_1\,,\ \ \dots\ ,\ \ 1/T_n\ \right]^{\mathsf{T}}$, with $T_i \neq T_j$ for $i \neq j$, and $\lambda > 0$. The error signal is given by

$$e(t) = y(t) - y_m(t) \tag{5.11}$$

where $y_m : \bar{\mathbb{R}}_+ \mapsto \mathbb{R}$ is a ramp function, which decreases with time, such that

$\exists\, t_m : y_m(t) \leq y^*$, for all $t \geq t_m$. For analysis purposes, we define $y_m(t)$ as

$$y_m(t) = y_{m0} - K_m t \tag{5.12}$$

with design constants $y_{m0} \in \mathbb{R}$ and $K_m > 0$. Thus, the objective of the controller is to make y track y_{m0}, while $x \notin \mathcal{D}_\Delta$, such that y reaches an arbitrarily small neighborhood of the minimum y^*.

5.3 Modulation Function Design

5.3.1 Matched Input Disturbance

From equations (5.10) and (5.11), and the output dynamics (5.9), the σ-dynamics can be written as

$$\dot{\sigma}(t) = k_p(x)u(t) + d_\sigma(e) = \sum_{j=1}^{n} \left[\frac{\partial h(x)}{\partial x_j}\right] u_j(t) - K_m + \lambda \operatorname{sign}(e) \tag{5.13}$$

which is valid $\forall x$. Thus, with respect to the σ-dynamics, the matched input disturbance is given by

$$d_\sigma(e) = -K_m + \lambda \operatorname{sign}(e) \tag{5.14}$$

Furthermore, we can also rewrite (5.13) as

$$\dot{\sigma}(t) = \sum_{j=1}^{n} |k_{pj}(x)| \operatorname{sign}(k_{pj}(x)) u_j(t) + d_\sigma(e) \tag{5.15}$$

a format which will be useful later to prove the stability of the proposed controller.

5.3.2 Modulation Function Design

The modulation function must be designed in such a way that a sliding-mode is enforced on some sliding-surface $\sigma = kT_j$, for some integer k and one of the parameters T_j. From the previous sections, we already know that it is necessary to select ρ such that the modulation function overcomes the matched disturbances. Therefore, consider the following choice of modulation functions,

$$\rho_j(t) = \rho(t) = \frac{|d_\sigma(e) + \delta|}{\min_j L_j} \tag{5.16}$$

with L_j from Assumption 27 and $\delta > 0$, which is very similar to (2.16b), the modulation function for SISO systems.

5.4 Stability and Convergence Analysis

In this section, the stability and convergence results of the real-time multi-variable extremum-seeking controller for systems with multiple decision variables and one objective are carried out by showing that the existence of ideal sliding-modes guarantees the attractiveness of the Δ-neighborhood, which is reached in finite time.

5.4.1 Existence of a Family of Integers

Before presenting the main results, we first establish a relation between the parameters T_j. This relation is important to guarantee the existence of a set of σ-intervals of the form $[k_i, k_{i+1}]$, for integers k_i and k_{i+1}, which are attractive and invariant with respect to $\sigma(t)$, for each fixed configuration of $\text{sign}(k_{pj})$, $j \in \{1, \ldots, n\}$. These attractive intervals are such that the trajectory $\sigma(t)$ converge to different sliding manifolds as the output $y(t)$ approaches the optimal solution. In practice, this relation between the different parameters T_j does not need to be so strict, but we enforce it to obtain the convergence and stability results. Thus, consider

$$T_j = N_j T \tag{5.17}$$

with $N_j \in \mathbb{N}$ as defined in the following lemma.

Lemma 2. *For any real s, let $\lfloor s \rfloor$ denote the greatest integer lower or equal to s. Consider positive integers N_i $(i = 1, \ldots, n)$ satisfying the properties*

(a) $N_1 > \sum_{j=2}^{n} N_j > N_2 > \sum_{j=3}^{n} N_j > N_3 > \ldots > N_{n-1} > N_n$.

(b) $\left\lfloor \frac{N_j}{N_k} \right\rfloor = \frac{N_j}{N_k}$ *is an even integer when* $k > j$ $(j = 2, \ldots, n-1)$.

Then the integers

$$\left\lfloor \frac{m_1 N_1 + m_2 N_2 + \ldots + m_n N_n}{N_i} \right\rfloor \quad \text{and} \quad m_i$$

with $m_1, \ldots, m_n \in \{0, 1\}$, have the same parity for $i = 1, \ldots, n$.

Proof. First note that, for all $k \in \{1, \ldots, n-1\}$, one has

$$\sum_{j=k+1}^{n} m_j N_j \leq \sum_{j=k+1}^{n} N_j$$

since $m_j = 0$ or $m_j = 1$. Moreover, from property (a), one can write that

$\sum_{j=k+1}^{n} N_j < N_k$ and, thus

$$\frac{\sum_{j=k+1}^{n} m_j N_j}{N_k} < \frac{\sum_{j=k+1}^{n} N_j}{N_k} < 1$$

and

$$\frac{\sum_{j=1}^{n} m_j N_j}{N_k} = \frac{\sum_{j=1}^{k-1} m_j N_j}{N_k} + m_k + \frac{\sum_{j=k+1}^{n} m_j N_j}{N_k} < \frac{\sum_{j=1}^{k-1} m_j N_j}{N_k} + m_k + 1$$

where the quantity

$$\frac{\sum_{j=1}^{k-1} m_j N_j}{N_k}$$

is an **even** integer according to property (b). In addition, since $\frac{\sum_{j=1}^{k-1} m_j N_j}{N_k} + m_k < \frac{\sum_{j=1}^{n} m_j N_j}{N_k}$, one can conclude that

$$\left\lfloor \frac{\sum_{j=1}^{n} m_j N_j}{N_k} \right\rfloor = \left\lfloor \frac{\sum_{j=1}^{k-1} m_j N_j}{N_k} \right\rfloor + m_k$$

and the integer $\left\lfloor \frac{\sum_{j=1}^{n} m_j N_j}{N_k} \right\rfloor$ has the same parity as m_k, recalling that $\frac{\sum_{j=1}^{k-1} m_j N_j}{N_k}$ is an **even** integer according to property (b). ■

Remark 11 (Possible Choice for N_j)**.** *One possible (conservative) choice for the parameters N_j that satisfies Lemma 2 is*

$$N_j = 2^{n-j}\,, \quad j = 1, \ldots, n\,.$$

The parameters m_j are not available to the control designer. Nonetheless, studying the changes of these parameters according to the sign $\text{sign}(\partial h/\partial x_j)$ of the partial derivatives is paramount in establishing the convergence results. Thus, consider two variations of $m_j(t) \in \{0, 1\}$, either

$$m_j(t) = \begin{cases} 1, & \text{sign}(k_{pj}(t)) = 1 \\ 0, & \text{sign}(k_{pj}(t)) = -1 \end{cases} \tag{5.18a}$$

or

$$m_j(t) = \begin{cases} 0, & \text{sign}(k_{pj}(t)) = 1 \\ 1, & \text{sign}(k_{pj}(t)) = -1 \end{cases} \tag{5.18b}$$

Lemma 3 (Families of Integers)**.** *Consider the partial derivatives $k_{pj} = \partial h/\partial x_j$ from the system (5.2b). Then, according to the definitions (5.18a) and (5.18b),*

there exist two families of integers $\mathcal{N}_m^+$ *and* $\mathcal{N}_m^-$*, such that*

$$(-1)^{\left\lfloor \frac{N}{N_j} \right\rfloor} \operatorname{sign}(k_{pj}) = \begin{cases} -1, & N \in \mathcal{N}_m^+ \\ 1, & N \in \mathcal{N}_m^- \end{cases} \tag{5.19}$$

Proof. From Lemma 2, the positive integers $\left\lfloor \frac{m_1N_1+m_2N_2+\ldots+m_nN_n}{N_i} \right\rfloor$ and m_i have the same parity. So, one can write

$$(-1)^{\left\lfloor \frac{N}{N_j} \right\rfloor} = (-1)^{m_j} = \begin{cases} 1\,, & m_j = 0 \\ -1\,, & m_j = 1 \end{cases} \tag{5.20}$$

where

$$N = 2mM + m_1N_1 + m_2N_2 + \ldots + m_nN_n \tag{5.21}$$

M is the least common multiple of the the numbers $\{N_1, \ldots, N_n\}$, and m is any arbitrary integer (positive or negative). In this sense, N represents a family of integers for each arbitrary m. When the integers m_j are all fixed, N represents a family of integers for each arbitrary m. Dividing N in (5.21) by N_j,

$$\left\lfloor \frac{N}{N_j} \right\rfloor = \frac{2mM}{N_j} + \left\lfloor \frac{m_1N_1 + m_2N_2 + \ldots + m_nN_n}{N_j} \right\rfloor \tag{5.22}$$

it follows that $\left\lfloor \frac{N}{N_j} \right\rfloor$ and $\left\lfloor \frac{m_1N_1+m_2N_2+\ldots+m_nN_n}{N_j} \right\rfloor$ have the same parity since $2mM/N_j$ is an even integer.

Now, considering the two particular choices of m_j from (5.18), define the two families of integers

$$\mathcal{N}_m^+(t) = \Bigg\{ N \in \mathbb{Z} : N = 2mM + m_1N_1 + \ldots + m_nN_n,\ m \in \mathbb{Z}, \\ m_j = \frac{1 + \operatorname{sign}(k_{pj}(t))}{2} \Bigg\} \tag{5.23a}$$

and

$$\mathcal{N}_m^-(t) = \Bigg\{ N \in \mathbb{Z} : N = 2mM + m_1N_1 + \ldots + m_nN_n,\ m \in \mathbb{Z}, \\ m_j = \frac{1 - \operatorname{sign}(k_{pj}(t))}{2} \Bigg\} \tag{5.23b}$$

Therfore, equation (5.20) can be further expanded in

$$(-1)^{\left\lfloor \frac{N}{N_j} \right\rfloor} = (-1)^{m_j} = \begin{cases} -1, & N \in \mathcal{N}_m^+ \text{ and } \operatorname{sign}(k_{pj}) = 1 \\ 1, & N \in \mathcal{N}_m^+ \text{ and } \operatorname{sign}(k_{pj}) = -1 \\ -1, & N \in \mathcal{N}_m^- \text{ and } \operatorname{sign}(k_{pj}) = -1 \\ 1, & N \in \mathcal{N}_m^- \text{ and } \operatorname{sign}(k_{pj}) = 1 \end{cases} \tag{5.24}$$

which when multiplied by $\operatorname{sign}(k_p)$ yields (5.19). ■

5.4.2 Existence of a Family of σ-Intervals

In this section, a possible choice for the parameters T_j is proposed. This choice guarantees that, for a fixed configuration of $\operatorname{sign}(k_p)$, there exists a family of σ-intervals of the form $[k_i, k_{i+1}]$ which are invariant with respect to $\sigma(t)$.

Lemma 4 (Family of σ-Intervals)**.** *Consider the σ-dynamics (5.15), written below to ease redability,*

$$\dot{\sigma}(t) = \sum_{j=1}^{n} |k_{pj}(x)| \operatorname{sign}(k_{pj}(x)) u_j(t) + d_\sigma(e) \tag{5.25}$$

and select T_j in the control input (5.10a) as

$$T_j = N_j T, \, \forall j \in \{1, \ldots, n\} \tag{5.26}$$

with $T > 0$ and the positive integers N_j selected according to Lemma 2. Then, for each permutation of $\operatorname{sign}(k_{pj})$, $j \in \{1, \ldots, n\}$, there exists a family of σ-intervals of the form $[k_i, k_{i+1}]$ which are invariant with respect to $\sigma(t)$.

Proof. Let $T_j = N_j T$ and expand equation (5.15) by including the control signal (5.10a) to obtain

$$\dot{\sigma}(t) = \sum_{j=1}^{n} |k_{pj}(t)| \, \rho_j(t) (-1)^{\left\lfloor \frac{\sigma(t)}{T N_j} \right\rfloor} \operatorname{sign}(k_{pj}(t)) + d_\sigma(t) \tag{5.27}$$

which holds $\forall x$. For $\sigma(t)/T$ in the neighborhood of $[N, N+\mu)$, where N is an integer of the family $N = 2mM + m_1 N_1 + m_2 N_2 + \ldots + m_n N_n$, for an arbitrary integer m,

$$N \leq \frac{\sigma(t)}{T} < N + \mu \tag{5.28}$$

and

$$\left\lfloor \frac{N}{N_j} \right\rfloor \leq \frac{N}{N_j} \leq \frac{\sigma(t)}{T N_j} < \frac{N}{N_j} + \frac{\mu}{N_j} \tag{5.29}$$

For a sufficiently small μ,

$$0 < \mu < \min_j \left\{ N_j \left(\left\lfloor \frac{N}{N_j} \right\rfloor + 1 - \frac{N}{N_j} \right) \right\} \tag{5.30}$$

which is independent of m, the following inequality holds $\forall j$,

$$\left\lfloor \frac{N}{N_j} \right\rfloor \leq \frac{N}{N_j} \leq \frac{\sigma(t)}{TN_j} < \frac{N}{N_j} + \frac{\mu}{N_j} \leq \left\lfloor \frac{N}{N_j} \right\rfloor + 1 \tag{5.31}$$

and we conclude that

$$\left\lfloor \frac{\sigma}{TN_j} \right\rfloor = \left\lfloor \frac{N}{N_j} \right\rfloor \tag{5.32}$$

Connecting the above equality with Lemma 3 and defining $N_m^+ \in \mathcal{N}_m^+$ and $N_m^- \in \mathcal{N}_m^-$, it follows that

$$(-1)^{\left\lfloor \frac{\sigma(t)}{TN_j} \right\rfloor} \operatorname{sign}(k_{pj}(t)) = \begin{cases} -1, & \sigma(t)/T \in [N_m^+,\, N_m^+ + \mu^+) \\ 1, & \sigma(t)/T \in [N_m^-,\, N_m^- + \mu^-) \end{cases} \tag{5.33a}$$

$$0 < \mu^+ < \min_j \left\{ N_j \left(\left\lfloor \frac{N_m^+}{N_j} \right\rfloor + 1 - \frac{N_m^+}{N_j} \right) \right\} \tag{5.33b}$$

$$0 < \mu^- < \min_j \left\{ N_j \left(\left\lfloor \frac{N_m^-}{N_j} \right\rfloor + 1 - \frac{N_m^-}{N_j} \right) \right\} \tag{5.33c}$$

Then, the so-called σ-intervals are the set $\mathcal{I}_\sigma$ defined as

$$\mathcal{I}_\sigma = \bigcup_{\forall m} I_\sigma(m) \tag{5.34a}$$

$$I_\sigma(m) = \begin{cases} [N_m^-,\, N_m^+] & , \quad N_m^- \leq N_m^+ \\ [N_m^-,\, N_{m+1}^+] & , \quad N_m^+ < N_m^- \end{cases} \tag{5.34b}$$

The set $\mathcal{I}_\sigma$ defines infinitely many closed intervals that, for a fixed $\operatorname{sign}(k_{pj})$, are attractive and invariant with respect to $\sigma(t)$. This conclusion follows from combining (5.27) with (5.33) and (5.34).

Moreover, for a fixed $\operatorname{sign}(k_{pj})$, the length of each interval $I_\sigma(m) \in \mathcal{I}_\sigma$ is constant and smaller or equal to $2M$ since

$$\left| N_m^+ - N_m^- \right| = \left| \left[m_1^+ N_1 + \ldots + m_n^+ N_n \right] - \left[m_1^- N_1 + \ldots + m_n^- N_n \right] \right| \leq 2M \tag{5.35}$$

■

Lemma 5 (Boundedness of σ for $x \notin \mathcal{D}_{\Delta_j}$)**.** *Let $[t_0, \bar{t}_0]$ be an interval such that $\forall t \in [t_0, \bar{t}_0]$, $x \notin \mathcal{D}_{\Delta_j}$, $\forall j$, i.e., the gradient does not change sign. Then, by using*

the modulation function design (5.16) *that ensures* $\sum_{j=1}^{n} |k_{pj}|\,\rho_j > |d_\sigma| + \delta$,

$$|\sigma(t) - \sigma(t_0)| \leq 2MT, \quad \forall t \in [t_0, \bar{t}_0] \tag{5.36}$$

Proof.

(i) If $\sigma(t_0) \in I_\sigma(m) \subset \mathcal{I}_\sigma$ for some m.

From Lemma 4, boundedness of $\sigma(t)$ must follow since near the lower bound of $I_\sigma(m)$ one has that $\dot{\sigma} \geq \delta$ and near the upper bound one has that $\dot{\sigma} \leq -\delta$.

(ii) If $\sigma(t_0) \notin \mathcal{I}_\sigma$.

The signal $\sigma(t)$ either stays outside $\mathcal{I}_\sigma$ $\forall t \in [t_0, \bar{t}_0]$ or there exists $t_0* \in [t_0, \bar{t}_0]$ such that $\sigma(t_0^*) \in I_\sigma(m) \subset \mathcal{I}_\sigma$ for some m. Thus, in both cases $|\sigma(t) - \sigma(t_0)| \leq 2MT$. ■

5.4.3 Attractiveness of Some $\mathcal{D}_{\Delta_j}$

From the previous demonstrations in this section, it is shown that, as long as $x \notin \mathcal{D}_{\Delta_j}$, the signal $\sigma(t)$ remains bounded. This result alone is enough to ensure that $|e(t)|$ decreases, $y(t)$ approaches the optimum value, and, thus, $x(t)$ shall reach $\mathcal{D}_{\Delta_j}$ for some $j \in \{1, \ldots, n\}$. To accompany this result, we must show that as $x(t)$ approaches $\mathcal{D}_{\Delta_j}$ none of the system signals escape in finite time. These properties are summarized in the following lemma.

Lemma 6 (Attractiveness of $\mathcal{D}_{\Delta_j}$)**.** *Consider system* (5.2)*, with control* (5.10)*,* (5.11)*,* (5.12)*, and* (5.16)*. Assume that, at $t = t_0$, $x(t_0) \notin \mathcal{D}_{\Delta_j}$ and let t_y^* be a finite time such that $y_m(t) < y^*$, for $t \geq t_y^*$. Then, $y(t)$ decreases at a rate $\dot{y}(t) \leq -(K_m + \lambda)$, for $t \geq t_y^*$ and $x \notin \mathcal{D}_{\Delta_j}$. Additionally, none of the signals, $x(t)$, $y(t)$, $u(t)$, and $\sigma(t)$, escape in finite time.*

Proof. We divide this proof in two steps. First, we show that finite-time escape is not possible. Second, we show that $y(t)$ decreases at the aforementioned rate.

(i) No Finite-Time Escape

When the modulation function is constant, the smoothness of the output function $h(\cdot)$ and the unbounded observability assumption Assumption 28 are enough to guarantee that finite-time escape does not occurs

(ii) Rate of Convergence

If finite-time escape is avoided, then there exists $t_{y^*} > 0$ so that for $t \geq t_{y^*}$, $y_m < y^* \iff \text{sign}\,(e) = 1$. Hence,

$$\sigma(t) = e(t) + \lambda \int_0^t \text{sign}(e(\tau))d\tau \tag{5.37}$$

satisfies

$$\sigma(t) = e(t) + \lambda \int_0^{t_{y^*}} \text{sign}(e(\tau))d\tau + \lambda \int_{t_{y^*}}^{t} (1)d\tau \tag{5.38}$$

$\forall t \geq t_{y^*}$. Therefore, one has that

$$\sigma(t) = y(t) - y_m(t) + \lambda t + C\,, \quad \forall t > t_{y^*} \tag{5.39}$$

where $C = \lambda[t_{y^*} + \int_0^{t_{y^*}} \text{sign}(e(\tau))d\tau]$ is a constant. One can further write

$$y(t) = \sigma(t) - (k_m + \lambda)t - C\,, \quad \forall t > t_{y^*} \tag{5.40}$$

In addition, since

$$\dot{\sigma}(t) = \dot{e}(t) + \lambda\,\text{sign}(e(t)) \tag{5.41}$$

one can write that

$$\dot{\sigma}(t) = \dot{e}(t) + \lambda\,, \quad \forall t \geq t_{y^*} \tag{5.42}$$

or, equivalently,

$$\dot{y} = \dot{\sigma}(t) - (k_m + \lambda)\,, \quad \forall t \geq t_{y^*} \tag{5.43}$$

assuring that y decreases when $\dot{\sigma} < k_m + \lambda$.

Therefore, y decreases with rate not lower than $k_m + \lambda$, that is, y must approach y^*. So, there exists at least one channel j such that $\text{sign}(k_{pj}(t))$ changes and, thus, x_j reaches $\mathcal{D}_{\Delta_j}$. Thus, after some finite time, there exists x_j inside of the $\mathcal{D}_{\Delta_j}$ region, for some value of j. ■

5.4.4 Dominant Configuration of Signs

To differentiate between the different channels, let us group the indices j in three disjoint sets, $\mathcal{P}$, $\mathcal{P}_\Delta$, and $\mathcal{P}_{\bar{\Delta}}$, such that $\mathcal{P} \cup \mathcal{P}_\Delta \cup \mathcal{P}_{\bar{\Delta}} = \{1, \ldots, n\}$.

$$\mathcal{P}(t) = \{j \in \mathbb{N} : L < |\partial h(t)/\partial x_j|\} \tag{5.44a}$$

$$\mathcal{P}_\Delta(t) = \{j \in \mathbb{N} : |\partial h(t)/\partial x_j| \leq L_j\} \tag{5.44b}$$

$$\mathcal{P}_{\bar{\Delta}}(t) = \{j \in \mathbb{N} : L_j < |\partial h(t)/\partial x_j| \leq L\} \tag{5.44c}$$

with L and L_j from Assumption 27, and p, p_Δ, and $p_{\bar{\Delta}}$ the number of elements in $\mathcal{P}$, $\mathcal{P}_\Delta$, and $\mathcal{P}_{\bar{\Delta}}$, respectively. The meaning of these sets is provided below.

$\mathcal{P}$ The collections of channels further away from their respective $\mathcal{D}_{\Delta_j}$ regions, for which the partial derivatives are larger than the largest minimum bound.

$\mathcal{P}_\Delta$ The collections of channels inside the $\mathcal{D}_{\Delta_j}$ regions, for which the partial derivatives are below their corresponding lower bounds L_j.

$\mathcal{P}_{\bar{\Delta}}$ The collections of channels close to the $\mathcal{D}_{\Delta_j}$ regions by an amount up to $L - L_j$, for which the partial derivatives are between the channel lower bound L_j and the maximum lower bound L.

So, the sign of k_{pj} is held constant $\forall j \in \mathcal{P} \cup \mathcal{P}_{\bar{\Delta}}$ and **can change only** for $j \in \mathcal{P}_\Delta$, only for the channels j for which $x \in \mathcal{D}_{\Delta_j}$.

Definition 11 (Dominant Configuration os Signs)**.** *At any given time instant t, there is a corresponding configuration of signs for the components $k_{pj}(t)$ of the gradient vector $k_p(t) = \partial h(t)/\partial x$. Given a particular configuration of signs, we define the* dominant configuration of signs *as the set*

$$\mathcal{D}(t) = \{\text{sign}(k_{pj}(t)) : j \in \mathcal{P}\} \tag{5.45}$$

containing all signs of partial derivatives of channels j for which $L < |\partial h(.)/\partial x_j|$[1]*.*

During some time interval $[t_0, t_1]$, $t_1 > t_0$, the inherited dominant configuration of signs is called *fixed* when the following items hold:

- $\mathcal{P} \neq \emptyset$ in $[t_0, t_1]$.
- If, for $t \in [t_0, t_1]$, some new channel (with index i) become a new member of $\mathcal{D}(t_0)$, then the sign of $k_{pi}(t)$ must be the same as in $t = t_0$.

Hence, the dominant configuration of signs $\mathcal{D}(t_0)$ changes to a new dominant configuration $\mathcal{D}(t_1)$, at $t = t_1$, if there exists at least one channel (with index k) belonging to $\mathcal{D}(t_1)$ such that $\text{sign}(k_{pk}(t_1)) \neq \text{sign}(k_{pk}(t_0))$.

Recall that, as long as all the partial derivatives retain their signs, $\text{sign}(k_{pj}(t))$ remains constant $\forall j$, there are two familes of integers $\mathcal{N}_m^-$ and $\mathcal{N}_m^+$, explained in Lemma 3, there exists a set $\mathcal{I}_\sigma$ of infinitely many disjoint sets $I_\sigma(m)$ that, according to Lemma 4, are invariant with respect to $\sigma(t)$. Furthermore, $\sigma(t)$ remains bounded according to Lemma 5.

In what follows, we provide a generalization of Lemma 5 by showing that the existence of a *fixed dominant configuration of signs* guarantees that $\sigma(t)$ remains bounded to an interval $[N_m^-, N_{m+1}^-]$, even when some channel $k \in \mathcal{P}_\Delta(t)$ changes sign of $k_{pk}(t)$ **inside** the $\mathcal{D}_{\Delta_k}$-neighborhood by crossing the level surface $\partial h(.)/\partial x_k = 0$.

[1] It is interesting to see this set as the one containing the signs of the partial derivatives of all the channels j that are far away from the level surfaces $\partial h(.)/\partial x_j = 0$

This also will ensure that y decreases as long as the dominant configuration of signs remains fixed.

Lemma 7 (Boundedness of σ for a Fixed Dominant Configuration of Signs)**.** *Consider σ-dynamics (5.13), with control (5.10), (5.11), (5.12), and (5.16). Let $t \in [t_0, \bar{t}_0]$ such that a fixed dominant configuration of signs $\mathcal{D}(t)$ is achieved for t in this interval. Then, for some m, it follows that $\sigma(t)/T \in I_\sigma(m) \subset \mathcal{I}_\sigma$ and, thus*

$$|\sigma(t) - \sigma(t_0)| \leq 2MT, \quad \forall t \in [t_0, \bar{t}_0] \tag{5.46}$$

Proof. First, rewrite the σ-dynamics as $(\forall x)$

$$\dot{\sigma} = \sum_{j \in \mathcal{P}} |k_{pj}|\, \rho_j (-1)^{\left\lfloor \frac{\sigma(t)}{TN_j} \right\rfloor} \operatorname{sign}(k_{pj}) + \sum_{j \in \mathcal{P}_{\bar{\Delta}}} |k_{pj}|\, \rho_j (-1)^{\left\lfloor \frac{\sigma(t)}{TN_j} \right\rfloor} \operatorname{sign}(k_{pj}) + d'_\sigma \tag{5.47}$$

where

$$d'_\sigma = \sum_{j \in \mathcal{P}_\Delta} |k_{pj}|\, \rho_j (-1)^{\left\lfloor \frac{\sigma(t)}{TN_j} \right\rfloor} \operatorname{sign}(k_{pj}) + d_\sigma \tag{5.48}$$

is regarded as a disturbance, which satisfies

$$|d'_\sigma| \leq \sum_{j \in \mathcal{P}_\Delta} L_j \rho_j + |d_\sigma| \tag{5.49}$$

since $|k_{pj}(t)| \leq L_j$, when $j \in \mathcal{P}_\Delta$. Moreover, one has that

$$(-1)^{\left\lfloor \frac{\sigma(t)}{TN_j} \right\rfloor} \operatorname{sign}(k_{pj}(t)) = \begin{cases} 1, & \sigma(t)/T \in [N_m^-, N_m^- + \mu^-) \\ -1, & \sigma(t)/T \in [N_m^+, N_m^+ + \mu^+) \\ -1, & \sigma(t)/T \in [N_{m+1}^+, N_{m+1}^+ + \mu^+) \end{cases}, \quad j \in \mathcal{P} \cup \mathcal{P}_{\bar{\Delta}} \tag{5.50a}$$

and

$$(-1)^{\left\lfloor \frac{\sigma(t)}{TN_j} \right\rfloor} \operatorname{sign}(k_{pj}(t)) \in \{-1, 1\} \quad j \in \mathcal{P}_\Delta \tag{5.50b}$$

Thus, for $\sigma(t) \in T[N_m^-, N_m^- + \mu^-)$, one can write

$$\dot{\sigma} = \sum_{j \in \mathcal{P}} |k_{pj}|\, \rho_j + \sum_{j \in \mathcal{P}_{\bar{\Delta}}} |k_{pj}|\, \rho_j + d'_\sigma \tag{5.51}$$

Since $j \in \mathcal{P} \implies 0 < L_j \leq L < |k_{pj}|$, and $\sum_{j\in\mathcal{P}_\Delta} |k_{pj}|\,\rho_j > 0$ when $\rho_j > 0$,

$$\dot{\sigma} \geq \sum_{j\in\mathcal{P}} L\rho_j + d'_\sigma\,. \tag{5.52}$$

Then, due to the design of the modulation function $\rho_i = \rho$, $i \in \{1,\ldots,n\}$, in (5.16), repeated below,

$$\rho = \frac{\bar{d}_\sigma + \delta}{\min_j L_j} = \frac{\bar{d}_\sigma + \delta}{L_m} \tag{5.53}$$

where $\bar{d}_\sigma > |d_\sigma|$ and $0 < L_m < \min\{L_1,\ldots,L_n\}$ one can conclude that

$$\sum_{j\in\mathcal{P}} L\rho_j = p(L\rho) \tag{5.54}$$

and

$$|d'_\sigma| \leq \rho \sum_{j\in\mathcal{P}_\Delta} L_j + |d_\sigma| \tag{5.55a}$$

or, equivalently,

$$-\rho \sum_{j\in\mathcal{P}_\Delta} L_j - |d_\sigma| \leq d'_\sigma \leq \rho \sum_{j\in\mathcal{P}_\Delta} L_j + |d_\sigma| \tag{5.55b}$$

leading to the conclusion that

$$\dot{\sigma} \geq p(L\rho) - \rho \sum_{j\in\mathcal{P}_\Delta} L_j - |d_\sigma| = \rho(Lp - \sum_{j\in\mathcal{P}_\Delta} L_j) - |d_\sigma|\,. \tag{5.56}$$

Therefore, by choosing L large enough and L_m small enough so that

$$Lp - \sum_{j\in\mathcal{P}_\Delta} L_j > L_m > 0 \tag{5.57}$$

then $\rho(Lp - \sum_{j\in\mathcal{P}_\Delta} L_j) > \rho L_m = \bar{d}_\sigma + \delta$ and

$$\dot{\sigma} \geq \rho(Lp - \sum_{j\in\mathcal{P}_\Delta} L_j) - |d_\sigma| \geq \delta \tag{5.58}$$

Analogously, when $\sigma(t) \in T[N_m^+, N_m^+ + \mu^+)$ or $\sigma(t) \in T[N_{m+1}^+, N_{m+1}^+ + \mu^+)$, one can further write

$$\dot{\sigma} = -\sum_{j\in\mathcal{P}} |k_{pj}|\,\rho - \sum_{j\in\mathcal{P}_{\bar{\Delta}}} |k_{pj}|\,\rho + d'_\sigma\,. \tag{5.59}$$

and, consequently,

$$\dot{\sigma} \leq -\sum_{j\in\mathcal{P}} |k_{pj}|\,\rho + d'_\sigma\,. \tag{5.60}$$

since $-\sum_{j\in\mathcal{P}_\Delta} |k_{pj}|\,\rho \leq 0$. In addition, $d'_\sigma \leq \rho\sum_{j\in\mathcal{P}_\Delta} L_j + |d_\sigma|$ and $|k_{pj}| > L$ for $j \in \mathcal{P}$, then one has that the following inequalities hold

$$\dot{\sigma} \leq -p(L\rho) + \rho \sum_{j\in\mathcal{P}_\Delta} L_j + |d_\sigma|\ , \tag{5.61}$$

and

$$\dot{\sigma} \leq -\rho(Lp - \sum_{j\in\mathcal{P}_\Delta} L_j) + |d_\sigma| \leq \underbrace{-\rho L_m + |d_\sigma|}_{=-\bar{d}_\sigma+|d_\sigma|-\delta} \leq -\delta\,. \tag{5.62}$$

Finally,

$$\dot{\sigma} \geq \delta\,,\ \forall x\,, \quad \sigma(t) \in T[N_m^-, N_m^- + \mu^-) \tag{5.63a}$$

and

$$\dot{\sigma} \leq \begin{cases} -\delta, & \sigma(t)/T \in [N_m^+, N_m^+ + \mu^+) \\ -\delta, & \sigma(t)/T \in [N_{m+1}^+, N_{m+1}^+ + \mu^+) \end{cases} \tag{5.63b}$$

which, together, imply $\sigma(t) \in I_\sigma(m) \subset \mathcal{I}_\sigma$, for $t \in [t_0, \bar{t}_0]$. ■

If the dominant configuration of signs remained fixed $\forall t \in [t_a, \infty)$, proving the convergence of y towards the minimum y^* would be a matter of applying Lemma 7. However, this cannot hold. To show that the fixed dominant configuration of signs changes, let $t_a > t_{y^*}$, such that

$$y(t) = \sigma(t) + (k_m + \lambda)t - C \tag{5.64a}$$

$$y(t_a) = \sigma(t_a) + (k_m + \lambda)t_a - C \tag{5.64b}$$

and, thus

$$\begin{aligned} y(t) - y(t_a) &= \sigma(t) - \sigma(t_a) - (k_m + \lambda)(t - t_a) \\ |y(t) - y(t_a)| &\leq |\sigma(t) - \sigma(t_a)| - (k_m + \lambda)(t - t_a) \\ |y(t) - y(t_a)| &\leq 2MT - (k_m + \lambda)(t - t_a) \end{aligned} \tag{5.65}$$

So, $y \to y^*$ since $y(t)$ reduces as long as the dominant configuration of signs remains fixed, and, consequently, for every channel j, $x \to \mathcal{D}_{\Delta_j}$ until $\mathcal{P} = \emptyset$. Therefore,

there exists $t_b > t_a$ such that the dominant configuration of signs changes.

5.4.5 Convergence Towards the Optimum

Having already established Lemmas 5 and 7, the main proof of this section, convergence of the output towards the optimum, is ready to be presented. For that, the challenge that remains is showing that $y \to y^*$, even though the dominant configuration of signs might change over time. For that, we show that, but for oscillations of order $\mathcal{O}(T)$, the decision variables $x(t)$ approach the optimum $\mathcal{D}_\Delta$ and remain arbitrarily close to this set thereafter.

Theorem 8. *Consider system (5.2), with control (5.10), (5.11), (5.12), and (5.16), all repeated here for simplicity.*

$$\dot{x}(t) = u(t) \tag{5.66a}$$

$$\dot{y}(t) = \frac{\partial h(x)}{\partial x}^{\mathsf{T}} u(t) \tag{5.66b}$$

$$e(t) = y(t) - y_m(t) \tag{5.66c}$$

$$y_m(t) = y_{m0} - K_m t \tag{5.66d}$$

$$\rho_j(t) = \rho(t) = \frac{|d_\sigma(e) + \delta|}{\min_j L_j} \tag{5.66e}$$

Recall $\mathcal{D}_\Delta = \bigcap_{\forall j} \mathcal{D}_{\Delta_j}$ from Definition 10. Then, by applying the control algorithm as described above, it follows that $\mathcal{D}_\Delta$ becomes attractive, $\mathcal{D}_\Delta$, and $y(t) \to y^*$. Furthermore, oscillations of the vector of decision variables $x(t)$ away from the minimizer are bounded, $\|x(t) - x^*\| \leq \Delta/2 + \mathcal{O}(T)$.

Proof.

(i) Attractiveness of $\mathcal{D}_\Delta$

Assume that we have some dominant configuration of signs $\mathcal{D}(t_{c_1})$, at $t = t_{c_1} > t_{y^*}$. This dominant configuration of signs changes only when, due to changes in one channel $j \in \mathcal{P}_\Delta$, $x(t)$ leaves the $\mathcal{D}_{\Delta_j}$-neighborhood and, also, the $|\partial h(t)/\partial x_j|$ becomes greater than L at $t = t_{c_2} > t_{c_1}$. So, recalling the relation between L and Δ defined in Definition 10 and Assumption 27,

$$|x_j(t_{c_2}) - x_j(t_{c_1})| \geq \Delta \tag{5.67}$$

Moreover, recalling that $\dot{x}_j = u_j$ and $|u_j| = \rho$, one has that:

$$x_j(t) = x_j(t_{c_1}) + \int_{t_{c_1}}^{t} u_j(\tau) d\tau\,, \quad t \in [t_{c_1}, t_{c_2}] \tag{5.68}$$

when σ is not in sliding motion, and

$$|x_j(t) - x_j(t_{c_1})| \leq \rho(t_{c_2} - t_{c_1}) \tag{5.69}$$

leading to the conclusion that the time required for changing the dominant configuration has the lower bound

$$t_{c_2} - t_{c_1} \geq \frac{\Delta}{\rho} \tag{5.70}$$

During this time interval $[t_{c_1}, t_{c_2}]$, the dominant configuration is fixed so that $\sigma(t) \in I_\sigma(m) \subset \mathcal{I}_\sigma$ is kept in an interval of length up to $2M$ and y decreases. Indeed, with $t_{y^*} \leq t_{c_1} \leq t \leq t_{c_2}$,

$$[y(t) - y(t_{c_1})] = [\sigma(t) - \sigma(t_{c_1})] + (k_m + \lambda)(t - t_{c_1}) \tag{5.71}$$

with σ satisfying

$$|\sigma(t) - \sigma(t_{c_1})| \leq 2MT \tag{5.72}$$

Hence, since $|\sigma(t) - \sigma(t_{c_1})| \leq 2MT$ one has that

$$-2MT \leq \sigma(t) - \sigma(t_{c_1}) \leq 2MT \tag{5.73}$$

and

$$[\sigma(t) - \sigma(t_{c_1})] - (k_m + \lambda)(t - t_{c_1}) \geq -2MT - (k_m + \lambda)(t - t_{c_1}) \tag{5.74}$$

from which one can conclude that

$$[y(t) - y(t_{c_1})] = [\sigma(t) - \sigma(t_{c_1})] - (k_m + \lambda)(t - t_{c_1}) \leq 0 \tag{5.75}$$

or, equivalently,

$$y(t) < y(t_{c_1}) \tag{5.76}$$

for

$$(t - t_{c_1}) \geq \frac{2M}{(k_m + \lambda)} T \tag{5.77}$$

It mean that y decreases after an interval of order $\mathcal{O}(T)$. Moreover, at $t = t_{c_2}$, one

has that

$$[y(t_{c_2}) - y(t_{c_1})] \geq -2MT + \frac{\Delta(k_m + \lambda)}{\rho} > 0 \tag{5.78}$$

since

$$(t_{c_2} - t_{c_1}) \geq \frac{\Delta}{\rho} \geq \frac{2M}{(k_m + \lambda)} T \tag{5.79}$$

for T sufficiently small such that

$$T \leq \frac{\Delta}{\rho} \frac{(k_m + \lambda)}{2M} = \frac{\Delta L_m}{(k_m + \lambda + \delta)} \frac{(k_m + \lambda)}{2M} \tag{5.80}$$

After $t = t_{c_2}$, $\sigma(t)$ moves towards a an interval $I_\sigma(m_2) \subset \mathcal{I}_\sigma$, reaching it at $t = t_{\sigma_2}$, and y might increase along the time interval $[t_{c_2}, t_{\sigma_2}]$. In fact, there are three possibilities:

(1) $\sigma(t)$ stays in the same interval $I_\sigma(m)$ $\forall t$.

(2) $\sigma(t)$ *decreases* towards a new interval $I_\sigma(m_2)$ while $t \in [t_{c_2}, t_{\sigma_2}]$.

(3) $\sigma(t)$ *increases* towards a new interval $I_\sigma(m_2)$ while $t \in [t_{c_2}, t_{\sigma_2}]$.

In the first two cases, y decreases and moves towards the minimum y^*. Focusing on the third case, assume that $\sigma(t) > \sigma(t_{c_2})$, $\forall t \in [t_{c_2}, t_{\sigma_2}]$. For what follows, recall the demonstration of Lemma 6. If the time derivative of $\sigma(t)$ is such that $\dot{\sigma}(t) \leq k_m + \lambda$, then $\dot{y} = \dot{\sigma} - (k_m + \lambda) \leq 0$, $t_{y^*} < t_{c_2} \leq t \leq t_{\sigma_2}$. Thus, y does not increase.

So, consider the case where $\dot{\sigma} > k_m + \lambda$. In this case, $\dot{y} = \dot{\sigma} - (k_m + \lambda) > 0$, $t_{y^*} < t_{c_2} \leq t \leq t_{\sigma_2}$. Moreover, since $\dot{\sigma} > k_m + \lambda$ one can write

$$\sigma(t_{\sigma_2}) - \sigma(t_{c_2}) \geq k_m + \lambda(t_{\sigma_2} - t_{c_2}) \tag{5.81}$$

Now, since the interval $I_\sigma(m_2)$ is located at most $2M$ apart, $|\sigma(t) - \sigma(t_{c_2})| \leq 2MT$ and, therefore,

$$(k_m + \lambda)(t_{\sigma_2} - t_{c_2}) \leq \sigma(t_{\sigma_2}) - \sigma(t_{c_2}) \leq 2MT \tag{5.82}$$

and

$$(t_{\sigma_2} - t_{c_2}) \leq \frac{2M}{(k_m + \lambda)} T \tag{5.83}$$

i.e., $t_{\sigma_2} - t_{c_2}$ is of order $\mathcal{O}(T)$. Consequently,

$$[y(t_{\sigma_2}) - y(t_{c_2})] = [\sigma(t_{\sigma_2}) - \sigma(t_{c_2})] - (k_m + \lambda)(t_{\sigma_2} - t_{c_2}) \tag{5.84}$$

implies

$$|y(t_{\sigma_2}) - y(t_{c_2})| \leq |\sigma(t_{\sigma_2}) - \sigma(t_{c_2})| + (k_m + \lambda)(t_{\sigma_2} - t_{c_2}) \leq 4MT \tag{5.85}$$

and y can increase only during the interval $[t_{c_2}, t_{\sigma_2}]$ and at most by a value of order $\mathcal{O}(T)$.

Taking $t = t_{\sigma_2}$ as the initial instant, and since $t_{\sigma_2} - t_{c_2}$ is of order $\mathcal{O}(T)$, one has that the next dominant configuration of signs change occurs (otherwise y decreases until it reaches the minimum) at some $t = t_{c_3} > t_{\sigma_2}$ such that

$$t_{\sigma_2} - t_{c_2} < \frac{\Delta}{\rho} \leq (t_{c_3} - t_{c_2}) \tag{5.86}$$

since $t_{c_3} - t_{c_2}$ is greater that $t_{\sigma_2} - t_{c_2}$, for T sufficiently small. With similar arguments, one can verify that y decreases during the interval $[t_{\sigma_2}, t_{c_3}]$, i.e.,

$$y(t_{c_3}) < y(t_{\sigma_2}) \leq y(t_{c_1}) \tag{5.87}$$

and can increase during the interval $[t_{\sigma_2}, t_{\sigma_3}]$ not more than a value of order $\mathcal{O}(T)$. This drives the conclusion that y approaches the extremum or the dominant configuration of signs disappear $\mathcal{P} = \emptyset$. In both cases, $x(t)$ reaches a neighborhood of order $\mathcal{O}(T)$ around $\mathcal{D}_\Delta$. Hence, it follows that $\mathcal{D}_\Delta$ is attractive.

(ii) Oscillations of order $\mathcal{O}(T)$

When some channel j escapes from this neighborhood of order $\mathcal{O}(T)$ around $\mathcal{D}_\Delta$, such that $\mathcal{P} \neq \emptyset$, a new dominant configuration of signs is obtained, guaranteeing that y diverges from y^* not more than a value of order $\mathcal{O}(T)$. ■

5.5 Numerical Simulation Example

To illustrate the proposed extremum-seeking controller, consider the optimization of the Rosenbrock function,

$$h(x) = 50 \sum_{k=1}^{N/2} \left(x_{2k-1}^2 - x_{2k}\right)^2 + (x_{2k-1} - 1)^2 \tag{5.88}$$

with $N = 2$, see fig. 5.1. It is possible to optimize functions of higher dimension, but we illustrate the problems with $N = 2$ because it is easier to visualize the results. We have chosen to optimize the Rosenbrock function because it is a challenging objective function, with a global minimum at $x^* = [1\ 1]^\mathsf{T}$.

We consider two simulation scenarios. In the first, we run the simulation with the control law as described in this section. In the second, we modify the control

law to implement a continuous unit-vector control, rewriting (5.10) as

$$u_j(t) = \rho_j(t) \frac{s(t)}{\max\left(\|s(t)\|, \epsilon\right)} \tag{5.89}$$

$$s(t) = \sin\left[\frac{\pi\sigma(t)}{T_j}\right] \tag{5.90}$$

The small parameter $\epsilon = 0.1$ is used to avoid the discontinuity when $s \to 0$. For both controllers, we set $\rho_1 = \rho_2 = 0.5$, $x(0) = \begin{bmatrix} -1 & -1 \end{bmatrix}$, $\lambda = 0.1$, $T_1 = 0.1$, and $T_2 = 0.2$. We set the reference model

$$y_m(t) = \max\left(y(0)\left(1 - t/20\right), 0\right) = \max\left(200 - 10t, 0\right) \tag{5.91}$$

such that the output should reach a small neighborhood of the minimum at $t = 20$. The sampling period is fixed at 10^{-4}.

The simulation results are shown in figs. 5.1 and 5.2. In both cases, it is possible to verify that the controllers can drive the output towards the minimum value. The performance is similar, but it is interesting to note that, in this case, the unit-vector control formulation outperforms the approach via sign function. This result is far from a generalization, and simulations with other objective functions display different results. Nonetheless, it is interesting to show that, depending on the problem at hand, one might benefit from other control formulations.

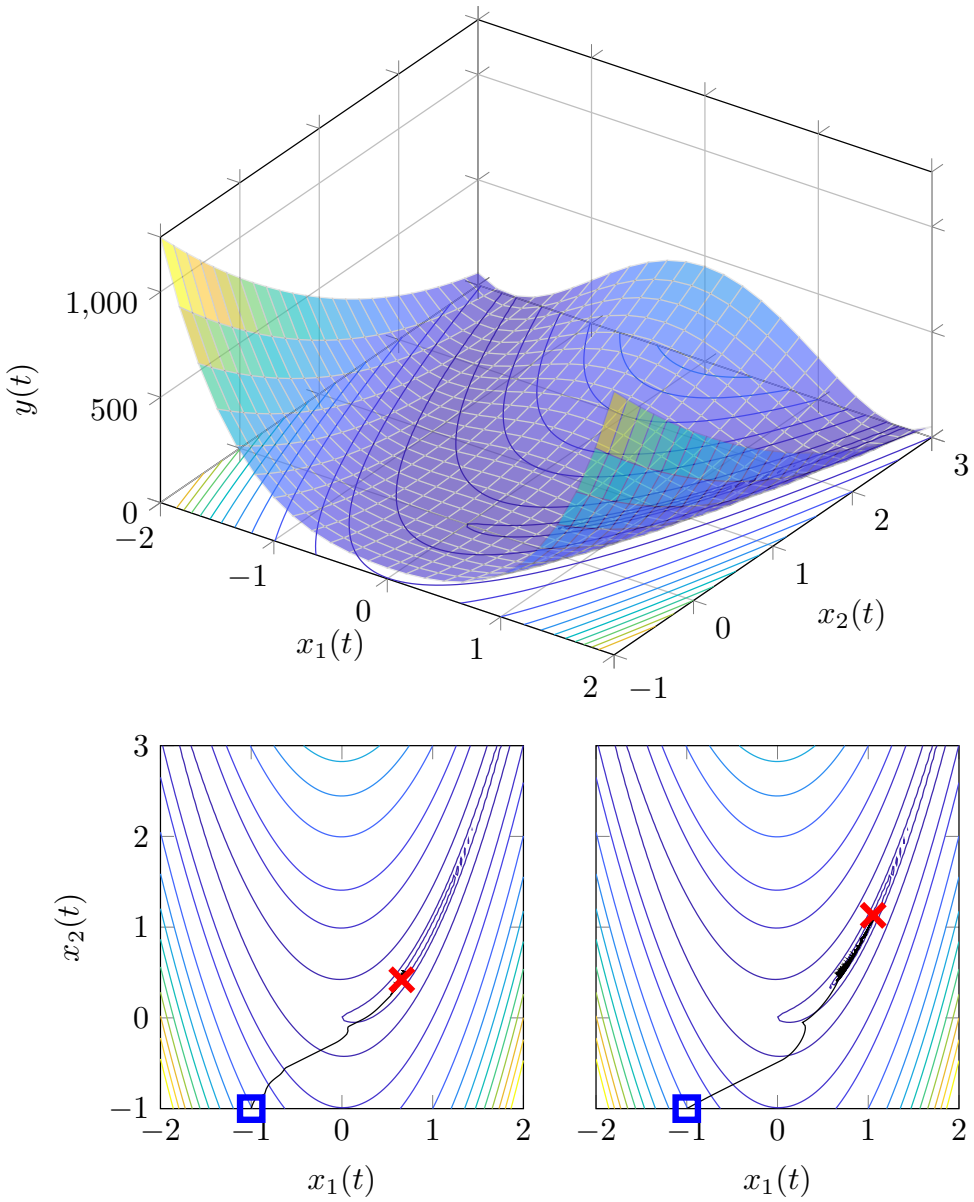


Figure 5.1: Two-dimensional Rosenbrock function at the top, and the systems trajectories when solving the two-dimensional optimization problem of section 5.5 using (left) the control with sign function (5.10) and (right) unit-vector control with a continuous switching element (5.89).

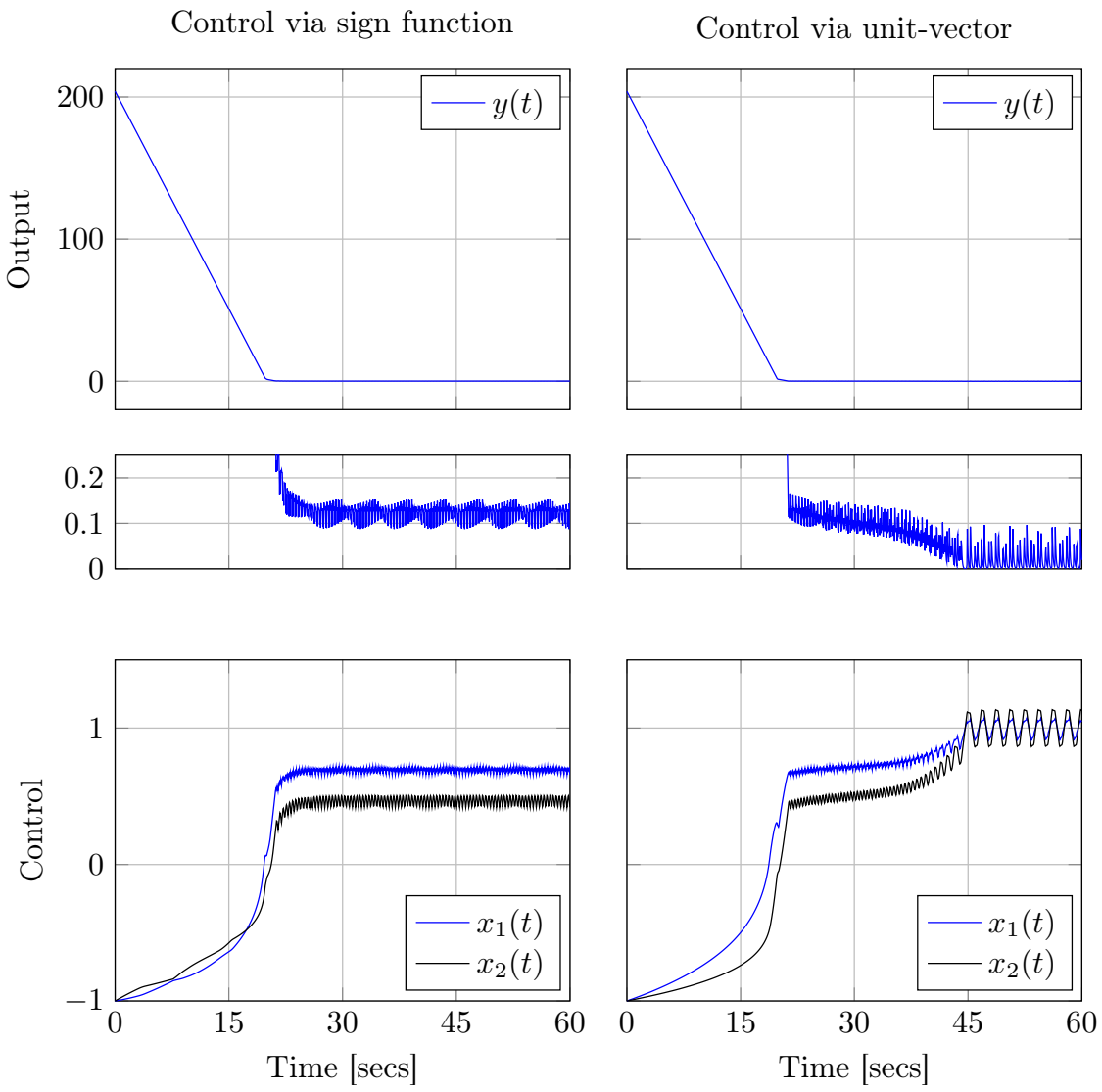


Figure 5.2: Output (objective function value) and control signals (decision variables) when solving the two-dimensional optimization problem of section 5.5 using (left) the control with sign function (5.10) and (right) unit-vector control with a continuous switching element (5.89).

Chapter 6

Dynamic Estimator for Extremum Consensus

Consensus estimation consists in applying an update rule such that, in a distributed network where each node has its inputs and estimates, every node converges to a desired performance index, which is a function of the network inputs. To converge to this desired index, each node can exchange information with a finite number of neighbor nodes.

In distributed optimization problems, it is appealing to consider consensus estimation as a valuable tool to perform scalarization of the multiple objective functions, see (Poveda and Quijano 2013; Guay et al. 2018; Salamah et al. 2018). In these works, the authors consider the average consensus estimator of Freeman et al. (2006) to perform scalarization through weighted sums. In all three works, the final objective is to solve the single-objective optimization problem, obtaining then a Pareto efficient solution. The downside of this approach is that, if the set of Pareto efficient solutions is non-convex, not all solutions are reachable after solving the single-objective optimization problem.

To counter this drawback, in this Thesis, we propose a novel dynamic consensus estimator capable of finding the overall maximum (minimum) input in a network. Thus, applying the proposed algorithm in a distributed optimization problem, one may consider the nodes inputs as the individual objective functions, and by estimating the overall maximum (minimum) objective in the network, one may apply scalarization via Chebyshev's distance to find Pareto efficient solutions for classes of optimization problems with non-convex Pareto sets. Naturally, if all objective functions are convex, Chebyshev's distance can still be used to find all Pareto efficient solutions, similarly to how weighted sums would be used.

6.1 Problem Formulation

There are two varieties of consensus estimators: static estimators and dynamic estimators. In *static consensus*, a snapshot of the nodes inputs at a given time is used to initialize the algorithm, but changes to these inputs are ignored thereafter. Static consensus is useful if the inputs vary slowly or do not vary at all. If the inputs vary through time, and such changes cannot be ignored, one should instead consider dynamic consensus. In *dynamic consensus*, algorithms are designed such that each node tracks the overall network performance index as it changes through time due to changes to the nodes inputs. In this section, we provide the mathematical background and the assumptions needed to develop our dynamic consensus estimator for maximum (minimum) estimation.

Consider a group of n labeled nodes, with labels belonging to the set $\mathcal{V} = \{1,\ 2,\ \dots\ ,\ n\}$, each one holding an input $u_j(t) \in \mathbb{R}$ and an estimate $x(t) \in \mathbb{R}$, both function of time. These nodes interact over a communication network, with topology represented by a directed graph $\mathcal{G} = (\mathcal{V},\ \mathcal{E})$, where $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ is the set of all edges which connect two nodes. It is said that a node i receives information from node j if and only if $(i, j) \in \mathcal{E}$. When this is the case, node i has access to the inputs and the estimates of node j at any given time t.

Many properties can be associated with graphs, but we are particularly interested in the following (see fig. 6.1 for illustrations of some of them).

Definition 12 (Undirected Path)**.** *An* undirected path *is a sequence of nodes* $i_1, i_2, \dots, i_p$ *such that either* $(i_j, i_{j+1}) \in \mathcal{E}$ *or* $(i_{j+1}, i_j) \in \mathcal{E}$*.*

Definition 13 (Weakly Connected)**.** *A graph is* weakly connected *if every pair of nodes lie on some undirected path.*

Definition 14 (Directed Cycle)**.** *A* directed cycle *is a sequence of nodes* $i_1, i_2, \dots, i_p$*, with* $i_1 = i_p$*, such that* $(i_j, i_{j+1}) \in \mathcal{E}$*.*

Definition 15 (Strongly Connected)**.** *A graph is* strongly connected *if every pair of nodes lie on some directed cycle. Considering undirected graphs, weakly connected* $\iff$ *strongly connected.*

Definition 16 (Directed Path)**.** *Similar to an undirected path, a* directed path *is a sequence of nodes* $i_1, i_2, \dots, i_p$ *such that* $(i_j, i_{j+1}) \in \mathcal{E}$*.*

Definition 17 (Shortest Path, Distance, Eccentricity, and Diameter)**.** *The* shortest path *between two nodes in a directed graph is the directed path with the least amount of edges. The* distance *between two nodes* $\delta(i, j)$ *is the number of edges in a shortest path, and the* eccentricity *of a node is its greatest distance to any other node. The*

diameter, *denoted $d(\mathcal{G})$, is the greatest eccentricity of all nodes in a graph.*

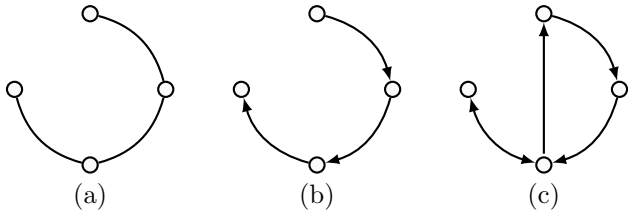


Figure 6.1: Examples of (a) an undirected strongly connected graph, (b) a weakly connected but not strongly connected directed graph, and (c) a strongly connected directed graph.

Assumption 29 (Graph Properties)**.** *For the remaining of this chapter, we use $\mathcal{G} = (\mathcal{V},\ \mathcal{E})$ to denote a directed and strongly connected graph, with nodes belonging to $\mathcal{V} = \{1,\ 2,\ \dots\ ,\ n\}$ and edges belonging to $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$.*

For a given node j,

$$\mathcal{U}_j(t) = \{u_i(t) \in \mathbb{R} : (j,i) \in \mathcal{E}\} \cup \{u_j(t)\} \tag{6.1a}$$

$$\mathcal{X}_j(t) = \{x_i(t) \in \mathbb{R} : (j,i) \in \mathcal{E}\} \cup \{x_j(t)\} \tag{6.1b}$$

are the sets containing all inputs and all estimates that node j is aware of, including its own input and estimate, which belong to the singletons $\{u_j(t)\}$ and $\{x_j(t)\}$.

Given a vector $v \in \mathbb{R}^p$, $v_j \in \mathbb{R}$ denotes its j-th component. Therefore, the vectors containing all inputs and all estimates in the network are $u(t), x(t) \in \mathbb{R}^n$, respectively.

Assumption 30 (Differentiable Input Vector)**.** *The input vector $u(t)$ is differentiable almost everywhere.*

Assumption 31 (Bounded Input Derivative)**.** *The input vector $u(t)$ is absolutely continuous and there is a known upper-bound $L > 0$ to its time derivative*

$$\max_{j \in \mathcal{V}} |\dot{u}_j(t)| \leq L\ ,\quad \forall t \tag{6.2}$$

which is defined almost everywhere.

Depending on the context, the function max(.) can denote either the maximum component of a vector or the maximum element in a set. Also, recall the sign function definition

$$\text{sign}(\zeta) = \begin{cases} -1 & ,\ \ \zeta < 0 \\ 1 & ,\ \ \zeta > 0 \end{cases} \tag{6.3}$$

However, we consider that if $\zeta = 0$ is not a sliding-surface, $\text{sign}(\zeta) = 0$ at $\zeta = 0$. Otherwise, $\text{sign}(\zeta)$ is undefined at $\zeta = 0$. We consider the following linear approximation of the sign function,

$$\underset{\epsilon}{\text{sigmoid}}(\zeta) = \begin{cases} \text{sign}(\zeta) & , \quad |\zeta| > \epsilon \\ \zeta/\epsilon & , \quad |\zeta| \leq \epsilon \end{cases} \tag{6.4}$$

for any positive scalar $\epsilon > 0$.

Considering the above definitions, the purpose of this chapter is to develop an algorithm such that all estimates x_j are driven toward the maximum input $\max(u)$ in the network. In other words, for $t > t^*$, $|x_j(t) - \max(u(t))| < \epsilon$, for all $j \in \mathcal{V}$ and for an arbitrarily small scalar $\epsilon > 0$, where the convergence time $t^* > 0$ can be made arbitrarily close to zero.

6.2 Dynamic Consensus Estimator

In this section, we formulate the problem and present the sliding-mode based maximum consensus algorithm.

Let $\mathcal{G}$ be the graph of a network such that Assumption 29 is satisfied. For this class of networks, consider the update rule

$$\tau \dot{x}_j = (1 + \alpha) \underset{\epsilon}{\text{sigmoid}}(\hat{e}_j) + \text{sign}(e_j) \tag{6.5a}$$

$$\hat{e}_j = \max(\mathcal{X}_j) - x_j \tag{6.5b}$$

$$e_j = \max(\mathcal{U}_j) - x_j \tag{6.5c}$$

where $\hat{e}_j(t) \in \mathbb{R}$ is the error between the node estimate and the maximum estimate it knows, $e_j(t) \in \mathbb{R}$ the error between the node estimate and the maximum input it knows, $\tau > 0$ is a scalar that controls the convergence rate, and $\alpha \in (0, 1)$ and $\epsilon > 0$ are other design parameters.

The first component of (6.5) is responsible for driving all estimates toward a common value, which is the highest in the entire network. The second component is responsible for driving the estimates toward the maximum input.

We assume that $u(t)$ is differentiable almost everywhere, and that its rate of change $\dot{u}(t)$ is bounded, such that Assumptions 30 and 31 are true. Assumption 31 is important to enable the control designer to select an appropriate value for the parameter τ in (6.5a).

Lemma 8 (Convergence to $\max(x)$)**.** *Let the graph $\mathcal{G}$ satisfy Assumptions 29 to 31, and let all nodes update their estimates through the update rule (6.5), starting at*

an initial time t_0. *Then, with* $\tau < \alpha/L$, *the network achieves a consensus and all estimates converge to* $\max(x) - x_j \leq \mathcal{O}(\epsilon)$ *after a finite time* $t_1 \leq t_0 + \mathcal{O}(\tau)$, *remaining therein for* $t \geq t_1$.

Proof. At any given time, a node's estimate is lower than or equal to the highest estimate that it is aware of, i.e. $x_j(t) \leq \max(\mathcal{X}_j(t))$, $\forall j \in \mathcal{V}$. Thus, considering also the ϵ-vicinity of $\max(\mathcal{X}_j)$, there are three possibilities for any node j: $x_j \leq \max(\mathcal{X}_j) - \epsilon$, $\max(\mathcal{X}_j) - \epsilon < x_j < \max(\mathcal{X}_j)$, and $x_j = \max(\mathcal{X}_j)$.

$\boldsymbol{x_j \leq \max(\mathcal{X}_j) - \epsilon}$**.** In this case, it follows that $\hat{e}_j \geq \epsilon \iff \text{sigmoid}_\epsilon(\hat{e}_j) = 1$, and (6.5a) becomes

$$\tau \dot{x}_j = 1 + \alpha + \text{sign}(e_j) \geq \alpha \tag{6.6}$$

with a solution

$$x_j(t) \geq x_j(t_0') + (\alpha/\tau)(t - t_0') \tag{6.7}$$

where $t_0' > t_0$ is a time instant at which x_j enters the region $x_j \leq \max(\mathcal{X}_j) - \epsilon$. Therefore, while x_j is not close to $\max(\mathcal{X}_j)$ it increases at a rate α/τ.

$\boldsymbol{x_j = \max(\mathcal{X}_j)}$**.** In this case, $\hat{e}_j = 0 \iff \text{sigmoid}_\epsilon(\hat{e}_j) = 0$, and (6.5a) becomes

$$\tau \dot{x}_j = \text{sign}(e_j) \tag{6.8}$$

such that the dynamics of e_j becomes

$$\dot{e}_j = \tau \frac{d}{dt} \max(\mathcal{U}_j) - \text{sign}(e_j) \tag{6.9}$$

Therefore, as long as $\tau < 1/L$, $e_j = 0$ is a sliding-surface while $x_j = \max(\mathcal{X}_j)$. Furthermore, to guarantee that the neighbors i of node j for which $\max(\mathcal{X}_i) = x_j$ have estimates converging to x_j, one must have $\alpha/\tau > L \iff \tau < \alpha/L$. This conclusion follows from (6.7).

These two cases are enough to show that $x_j \to \max(x)$, $\forall j \in \mathcal{V}$, entering an ϵ-vicinity of $\max(x)$ in finite time, since all estimates are either increasing, according to (6.7) or bounded by their highest known input until another estimate exceeds it. When this happens, the exceeded estimate, for instance x_i, must track its current maximum known estimate $\max(\mathcal{X}_i)$. This process repeats itself and propagates through the network until all estimates reach $x_j \geq \max(x) - d(\mathcal{G})\,\epsilon$, converging in a finite time

$$t_1 \leq t_0 + \Delta \left(\frac{\alpha}{\tau} - L\right)^{-1} \leq t_0 + \mathcal{O}(\tau) \tag{6.10a}$$

$$\Delta = \max\left(\begin{bmatrix} u(t_0) \\ x(t_0) \end{bmatrix}\right) - \min(x(t_0)) \tag{6.10b}$$

The propagation is guaranteed because $\mathcal{G}$ is strongly connected. The factor $d(\mathcal{G})$ acts as a worst case bound because the estimate error might propagate through the network from $\max(x)$ to the node furthest from it. This distance is, at most, equal to the graph diameter $d(\mathcal{G})$. ■

From the perspective of consensus, Lemma 8 is enough to show that the network reaches a consensus with the proposed technique. It does not, however, establish any relationship between the nodes inputs and the consensus value. The next theorem shows that the update rule (6.5) is able to enforce tracking of the maximum network input.

Theorem 9 (Convergence to $\max(u)$)**.** *Let the graph $\mathcal{G}$ satisfy Assumptions 29 to 31, and let all nodes update their estimates through the update rule (6.5), starting at an initial time t_0. Then, with $\tau < \alpha/L$, the network achieves a consensus and all estimates converge to $|\max(u) - x_j| < \mathcal{O}(\epsilon)$ after a finite time $t^* \leq t_0 + \mathcal{O}(\tau)$, remaining therein for $t \geq t^*$.*

Proof. From Lemma 8 we already know that $\max(\mathcal{X}_j) - x_j \leq \epsilon$ and $\max(x) - x_j \leq d(\mathcal{G})\,\epsilon$ hold $\forall j \in \mathcal{V}$ for $t \geq t_1$. Thus, it suffices to show that $\max(x) \to \max(u)$. The proof is then split into two parts. First, we show that if $\max(x) < \max(u)$, all estimates increase. Otherwise, if $\max(x) > \max(u)$, all estimates decrease.

$\max(x) < \max(u)$. Let $j : u_j = \max(u)$, which implies $u_j = \max(\mathcal{U}_j)$, and let

$$k \in \{i \in \mathcal{V} : (j, i) \in \mathcal{E}\} \cup \{j\} \tag{6.11}$$

that is, k correspond to all nodes that node j has access to, including itself. Considering these nodes and $t > t_1$, such that $\max(\mathcal{X}_j) - x_j \leq \epsilon \iff \text{sigmoid}_\epsilon(\hat{e}_j) = \hat{e}_j/\epsilon$,

$$\tau \dot{x}_k = \frac{1+\alpha}{\epsilon}\,\hat{e}_k + \text{sign}(e_k) \tag{6.12}$$

Since $\max(x) < \max(u)$, then $\text{sign}(e_k) = \text{sign}(\max(u) - x_k) = 1$. Furthermore, from its definition, it follows that $\hat{e}_k \geq 0$, and equation (6.12) can be converted to the inequality

$$\tau \dot{x}_k \geq 1 \tag{6.13}$$

which yields

$$x_k(t) \geq x_k(t_2) + (t - t_2)/\tau \tag{6.14}$$

where $t_2 \geq t_1$ is any time instant for which $\max(x) < \max(u)$. Hence, for $t > t_1$, $\max(\mathcal{X}_j)$ must reach $\max(u)$ in a finite time

$$t_1^* \leq t_1 + \Delta^* \left(\frac{1}{\tau} - L\right)^{-1} \tag{6.15a}$$

$$\Delta^* = |\max(u(t_1)) - \max(\mathcal{X}_j(t_1))| \tag{6.15b}$$

From Lemma 8, $|x_j - \max(u)| \leq d(\mathcal{G})\,\epsilon$, $\forall j \in \mathcal{V}$, is reached in a finite time bounded by (6.15).

$\max(x) > \max(u)$. Consider the error function

$$e = \max(x) - \max(u) = 0 \tag{6.16}$$

with time derivative

$$\tau\dot{e} = \text{sign}(e_i) - \tau \frac{d}{dt} \max u(t) \tag{6.17}$$

where $i : x_i = \max(x)$ and $\text{sigmoid}_\epsilon(x_i) = 0$ was omitted because x_i is not in sliding-mode. Since $\max(x) > \max(u)$, then $\text{sign}(e_i) = -1$. Therefore, the following is valid whenever $\max(x) > \max(u)$:

$$\tau\dot{e} \leq \tau L - 1 \tag{6.18a}$$

$$e(t) \leq e(t_3) - \left(\frac{1}{\tau} - L\right)(t - t_3) \tag{6.18b}$$

where $t_3 \geq t_1$ is any time instant for which $\max(x) > \max(u)$. Hence, for $t > t_1$ and $\max(x) > \max(u)$, the maximum estimate $\max(x)$ reaches $\max(u)$ in a finite time

$$t_2^* \leq t_1 + \Delta^* \left(\frac{1}{\tau} - L\right)^{-1} \tag{6.19}$$

which equals the bound (6.15) for the previous case. Once again, invoking Lemma 8, $|x_j - \max(u)| \leq d(\mathcal{G})\,\epsilon$, $\forall j \in \mathcal{V}$, is reached in a finite time bounded by (6.19).

Finally, since it was shown that $\max(x)$ is driven toward $\max(u)$, we conclude from (6.10), (6.15), and (6.19) that $|\max(u) - x_j| \leq d(\mathcal{G})\,\epsilon \leq \mathcal{O}(\epsilon)$ is reached in a finite time

$$t^* \leq t_0 + t_1 + \Delta^* \left(\frac{1}{\tau} - L\right)^{-1} \leq t_0 + \mathcal{O}(\tau) \tag{6.20}$$

■

Remark 12 (Finite Time Upper-Bounds)**.** *It is worth pointing out that both conver-*

gence time upper-bounds (6.10) *and* (6.20) *are very conservative, since, to compute them, it is assumed that the inputs are always growing at their maximum possible rate* $L \geq \max_{j\in\mathcal{V}} |\dot{u}_j|$.

Remark 13 (Application to Minimum Consensus Estimation)**.** *The proposed consensus estimator can be easily applied to the problem of finding the minimum overall input in a network, instead of the maximum overall input. For that, one need only change the* max(.) *functions by* min(.) *functions, and no other modification is needed, such that the consensus dynamics is written as follows.*

$$\tau\dot{x}_j = (1+\alpha)\ \underset{\epsilon}{\text{sigmoid}}(\hat{e}_j) + \text{sign}(e_j) \tag{6.21a}$$

$$\hat{e}_j = \min\left(\mathcal{X}_j\right) - x_j \tag{6.21b}$$

$$e_j = \min\left(\mathcal{U}_j\right) - x_j \tag{6.21c}$$

Theorem 9 guarantees that all estimates track the maximum network input, with an arbitrarily small error of order $\mathcal{O}(\epsilon)$. Naturally, in practical discrete-time implementations, even though the theory ensures arbitrarily fast convergence rates and small tracking errors, there is a tradeoff between improving these values and selecting an appropriate sampling period. The faster the system dynamics, the smaller the sampling period. Likewise, the smaller the desired tracking error, the smaller the sampling period. To help implementing the proposed consensus algorithm (6.5), we highlight some implementation guidelines.

6.2.1 Discrete-Time Implementation Guidelines

To avoid undesired chattering, we have experienced better results using the trapezoidal integration rule and using $\text{sigmoid}_\epsilon(e_j)$ instead of $\text{sign}(e_j)$ in (6.5a). For the initial states, we suggest using $x_j(t_0) = u_j(t_0)$. Let $p_{\text{error}} > 0$ and $p_{\text{rate}} > L$ denote the desired maximum error and minimum convergence rate, with L from Assumption 31. Using these specifications, the control parameters are defined as

$$\epsilon = p_{\text{error}}/d(\mathcal{G}) \tag{6.22a}$$

$$\tau = \alpha/p_{\text{rate}} \tag{6.22b}$$

Although we let $\alpha \in (0,1)$, we usually select $\alpha = 0.5$.

There are two basic rules to select the sampling period. It can be either a function of the convergence rate or a function of the desired tracking error. To ensure convergence and that both specification are met, the sampling period t_s

should be

$$t_s \leq \min\left(\epsilon\ ,\ \tau/100\right) \tag{6.23}$$

Naturally, if the sampling period t_s is pre-defined, good choices of ϵ and τ are

$$\epsilon \geq t_s \tag{6.24a}$$

$$\tau \geq 100\, t_s \tag{6.24b}$$

There is a margin on τ, such that the 10^2 factor can be relaxed to 10 without much impact on performance.

Finally, we stress that it is essential to select an appropriate sampling period. Otherwise, there might be convergence issues that may hinder the algorithm performance. If reducing the sampling period or increasing τ are not viable options, one might actually lower the parameter τ. Although this seems counterintuitive, it mitigates the problem, at the expense of increasing chattering.

6.2.2 Numerical Simulation Example

In this section, we illustrate the proposed consensus algorithm properties through two numerical simulations. The first one serves to illustrate the algorithm tracking performance and its convergence properties, while the second displays its robustness to the network size. The graphs topologies for these simulations are illustrated in fig. 6.2. Note that, regarding the network connectivity, the second topology is the worst possible, since every node has the same eccentricity, which equals the network diameter $d(\mathcal{G}_2) = 99$.

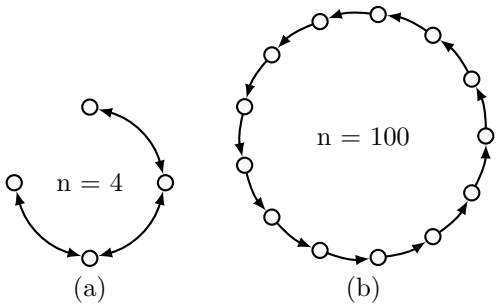


Figure 6.2: Topology of the two graphs, $\mathcal{G}_a$ on the left and $\mathcal{G}_b$ on the right, considered in the simulations. For simplicity, we draw only 13 nodes of $\mathcal{G}_b$, but the actual simulation runs with 100 nodes.

As suggested in section 6.2.1, all simulations are solved using the trapezoidal integration rule.

Small Network

The first simulation consists of a graph $\mathcal{G}_a$ with four nodes connected as in fig. 6.2(a), each with a sinusoidal input

$$u_j(t) = \sin(2\pi t/j) \tag{6.25}$$

Nodes are numbered starting from the topmost node in fig. 6.2(a) and increase clockwise until the last node is reached. The target consensus value $\max(u)$ is shown in fig. 6.3. All parameters, together with those of the other simulation, are listed in table 6.1. The control parameters are chosen according to (6.22) and (6.23) to achieve a desired maximum consensus error $p_{\text{error}} = 0.01$ and a desired minimum convergence rate $p_{\text{rate}} = 2\pi$.

Table 6.1: Parameters used to run the simulations on graphs $\mathcal{G}_a$ and $\mathcal{G}_b$ of fig. 6.2.

	τ	α	ϵ	t_s
Simulation of $\mathcal{G}_a$	$8 \cdot 10^{-2}$	$5 \cdot 10^{-1}$	$2.5 \cdot 10^{-3}$	10^{-4}
Simulation of $\mathcal{G}_b$	$5 \cdot 10^{-2}$	$5 \cdot 10^{-1}$	10^{-3}	10^{-4}

Large Network with Sparse Connectivity

This simulation consists of a graph $\mathcal{G}_b$ with one hundred nodes connected as in fig. 6.2(b), labeled clockwise, with ramp inputs

$$u_j(t) = \left(\frac{j}{n}\right) t + b_j \tag{6.26a}$$

$$b_{j-1} = \left(\frac{j-1}{jn}\right) t_f + b_j \tag{6.26b}$$

$$b_n = 1 - t_f \quad , \quad t_f = 10 \tag{6.26c}$$

Out of curiosity, note that $\lim_{n\to\infty} \max(u(t)) = [(t/t_f)^2 - 1]\, t_f/2 + 1$, for $t \in [0, t_f]$, with $u(t)$ from (6.26), is a parabola. The target consensus value $\max(u)$ is shown in fig. 6.3. All control and simulation parameters are listed in table 6.1. The control parameters are chosen according to (6.22) and (6.23) to achieve a desired maximum consensus error $p_{\text{error}} = 0.1$ and a desired minimum convergence rate $p_{\text{rate}} = 10$.

Simulation Results

The simulation results are shown in fig. 6.4 for both scenarios. Note that, as expected, during both simulations all nodes estimates converge to the maximum network input and proceed to track this value afterwards. The bottom graphs display the tracking errors, which remain smaller than the prescribed values of 0.01 and 0.1.

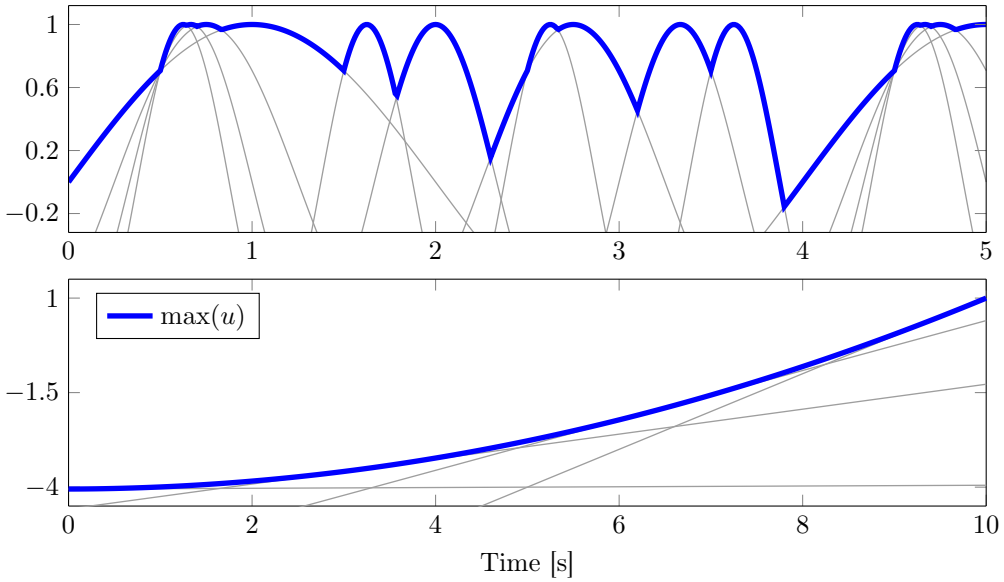


Figure 6.3: Maximum consensus of networks $\mathcal{G}_a$ and $\mathcal{G}_b$ shown in fig. 6.2, with inputs (6.25) and (6.26). In blue the value of $\max(u)$ and in light gray the nodes inputs u_j. Only the inputs u_1, u_{33}, u_{66}, and u_{100} are shown for $\mathcal{G}_b$.

Analyzing the results from the first simulation (topmost plots in fig. 6.4), we observe some of the convergence properties of the proposed consensus algorithm. A lot can be said about the interaction between nodes 2 (in solid black), 3 (in dashed black), and 4 (in solid red). Focusing on a short frame at the begining of the simulation, one can study several convergence properties discussed in the proof of Lemma 8, and also get a good feeling of the algorithm transient behavior.

When the simulation starts, node 4 knows no higher estimate than its own, and, hence, x_4 tracks its own input, since $\max(\mathcal{U}_4) = u_4$. Meanwhile, not only does node 3 knows an estimate higher than its own, $\max(\mathcal{X}_3) = x_2$, but also $\max(\mathcal{U}_3) = u_2 > x_3$. Thus, node's 3 estimate increases at a rate $(2 + \alpha)/\tau$. At approximately $t = 8.4$ ms x_3 surpasses x_4, and node 4 will then start increasing its estimate at a rate α/τ, since $\max(\mathcal{X}_4) = u_4 < x_4$, and follow x_3. Shortly after, at approximately $t = 11$ ms, x_3 surpasses $\max(\mathcal{U}_3) = x_2$, node's 3 highest known input, and x_3 will then start increasing at a rate α/τ, the same as x_4. Both estimates continue to grow until x_3 reaches the ϵ-vicinity below x_2 at approximately $t = 157$ ms. From this moment on, the network reaches a consensus and all estimates track $\max(u)$.

The results of the second simulation illustrate a phenomenon that, although very unlikely, can occur on some occasions. It is the error propagation from one node to another across a long path in the network, in this case, across the entire network. In the current example, it happens because $\max(u)$ is always changing from u_j to u_{j+1} as time goes by. These consecutive changes imply that x_{j-1} is always chasing x_j. Since all estimates are increasing at the same rate, without ever reaching one

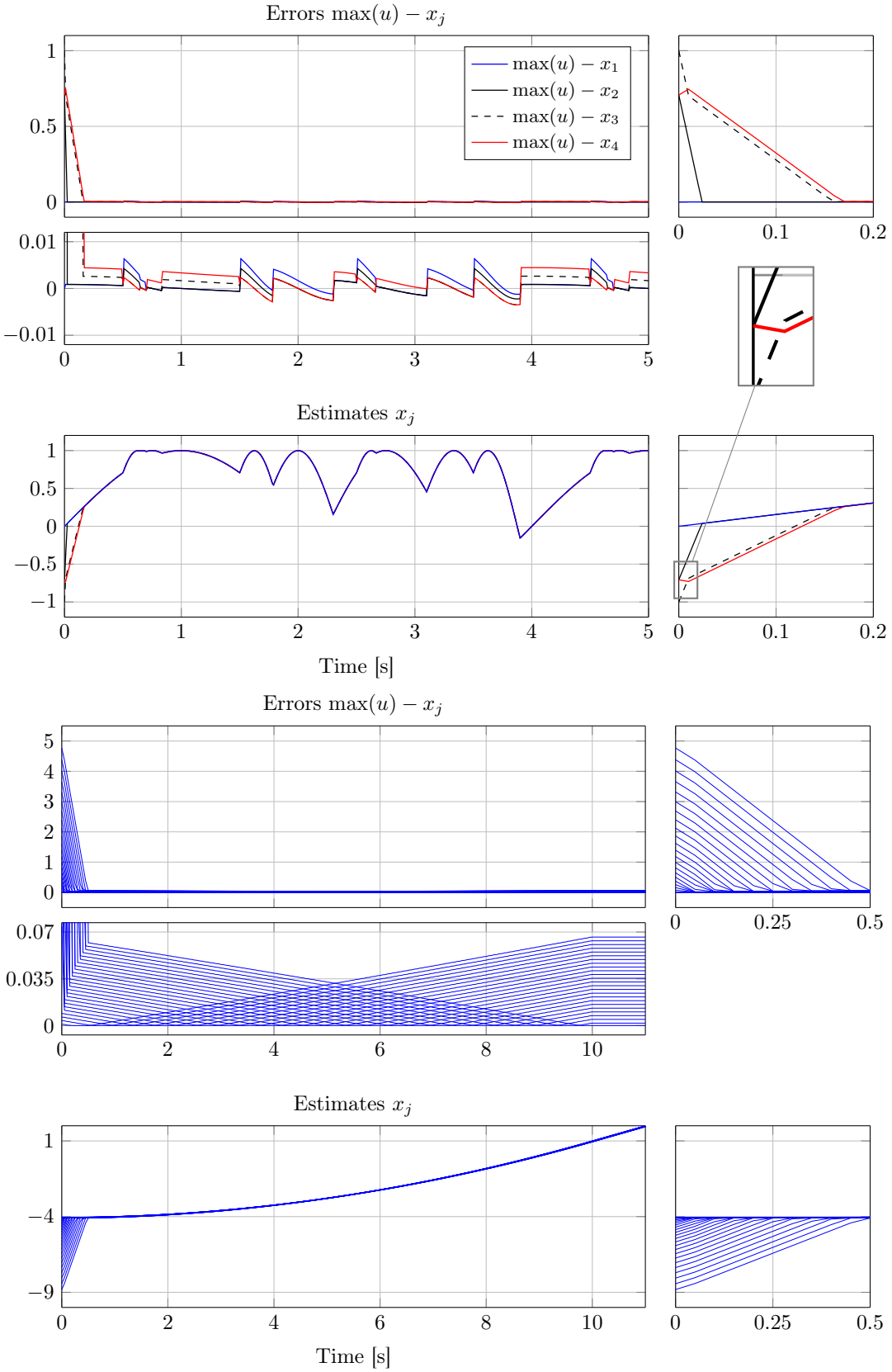


Figure 6.4: Simulation results for networks $\mathcal{G}_a$ and $\mathcal{G}_b$, with inputs (6.25) and (6.26), using the proposed consensus algorithm (6.5). At the top, the solution of $\mathcal{G}_a$. At the bottom, the solution of $\mathcal{G}_b$, where only 25 out of the 100 nodes are shown.

another, the final consensus error for each node becomes

$$|\max(u) - x_j|_{t \geq 10} \approx \delta(j, 100)\, r\, \epsilon = (100 - j)\, r\, \epsilon$$

where $\delta(j, 100)$ is the distance from node j to node 100 and $r \in (0, 1)$ a ratio which determines the separation between each estimate. On this simulation scenario, we have $r \approx 0.7$. On average, the overall final consensus errors is reduced if the node inputs of this second example are randomly reordered or, even better, if more edges are added to the network, such that the network diameter $d(\mathcal{G}_2)$ decreases.

6.3 Privacy-Preserving Dynamic Consensus Estimator

As it is, the proposed consensus estimator fails to provide privacy with respect to the nodes inputs and estimates. From the update rule (6.5), it is clear that any node in the network is aware of the inputs and estimates from all of its neighbors. In this section, we propose a modification to the proposed consensus estimator such that it is able to deliver privacy with respect to the nodes inputs. For that, we avoid transmitting the nodes inputs to their neighbors. Instead, we rely on the neighbor nodes telling to which direction should a node update its estimate.

Remark 14. *In this section, we develop a privacy preserving consensus estimator. We do not, however, develop the stability nor the convergence proof, but leave them as a topic for future research.*

Consider the following modification to the consensus dynamics (6.5),

$$\tau \dot{x}_j = (1 + \alpha) \operatorname*{sigmoid}_{\epsilon}(\hat{e}_j) + \operatorname{sign}(u_j - x_j) \tag{6.27a}$$

$$\hat{e}_j = \max(\mathcal{X}_j) - x_j \tag{6.27b}$$

$$\mathcal{X}_j = \{x \in \mathbb{R} : x = x_i + N_i,\ (j, i) \in \mathcal{E}\} \cup \{x_j\} \tag{6.27c}$$

where N_i is a zero-mean random variable that adds noise to the estimate advertised by node i, and $\mathcal{X}_j$ is modified to reflect that node j only knows the estimates of its neighbors plus some noise N_i.

The idea behind the above update rule is to protect the information about the node input and to mask the information about the node estimate. The node input is protected since it is never transmitted to any neighbor, and the estimates are masked by adding noise before transmitting them. By not transmitting the inputs, one downside is an increase in the convergence time. Furthermore, the noise added to the estimates cause a deviation that is proportional to the noise power.

This privacy preserving algorithm can also be modified to serve as a minimum consensus estimator. For that, one need only swap the max(.) function by the min(.) function in (6.27b).

We do not prove the convergence or stability of the proposed privacy-preserving dynamic consensus estimator. Nonetheless, it is not unexpected that the modified consensus algorithm retains the properties of the one from section 6.2, but for a small residual error. The residual error is caused by the noise added to the argument of the $\text{sigmoid}_\epsilon(.)$ function in eq. (6.27a), and, thus, can be made arbitrarily small by admitting more disclosure of the nodes estimates.

Chapter 7

Conclusion

In this Thesis, the main topic of study has been real-time optimization of single-objective and multi-objective problems, with single or multiple decision variables, and subject to the dynamics of an underlying process. Other contributions are made on global output tracking and dynamic consensus estimation. Although we start the thesis by describing the results on output-tracking, we follow a different path in this conclusion to highlight what we believe are the most relevant contributions.

Two distinct extremum-seeking sliding-mode controllers using only input-output information and periodic switching functions were introduced to solve multi-objective problems from two different perspectives. First, in section 4.2, we develop an algorithm capable of driving the trajectories of a process towards an equilibrium that corresponds to the Nash equilibrium of a multi-objective output map. Changing the way we compute the periodic switching function amplitude, in section 4.3, we show that the same algorithm is also able to optimize multi-objective problems subject to the dynamics of a nonlinear input-affine process. Second, in chapter 5, we develop an algorithm capable of driving the system trajectories toward an equilibrium that corresponds to the optimizer of a single-objective optimization problem with multiple decision variables. Assuming convex objective functions and taking the convex combination of the multiple objectives, a common technique for scalarization, we show that the solution is Pareto efficient. Our results extend the current theory of sliding-mode extremum-seeking control by significantly widening the class of solvable optimization problems to include those with nonlinear input-affine underlying processes, without relying on any time-scale separation nor producing undesired chattering.

A novel dynamic consensus estimator for maximum-value and minimum-value estimation (section 6.2) is introduced and generalized (section 6.3) to avoid the disclosure of a node's private estimates and measurements. Although the literature is quite vast on consensus estimation, and especially maximum-value estimation has been extensively studied by several authors, to the best of our knowledge, our algo-

rithm is the first to address dynamic maximum-value and minimum-value consensus explicitly.

Even though single-objective real-time optimization via extremum-seeking-control and trajectory tracking of systems with unknown control direction via output-feedback are topics that have received much attention in the past three decades, we have also managed to provide some contributions to them in chapters 2 and 3. For both cases, we have extended the current theory on sliding-mode control via switching functions to admit chattering-free control laws by resorting to continuous switching elements. Furthermore, we generalize the sliding-mode controller such that the proposed control law covers several different types of sliding surfaces. Both of these contributions, but especially the latter, extend the possible practical implementations in the sense of expanding the design choices available to the control designer.

Additionally, we show that the proposed sliding-mode controllers for extremum-seeking and output tracking can be applied to arbitrary nonlinear systems, given that the closed-loop satisfy the conditions to apply Tikhonov's theorem.

Finally, throughout the manuscript, we have illustrated most of the proposed algorithms through numerical simulations, inspired by theory and from practical applications. Notably, in section 4.2.6, an engineering application concerning the multi-variable real-time optimization for the output signal power spectrum of Raman optical amplifiers was considered to evaluate the performance of the multi-objective extremum-seeking control for Nash equilibrium seeking. The parameters of the optical fiber used in the simulations describe a Raman optical amplifier based on the *TrueWave®Reach - Low Water Peak*(RFLWP) optical fiber, with model parameters experimentally characterized by the company OFS Fitel Denmark ApS. The results are consistent with those of the literature, and performance improvements are obtained with the proposed scheme to pursue a flat signal spectrum and output signal power regulation.

7.1 Future Works

A common next step of every technique developed in this Thesis is experimental validation. Even though we thought out the simulations to display some practical implementation aspects, such as discrete-time implementation, control amplitude limitation, and chattering alleviation, experimental validation is indispensable to evaluate the proposed algorithms appropriately.

Regarding the multi-variable optimization algorithm proposed in chapter 5, there are three future developments that could improve the algorithm significantly:

- proof of convergence of the control design via unit-vector control, illustrated

in section 5.5, with control law (5.10) rewritten as

$$u_j(t) = \rho(t)\frac{s(t)}{\max\left(\|s(t)\|, \epsilon\right)} \tag{7.1}$$

$$s(t) = \sin\left[\frac{\pi\sigma(t)}{T_j}\right] \tag{7.2}$$

where we use $\max\left(\|s(t)\|, \epsilon\right)$ instead of $\|s(t)\|$ to make the control law continuous;

- generalization to input-affine nonlinear systems;
- distributed implementation relying on scalarization via consensus estimation using average consensus or the proposed maximum (minimum) consensus estimator.

It is worth noting that many authors suggest the approximation of $\|v\|$ using $\|v\| + \epsilon$. However, we have experienced that this approach delivers worse results when compared to $\max\left(\|v\|, \epsilon\right)$. Naturally, more experimentation is necessary before drawing any conclusions.

Extensively used in this Thesis and by many authors in the past three decades, the control design via periodically switching functions, sign(sin(.)), can still be significantly improved and generalized. Taking inspiration from the example of unit-vector control for the multi-variable case, we suggest the following generic periodically switching function:

$$u(t) = \rho(t)\,\phi\big(\sin\left[\pi T^{-1}\sigma(t)\right]\big) \tag{7.3}$$

where $\phi(.)$ can be any control law based on the output error that is capable of making the origin of the closed-loop process asymptotically stable. For instance, if one whishes to apply super-twisting control, then

$$\phi(s) = \phi_1(s) + \int_0^t \phi_2(s)d\tau \tag{7.4a}$$

$$\phi_1(s) = \begin{cases} -\kappa_1 \left|s_0\right|^{0.5}\operatorname{sign}(s) & , \quad |s| > s_0 \\ -\kappa_1 \left|s\right|^{0.5}\operatorname{sign}(s) & , \quad |s| \leq s_0 \end{cases} \tag{7.4b}$$

$$\phi_2(s) = \begin{cases} -\phi_M & \\ -\kappa_2 \operatorname{sign}(s) & , \quad |\phi| \leq \phi_M \end{cases} \tag{7.4c}$$

The original periodically switching function is recovered by selecting $\phi(.) = \operatorname{sign}(.)$. Extensive studies must still be conducted to characterize $\phi(.)$ properly, and to develop the stability and convergence results.

As discussed in section 1.3.1, the stabilization controller for SISO systems with unknown output sign can be related to (Scheinker and Krstić 2013; Scheinker and Krstić 2016), where the authors provide general results based on the minimization of control Lyapunov functions. Such a formulation allows one to consider stabilization as a consequence of solving an optimization problem via extremum-seeking control. Following this same direction with the controller proposed in section 3.3 might yield interesting new results. Also, one thing to keep in mind is that, when using control Lyapunov function, the objective function is known to the control designer, a fact that might be exploited to improve the performance of the sliding-mode controller.

Regarding the dynamic consensus estimator, we have already indicated a modification to ensure privacy preservation in section 6.3. Although we have tested in simulations this modified version of the algorithm, its stability and convergence properties still need to be studied. The other developments in the proposed consensus algorithm that we see as the most promising are:

- adapt the algorithm to enable a discrete-time implementation;
- improve the algorithm to ensure robustness to network delays;
- improve the algorithm to ensure robustness to malicious network attacks.

Bibliography

Adetola, V. and M. Guay (2006). "Adaptive output feedback extremum seeking receding horizon control of linear systems". In: *J. of Process Control* 16.5, pp. 521–533 (cit. on pp. 20, 27).

Adetola, V. and M. Guay (2007). "Parameter convergence in adaptive extremum-seeking control". In: *Automatica* 43.1, pp. 105–110 (cit. on p. 20).

Adetola, V. and M. Guay (2010). "Integration of real-time optimization and model predictive control". In: *J. of Process Control* 20.2, pp. 125–133 (cit. on pp. 20, 27).

Agrawal, G. P. (2001). *Nonlinear fiber optics*. Academic Press (cit. on p. 88).

Aizerman, M. and E. Pyatnitskii (1974). "Foundations of a theory of discontinuous systems". In: *Automation and Remote Control* 35.7, pp. 1066–1079 (cit. on pp. 14, 157, 166).

Aminde, N. O., T. R. Oliveira, and L. Hsu (2013). "Global output-feedback extremum seeking control via monitoring functions". In: *IEEE Conf. on Dec. and Contr.* IEEE, pp. 1031–1036 (cit. on p. 34).

Angeli, D. and E. D. Sontag (1999). "Forward completeness, unboundedness observability, and their Lyapunov characterizations". In: *Syst. and Control Lett.* 38.4-5, pp. 209–217 (cit. on pp. 69, 107).

Ariyur, K. B. and M. Krstic (2002). "Analysis and design of multivariable extremum seeking". In: *American Contr. Conf.* Vol. 4. IEEE, pp. 2903–2908 (cit. on pp. 25, 27, 28).

Barbot, J.-P., T. Boukhobza, and M. Djemai (1996). "Sliding mode observer for triangular input form". In: *IEEE Conf. on Dec. and Contr.* Vol. 2. IEEE, pp. 1489–1490 (cit. on p. 13).

Bartolini, G. and T. Zolezzi (1985). "Variable structure systems nonlinear in the control law". In: *IEEE Trans. Automat. Contr.* 30.7, pp. 681–684 (cit. on p. 14).

Basar, T. and G. J. Olsder (1999). "Dynamic noncooperative game theory, ser". In: *Classics in Applied Mathematics. SIAM* (cit. on p. 68).

Bauso, D., L. Giarré, and R. Pesenti (2009). "Consensus for networks with unknown but bounded disturbances". In: *SIAM J. on Contr. and Optim.* 48.3, pp. 1756–1770 (cit. on p. 29).

Bhaya, A. and E. Kaszkurewicz (2006). *Control perspectives on numerical algorithms and matrix problems*. Vol. 10. SIAM (cit. on p. 7).

Bhaya, A. and E. Kaszkurewicz (2007). "A control-theoretic approach to the design of zero finding numerical methods". In: *IEEE Trans. Automat. Contr.* 52.6, pp. 1014–1026 (cit. on p. 7).

Bromage, J. (2004). "Raman Amplification for Fiber Communications Systems". In: *J. of Lightwave Technology* 22.1, pp. 79–93 (cit. on pp. 86, 88, 167).

Bromage, J., K. Rottwitt, and M. E. Lines (2002). "A Method to Predict the Raman Gain Spectra of Germanosilicate Fibers With Arbitrary Index Profiles". In: *IEEE Photon. Technol. Lett.* 14.1, pp. 24–26 (cit. on pp. 88, 89).

Camacho, E. F. and C. Bordons Alba (2007). *Model Predictive Control*. 2nd ed. Springer (cit. on p. 56).

Chen, C.-T. (1999). *Linear system theory and design*. 2nd ed. New York: Oxford University Press, Inc. (cit. on p. 5).

Chicone, C. (2006). *Ordinary differential equations with applications*. Vol. 34. Springer-Verlag New York (cit. on p. 157).

Choi, J.-Y., M. Krstić, K. B. Ariyur, and J. S. Lee (2002). "Extremum seeking control for discrete-time systems". In: *IEEE Trans. Automat. Contr.* 47.2, pp. 318–323 (cit. on pp. 16, 27).

Clarke, F. H. (1990). *Optimization and nonsmooth analysis*. Vol. 5. Siam (cit. on p. 14).

Cunha, J. P. V., L. Hsu, R. R. Costa, and F. Lizarralde (2003). "Output-feedback model-reference sliding mode control of uncertain multivariable systems". In: *IEEE Trans. Automat. Contr.* 48.12, pp. 2245–2250 (cit. on p. 12).

DeCarlo, R. A., S. H. Zak, and G. P. Matthews (1988). "Variable structure control of nonlinear multivariable systems: a tutorial". In: *Proceedings of the IEEE* 76.3, pp. 212–232 (cit. on p. 10).

Dehaan, D. and M. Guay (2005). "Extremum-seeking control of state-constrained nonlinear systems". In: *Automatica* 41.9, pp. 1567–1574 (cit. on pp. 20, 27).

Dhaliwal, S. and M. Guay (2012). "Set-based adaptive estimation for a class of nonlinear systems with time-varying parameters". In: *IFAC Proceedings Volumes* 45.15, pp. 391–395 (cit. on p. 25).

Dower, P. M. and P. M. Farrell (2006). "On linear control of backward pumped raman amplifiers". In: *14th IFAC Symposium on System Identification*. Newcastle, Australia, pp. 547–552 (cit. on pp. 93, 94).

Dower, P. M., P. Farrell, and D. Nesić (2008). "Extremum seeking control of cascaded Raman optical amplifiers". In: *IEEE Trans. on Control Syst. Tech.* 16.3, pp. 396–407 (cit. on pp. 85–88, 94, 167–169).

Drakunov, S. V. (1992). "Sliding-mode observers based on equivalent control method". In: *IEEE Conf. on Dec. and Contr.* IEEE, pp. 2368–2369 (cit. on pp. 13, 22).

Drakunov, S., Ü. Özgüner, P. Dix, and B. Ashrafi (1995). "ABS control using optimum search via sliding modes". In: *IEEE Trans. on Control Syst. Tech.* 3.1, pp. 79–85 (cit. on pp. 21, 36, 55).

Drakunov, S. and V. Utkin (1995). "Sliding mode observers. Tutorial". In: *IEEE Conf. on Dec. and Contr.* Vol. 4. IEEE, pp. 3376–3378 (cit. on p. 13).

Drakunov, S. and Ü. Özgüner (1992). "Optimization of nonlinear system output via sliding mode approach". In: *Proceedings of the IEEE International Workshop on Variable Structure and Lyapunov Control of Uncertain Dynamical Systems, Sheffield, UK*, pp. 61–62 (cit. on pp. 7, 21, 27, 31, 34).

Dürr, H.-B., M. Krstić, A. Scheinker, and C. Ebenbauer (2017). "Extremum seeking for dynamic maps using Lie brackets and singular perturbations". In: *Automatica* 83, pp. 91–99 (cit. on pp. 18, 27).

Dürr, H.-B., M. S. Stanković, C. Ebenbauer, and K. H. Johansson (2013). "Lie bracket approximation of extremum seeking systems". In: *Automatica* 49.6, pp. 1538–1552 (cit. on p. 18).

Feng, Y., X. Yu, and Z. Man (2002). "Non-singular terminal sliding mode control of rigid manipulators". In: *Automatica* 38.12, pp. 2159–2167 (cit. on p. 12).

Ferreira, L. V., E. Kaszkurewicz, and A. Bhaya (2005). "Solving systems of linear equations via gradient systems with discontinuous righthand sides: application to LS-SVM". In: *IEEE Transactions on Neural Networks* 16.2, pp. 501–505 (cit. on p. 7).

Filippov, A. F. (1964). "Differential equations with discontinuous right-hand side". In: *American Math. Soc. Translations* 42.2, pp. 199–231 (cit. on pp. 14, 34).

Filippov, A. F. (1988). *Differential equations with discontinuous righthand sides: control systems*. Vol. 18. Springer Netherlands (cit. on pp. 14, 34, 157, 158).

Floquet, T. and J.-P. Barbot (2006). "A canonical form for the design of unknown input sliding mode observers". In: ed. by C. Edwards, E. F. Collet, and L. Fridman. Vol. 334. Lecture Notes in Control and Information Sciences. Spinger Verlag (cit. on p. 13).

Freeman, R. A., P. Yang, and K. M. Lynch (2006). "Stability and convergence properties of dynamic average consensus estimators". In: IEEE, pp. 338–343 (cit. on pp. 26, 127).

Fridman, L. and A. Levant (2002). "Higher order sliding modes". In: *Sliding Mode Control in Engineering*. Ed. by W. Perruquetti and J. P. Barbot. Marcel Dekker. Chap. 3, pp. 53–101 (cit. on p. 13).

Frihauf, P., M. Krstic, and T. Basar (2011). "Nash equilibrium seeking in noncooperative games". In: *IEEE Trans. Automat. Contr.* 57.5, pp. 1192–1207 (cit. on pp. 25, 27, 28, 32, 73).

Fu, L. and Ü. Özgüner (2009). "Variable structure extremum seeking control based on sliding mode gradient estimation for a class of nonlinear systems". In: *American Control Conference, 2009. ACC'09.* IEEE, pp. 8–13 (cit. on p. 22).

Fu, L. and Ü. Özgüner (2011). "Extremum seeking with sliding mode gradient estimation and asymptotic regulation for a class of nonlinear systems". In: *Automatica* 47.12, pp. 2595–2603 (cit. on pp. 22, 27).

Ghaffari, A., M. Krstić, and D. Nešić (2012). "Multivariable Newton-based extremum seeking". In: *Automatica* 48.8, pp. 1759–1767 (cit. on pp. 18, 25, 27, 28).

Guay, M. (2016). "A perturbation-based proportional integral extremum-seeking control approach". In: *IEEE Trans. Automat. Contr.* 61.11, pp. 3370–3381 (cit. on pp. 19, 27).

Guay, M. and D. Dochain (2014a). "A proportional-integral extremum-seeking controller design technique". In: *IFAC World Congress*, pp. 337–342 (cit. on p. 19).

Guay, M. and D. Dochain (2014b). "A utopic multi-objective extremum-seeking controller design technique". In: *IEEE Conf. on Dec. and Contr.* IEEE, pp. 3438–3443 (cit. on pp. 25, 27, 28).

Guay, M. and D. Dochain (2015). "A time-varying extremum-seeking control approach". In: *Automatica* 51, pp. 356–363 (cit. on pp. 19, 25–27).

Guay, M. and D. Dochain (2017). "A proportional-integral extremum-seeking controller design technique". In: *Automatica* 77, pp. 61–67 (cit. on pp. 19, 26, 27).

Guay, M., I. Vandermeulen, S. Dougherty, and P. J. McLellan (2018). "Distributed extremum-seeking control over networks of dynamically coupled unstable dynamic agents". In: *Automatica* 93, pp. 498–509 (cit. on pp. 25–28, 32, 127).

Guay, M. and T. Zhang (2003). "Adaptive extremum seeking control of nonlinear dynamic systems with parametric uncertainties". In: *Automatica* 39.7, pp. 1283–1293 (cit. on pp. 20, 27).

Gutman, S. and G. Leitmann (1975). "Stabilizing control for linear systems with bounded parameter and input uncertainty". In: *IFIP Technical Conference on Optimization Techniques*. Springer, pp. 729–755 (cit. on p. 12).

Gutman, S. (1979). "Uncertain dynamical systems–A Lyapunov min-max approach". In: *IEEE Trans. Automat. Contr.* 24.3, pp. 437–443 (cit. on p. 12).

Haskara, I. and Ü. Özgüner (1999). "Equivalent value filters in disturbance estimation and state observation". In: *Lecture Notes in Control and Information Science* 247 (cit. on p. 13).

Haskara, I., Ü. Özgüner, and J. Winkelman (2000). "Extremum control for optimal operating point determination and set point optimization via sliding modes". In: *J. of Dynamic Syst., Measurement, and Control* 122.4, pp. 719–724 (cit. on pp. 21, 22, 27, 31, 34).

He, J., J. Chen, P. Cheng, and X. Cao (2014a). "Secure time synchronization in wirelesssensor networks: A maximumconsensus-based approach". In: *IEEE Trans. on Parallel and Distributed Syst.* 25.4, pp. 1055–1065 (cit. on p. 29).

He, J., P. Cheng, L. Shi, J. Chen, and Y. Sun (2014b). "Time synchronization in WSNs: A maximum-value-based consensus approach". In: *IEEE Trans. Automat. Contr.* 59.3, pp. 660–675 (cit. on p. 29).

Headley, C. and G. P. Agrawal (2005). *Raman Amplification in Fiber Optical Communication Systems*. Academic Press (cit. on p. 84).

Hsu, L. (1997). "Smooth sliding control of uncertain systems based on a prediction error". In: *Int. J. of Robust and Nonlinear Control* 7.4, pp. 353–372 (cit. on p. 12).

Hung, J. Y., W. Gao, and J. C. Hung (1993). "Variable structure control: A survey". In: *IEEE Trans. on Industrial Elec.* 40.1, pp. 2–22 (cit. on p. 10).

Kalyanmoy, D. (2001). *Multi objective optimization using evolutionary algorithms*. John Wiley and Sons (cit. on p. 25).

Khalil, H. K. (2002). *Nonlinear systems*. 3rd ed. Prentice Hall (cit. on pp. 13, 22, 32, 50, 51, 104, 161, 165).

Khalil, H. K. and L. Praly (2014). "High-gain observers in nonlinear feedback control". In: *International Journal of Robust and Nonlinear Control* 24.6, pp. 993–1015 (cit. on p. 6).

Kia, S. S., J. Cortés, and S. Martinez (2015). "Dynamic average consensus under limited control authority and privacy requirements". In: 25.13, pp. 1941–1966 (cit. on p. 29).

Kia, S. S., B. Van Scoy, J. Cortes, R. A. Freeman, K. M. Lynch, and S. Martinez (2019). "Tutorial on Dynamic Average Consensus: The Problem, Its Applications, and the Algorithms". In: *IEEE Control Syst. Mag.* 39.3, pp. 40–72 (cit. on p. 29).

Kidorf, H., K. Rottwitt, M. Nissov, M. Ma, and E. Rabarijaona (1999). "Pump interactions in a 100-nm bandwidth raman amplifier". In: *IEEE Photon. Technol. Lett.* 11.5, pp. 530–532 (cit. on pp. 84, 86).

Kokotovic, P. V. (1992). "The joy of feedback: nonlinear and adaptive". In: *IEEE Control Syst.* 12.3, pp. 7–17 (cit. on p. 5).

Krichman, M., E. D. Sontag, and Y. Wang (2001). "Input-output-to-state stability". In: 39.6, pp. 1874–1928 (cit. on p. 38).

Krstić, M. (2009). *Delay Compensation for Nonlinear, Adaptive, and PDE Systems*. 1st. Birkhauser (cit. on p. 168).

Krstić, M. and A. Smyshlyaev (2008). *Boundary Control of PDEs: A Course on Backstepping Designs*. 1st. SIAM (cit. on p. 168).

Krstić, M. (2013). "Extremum Seeking Control". In: *Encyclopedia of Systems and Control*. Ed. by J. Baillieul and T. Samad. London: Springer London, pp. 1–5. ISBN: 978-1-4471-5102-9 (cit. on p. 20).

Krstić, M. and H.-H. Wang (2000). "Stability of extremum seeking feedback for general nonlinear dynamic systems". In: *Automatica* 36.4, pp. 595–601 (cit. on pp. 15, 16, 18, 25, 27).

Lara-Cisneros, G., R. Femat, and D. Dochain (2017). "Robust sliding mode-based extremum-seeking controller for reaction systems via uncertainty estimation approach". In: *Int. J. of Robust and Nonlinear Control* 27.16, pp. 3218–3235 (cit. on pp. 22, 23, 27, 31).

Levant, A. (1993). "Sliding order and sliding accuracy in sliding mode control". In: *Int. J. of Contr.* 58.6, pp. 1247–1263 (cit. on pp. 10, 11, 13).

Levant, A. (1998). "Robust exact differentiation via sliding mode technique". In: *Automatica* 34.3, pp. 379–384 (cit. on p. 14).

Levant, A. (2003). "Higher-order sliding modes, differentiation and output-feedback control". In: *Int. J. of Contr.* 76.9-10, pp. 924–941 (cit. on p. 13).

Liberzon, D. (2011). *Calculus of variations and optimal control theory: a concise introduction*. Princeton University Press (cit. on pp. 8, 55).

Lizarralde, F., J. C. Monteiro, R. Pereira, and L. Hsu (2017). "Sliding Mode Based Extremum Seeking Control of Two-Phase Flow Micro-Thermal-Fluid Cooling Systems". In: *IFAC World Congress* 50.1. 20th IFAC World Congress, pp. 5133–5138. ISSN: 2405-8963 (cit. on pp. 22, 27, 30, 34, 35).

Lyapunov, A. M. (1992). "The general problem of the stability of motion". In: *Int. J. of Contr.* 55.3, pp. 531–534 (cit. on pp. 160, 161).

Manitara, N. E. and C. N. Hadjicostis (2013). "Privacy-preserving asymptotic average consensus". In: *European Contr. Conf.* IEEE, pp. 760–765 (cit. on p. 29).

Marler, R. T. and J. S. Arora (2004). "Survey of multi-objective optimization methods for engineering". In: *Structural and multidisciplinary optimization* 26.6, pp. 369–395 (cit. on p. 24).

Marler, R. T. and J. S. Arora (2010). "The weighted sum method for multi-objective optimization: new insights". In: *Structural and Multidisciplinary Optimization* 41.6, pp. 853–862 (cit. on p. 24).

Mazenc, F., L. Praly, and W. Dayawansa (1994). "Global stabilization by output feedback: examples and counterexamples". In: *Syst. and Control Lett.* 23.2, pp. 119–125 (cit. on p. 6).

Miettinen, K. (1998). *Nonlinear multiobjective optimization*. Vol. 12. Springer. ISBN: 978-0-7923-8278-2 (cit. on p. 25).

Mills, G. and M. Krstić (2015). "Maximizing higher derivatives of unknown maps with extremum seeking". In: *IEEE Conf. on Dec. and Contr.* IEEE, pp. 5648–5653 (cit. on p. 19).

Mills, G. and M. Krstić (2018). "Maximizing Map Sensitivity and Higher Derivatives Via Extremum Seeking". In: *IEEE Trans. Automat. Contr.* (Cit. on pp. 19, 27).

Mo, Y. and R. M. Murray (2016). "Privacy preserving average consensus". In: *IEEE Trans. Automat. Contr.* 62.2, pp. 753–765 (cit. on p. 29).

Moase, W. H. and C. Manzie (2012). "Fast extremum-seeking for Wiener–Hammerstein plants". In: *Automatica* 48.10, pp. 2433–2443 (cit. on pp. 18, 19, 27).

Moase, W. H., C. Manzie, and M. J. Brear (2010). "Newton-like extremum-seeking for the control of thermoacoustic instability". In: *IEEE Trans. Automat. Contr.* 55.9, pp. 2094–2105 (cit. on pp. 18, 25, 27).

Monteiro, J. C. and A. J. Peixoto (2020). "Convergence and Stability Properties of a Dynamic Maximum Consensus Estimator". In: *IFAC World Congress*. International Federation of Automatic Control (cit. on p. 31).

Moradian, H. and S. S. Kia (2018). "On robustness analysis of a dynamic average consensus algorithm to communication delay". In: 6.2, pp. 633–641 (cit. on p. 29).

Nash, J. (1950). "Equilibrium points in n-person games". In: *Proceedings of the national academy of sciences* 36.1, pp. 48–49 (cit. on pp. 24, 66, 67).

Nash, J. (1951). "Non-cooperative games". In: *Annals of mathematics*, pp. 286–295 (cit. on p. 66).

Nešić, D., Y. Tan, W. H. Moase, and C. Manzie (2010a). "A unifying approach to extremum seeking: Adaptive schemes based on estimation of derivatives". In: *IEEE Conf. on Dec. and Contr.* IEEE, pp. 4625–4630 (cit. on p. 15).

Nešić, D. (2009). "Extremum seeking control: Convergence analysis". In: *European Contr. Conf.* IEEE, pp. 1702–1715 (cit. on p. 17).

Nešić, D., A. Mohammadi, and C. Manzie (2010b). "A systematic approach to extremum seeking based on parameter estimation". In: *IEEE Conf. on Dec. and Contr.* IEEE, pp. 3902–3907 (cit. on pp. 15, 18).

Nešić, D., Y. Tan, C. Manzie, A. Mohammadi, and W. Moase (2012). "A unifying framework for analysis and design of extremum seeking controllers". In: *IEEE Chinese Conf. on Dec. and Contr.* IEEE, pp. 4274–4285 (cit. on p. 15).

Nunes, E. V. L., L. Hsu, and F. Lizarralde (2009). "Global exact tracking for uncertain systems using output-feedback sliding mode control". In: *IEEE Trans. Automat. Contr.* 54.5, pp. 1141–1147 (cit. on p. 14).

Olfati-Saber, R., J. A. Fax, and R. M. Murray (2007). "Consensus and cooperation in networked multi-agent systems". In: 95.1, pp. 215–233 (cit. on p. 28).

Olfati-Saber, R. and R. M. Murray (2004). "Consensus problems in networks of agents with switching topology and time-delays". In: *IEEE Trans. Automat. Contr.* 49.9, pp. 1520–1533 (cit. on p. 29).

Oliveira, T. R., A. Estrada, and L. M. Fridman (2017a). "Global and exact HOSM differentiator with dynamic gains for output-feedback sliding mode control". In: *Automatica* 81, pp. 156–163 (cit. on p. 14).

Oliveira, T. R., L. Hsu, and A. J. Peixoto (2011). "Output-feedback global tracking for unknown control direction plants with application to extremum-seeking control". In: *Automatica* 47, pp. 2029–2038 (cit. on pp. 22, 35–37, 44).

Oliveira, T. R., M. Krstić, and D. Tsubakino (2015). "Multiparameter extremum seeking with output delays". In: IEEE, pp. 152–158 (cit. on pp. 19, 27).

Oliveira, T. R., M. Krstić, and D. Tsubakino (2017b). "Extremum seeking for static maps with delays". In: *IEEE Trans. Automat. Contr.* 62.4, pp. 1911–1926 (cit. on pp. 19, 27).

Oliveira, T. R., A. J. Peixoto, and L. Hsu (2010a). "Sliding mode control of uncertain multivariable nonlinear systems with unknown control direction via switching and monitoring function". In: *IEEE Trans. Automat. Contr.* 55.4, pp. 1028–1034 (cit. on pp. 12, 34).

Oliveira, T. R., A. J. Peixoto, and L. Hsu (2010b). "Solving the unknown control direction problem for strongly nonlinear systems by means of periodic switching function and norm observers". In: *International Workshop on Variable Structure Systems (VSS)*. IEEE, pp. 118–123 (cit. on pp. 22, 31).

Oliveira, T. R., A. J. Peixoto, and L. Hsu (2012). "Global real-time optimization by output-feedback extremum-seeking control with sliding modes". In: *J. of the Franklin Institute* 349.4, pp. 1397–1415 (cit. on pp. 22, 27, 31, 34, 44, 55).

Palais, J. (1998). *Fiber Optic Communications*. 4th. Prentice Hall (cit. on pp. 84, 85).

Pan, Y., T. Acarman, and Ü. Özgüner (2002). "Nash solution by extremum seeking control approach". In: *IEEE Conf. on Dec. and Contr.* Vol. 1. IEEE, pp. 329–334 (cit. on pp. 26, 28, 32).

Pan, Y., Ü. Özgüner, and T. Acarman (2003). "Stability and performance improvement of extremum seeking control with sliding mode". In: *Int. J. of Contr.* 76.9-10, pp. 968–985 (cit. on p. 22).

Peixoto, A. J. and T. R. Oliveira (2012). "Extremum seeking control via sliding mode and periodic switching function applied to Raman optical amplifiers". In: *American Contr. Conf.* IEEE, pp. 5377–5382 (cit. on pp. 26, 28, 32, 71, 73, 77, 85–88).

Peixoto, A. J. and T. R. Oliveira (2014). "Global output-feedback extremum seeking control for nonlinear systems with arbitrary relative degree". In: *IFAC World Congress* 47.3, pp. 5473–5478 (cit. on pp. 22, 23, 31).

Peixoto, A. J. and T. R. Oliveira (2016). "Global output-feedback extremum seeking for a class of nonlinear dynamic systems with arbitrary relative degree". In: *International Journal of Control* 89.9, pp. 1821–1837 (cit. on pp. 22, 27).

Peixoto, A. J., T. R. Oliveira, D. Pereira-Dias, and J. C. Monteiro (2020). "Multivariable extremum-seeking by periodic switching functions with application to Raman optical amplifiers". In: *Contr. Eng. Practice* 96, p. 104278 (cit. on pp. 26, 27, 30, 32, 68, 71, 73, 83, 94, 95).

Polyakov, A. and L. Fridman (2014). "Stability notions and Lyapunov functions for sliding mode control systems". In: *J. of the Franklin Institute* 351.4, pp. 1831–1865 (cit. on pp. 10, 14, 34, 162, 164–166).

Poveda, J. and N. Quijano (2013). "Distributed extremum seeking for real-time resource allocation". In: *American Contr. Conf.* IEEE, pp. 2772–2777 (cit. on pp. 25–28, 127).

Ren, W. and R. W. Beard (2005). "Consensus seeking in multiagent systems under dynamically changing interaction topologies". In: *IEEE Trans. Automat. Contr.* 50.5, pp. 655–661 (cit. on p. 28).

Rotea, M. A. (2000). "Analysis of multivariable extremum seeking algorithms". In: *American Contr. Conf.* Vol. 1. 6. IEEE, pp. 433–437 (cit. on pp. 16, 27).

Rušiti, D., G. Evangelisti, T. R. Oliveira, M. Gerdts, and M. Krstić (Jan. 2019). "Stochastic Extremum Seeking for Dynamic Maps With Delays". In: *IEEE Contr. Syst. Lett.* 3.1, pp. 61–66. ISSN: 2475-1456 (cit. on p. 27).

Rušiti, D., T. R. Oliveira, G. Mills, and M. Krstić (2016). "Newton-based extremum seeking for higher derivatives of unknown Maps with Delays". In: *IEEE Conf. on Dec. and Contr.* IEEE, pp. 1249–1254 (cit. on pp. 19, 27).

Rušiti, D., T. R. Oliveira, G. Mills, and M. Krstić (2018). "Deterministic and Stochastic Newton-based extremum seeking for higher derivatives of unknown maps with delays". In: *European J. of Contr.* 41, pp. 72–83 (cit. on p. 19).

Salamah, Y. B., L. Fiorentini, and Ü. Özgüner (2018). "Cooperative Extremum Seeking Control via Sliding Mode for Distributed Optimization". In: *IEEE Conf. on Dec. and Contr.* IEEE, pp. 1281–1286 (cit. on pp. 26–28, 32, 127).

Salamah, Y. B. and Ü. Özgüncr (2018). "Sliding mode multivariable extremum seeking control with application to wind farm power optimization". In: *American Contr. Conf.* IEEE, pp. 5321–5326 (cit. on pp. 26–28, 32).

Scheinker, A. and M. Krstić (2013). "Minimum-seeking for CLFs: Universal semiglobally stabilizing feedback under unknown control directions". In: *IEEE Trans. Automat. Contr.* 58.5, pp. 1107–1122 (cit. on pp. 32, 144).

Scheinker, A. and M. Krstić (2016). *Model-free Stabilization by Extremum Seeking*. Springer (cit. on pp. 32, 144).

Sharafi, J., W. H. Moase, and C. Manzie (2015). "Fast extremum seeking on Hammerstein plants: A model-based approach". In: *Automatica* 59, pp. 171–181 (cit. on p. 20).

Sharafi, J., W. H. Moase, R. C. Shekhar, and C. Manzie (2013). "Fast model-based extremum seeking on Hammerstein plants". In: *IEEE Conf. on Dec. and Contr.* IEEE, pp. 6226–6231 (cit. on pp. 20, 27).

Shevitz, D. and B. Paden (1994). "Lyapunov stability theory of nonsmooth systems". In: *IEEE Trans. Automat. Contr.* 39.9, pp. 1910–1914 (cit. on p. 14).

Shi, G. and K. H. Johansson (2013). "Robust consensus for continuous-time multiagent dynamics". In: *SIAM Journal on Control and Optimization* 51.5, pp. 3673–3691 (cit. on p. 29).

Shim, H. and A. Teel (2001). "Further results on the nonlinear separation principle: the general 'asymptotically controllable' case". In: *IFAC Proceedings Volumes* 34.6, pp. 1457–1462 (cit. on p. 6).

Spanos, D. P., R. Olfati-Saber, and R. M. Murray (2005). "Dynamic consensus on mobile networks". In: *IFAC World Congress*. Citeseer, pp. 1–6 (cit. on p. 28).

Sussman, H. J. and P. V. Kokotovic (1991). "The peaking phenomenon and the global stabilization of nonlinear systems". In: *IEEE Trans. Automat. Contr.* 36.4 (cit. on p. 13).

Takahashi, R. H. and P. L. Peres (1999). "Unknown input observers for uncertain systems: A unifying approach". In: *European J. of Contr.* 5.2-4, pp. 261–275 (cit. on p. 6).

Tan, Y., D. Nešić, and I. Mareels (2006). "On non-local stability properties of extremum seeking control". In: *Automatica* 42.6, pp. 889–903 (cit. on pp. 17, 27).

Tan, Y., D. Nešić, and I. Mareels (2008). "On the choice of dither in extremum seeking systems: A case study". In: *Automatica* 44.5, pp. 1446–1450 (cit. on p. 17).

Tan, Y., D. Nešić, I. M. Mareels, and A. Astolfi (2009). "On global extremum seeking in the presence of local extrema". In: *Automatica* 45.1, pp. 245–251 (cit. on pp. 17, 27).

Teel, A. and L. Praly (1994). "Global stabilizability and observability imply semi-global stabilizability by output feedback". In: *Syst. and Control Lett.* 22.5, pp. 313–325 (cit. on p. 6).

Teel, A. and L. Praly (1995). "Tools for semiglobal stabilization by partial state and output feedback". In: *SIAM J. on Contr. and Optim.* 33.5, pp. 1443–1488 (cit. on p. 6).

Utkin, V. (1977). "Variable structure systems with sliding modes". In: *IEEE Trans. Automat. Contr.* 22.2, pp. 212–222 (cit. on pp. 3, 10, 13, 14).

Utkin, V. I. (1992). *Sliding modes in control and optimization*. Springer-Verlag Berlin Heidelberg (cit. on pp. 7, 13, 14, 21, 34, 157, 166).

Venkataraman, S. and S. Gulati (1993). "Control of nonlinear systems using terminal sliding modes". In: *J. of Dynamic Syst., Measurement, and Control* 115.3, pp. 554–560 (cit. on p. 11).

Walsh, G. C. (2000). "On the application of multi-parameter extremum seeking control". In: *American Contr. Conf.* Vol. 1. 6. IEEE, pp. 411–415 (cit. on pp. 16, 27).

Wang, X., J. He, P. Cheng, and J. Chen (2019). "Differentially Private Maximum Consensus: Design, Analysis and Impossibility Result". In: *IEEE Trans. on Network Sci. and Eng.* 6.4, pp. 928–939 (cit. on p. 29).

Wendell, R. E. and D. N. Lee (1977). "Efficiency in multiple objective optimization problems". In: *Mathematical Programming* 12.1, pp. 406–414 (cit. on p. 105).

Xiong, Y. and M. Saif (2001). "Sliding mode observer for nonlinear uncertain systems". In: *IEEE Trans. Automat. Contr.* 46.12, pp. 2012–2017 (cit. on p. 13).

Ye, M., G. Hu, and S. Xu (2020). "An extremum seeking-based approach for Nash equilibrium seeking in N-cluster noncooperative games". In: *Automatica* 114, p. 108815 (cit. on p. 28).

Young, K. D., V. I. Utkin, and Ü. Özgüner (1999). "A control engineer's guide to sliding mode control". In: *IEEE Trans. on Control Syst. Tech.* 7.3, pp. 328–342 (cit. on pp. 10, 13).

Yu, H. and Ü. Özgüner (2002). "Extremum-seeking control strategy for ABS system with time delay". In: *American Contr. Conf.* Vol. 5. IEEE, pp. 3753–3758 (cit. on pp. 21, 27).

Yu, H. and Ü. Özgüner (2003). "Smooth extremum-seeking control via second order sliding mode". In: *American Contr. Conf.* Vol. 4. IEEE, pp. 3248–3253 (cit. on p. 27).

Yu, X. and M. Zhihong (2002). "Fast terminal sliding-mode control design for nonlinear dynamical systems". In: *IEEE Trans. on Circuits and Sys. I: Fundamental Theory and Applications* 49.2, pp. 261–264 (cit. on p. 12).

Zak, M. (1988). "Terminal attractors for addressable memory in neural networks". In: *Physics Letters A* 133.1-2, pp. 18–22 (cit. on p. 11).

Appendix A

Mathematical Background

A.1 Notation

- $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ is the set of extended real numbers, $\mathbb{R}_+ = \{x \in \mathbb{R} | x > 0\}$, and $\bar{\mathbb{R}}_+ = \mathbb{R}_+ \cup \{+\infty\}$.
- The Euclidean norm of a vector $x \in \mathbb{R}^n$ is denoted $\|x\|$.
- The i-th element of a vector $x \in \mathbb{R}^n$ is denoted either x_i or $(x)_i$.
- The absolute value $|x|$ of a vector $x \in \mathbb{R}^n$ is taken element-wise, such that $|x| = [|x_1| \ \ |x_2| \dots |x_n|]^\mathsf{T}$
- The sign function is defined as

$$\underset{\alpha}{\text{sign}}(x) = \begin{cases} +1 & , \quad x > 0 \\ -1 & , \quad x < 0 \\ 0 & , \quad x = 0 \end{cases} \tag{A.1}$$

- The geometric sum of two sets $M_1, M_2 \subseteq \mathbb{R}^n$ is denoted by "+",

$$M_1 + M_2 = \bigcup \{x_1 + x_2\} \ , \ x_1 \in M_1 \ , \ x_2 \in M_2 \tag{A.2}$$

- The set of continuous functions defined on $\Omega \subseteq \mathbb{R}^n$, continuously differentiable up to order k is denoted $C^k(\Omega)$. If $\Omega = \mathbb{R}^n$, the domain is omitted and $C^k(\Omega) = C^k$.
- If $V : \mathbb{R}^n \mapsto \mathbb{R}$ belongs to C^1, its gradient is defined as the column vector $\nabla V(x) = \left[\frac{\partial V}{\partial x_1} \ \frac{\partial V}{\partial x_2} \ \dots \ \frac{\partial V}{\partial x_n}\right]^\mathsf{T}$. If $\sigma : \mathbb{R}^n \mapsto \mathbb{R}^p$, its gradient is a matrix with columns given by the gradients of its components $\nabla\sigma(x) = [\nabla\sigma_1(x) \ \nabla\sigma_2(x) \ \dots \ \nabla\sigma_p(x)]^\mathsf{T}$.

- The power set of M, set of all possible combinations of subsets of M, is denoted 2^M.

A.2 Discontinuous Systems

The mathematical foundations for SMC lie in Filippov's theory of ordinary differential equations with discontinuous right-hand side (Filippov 1988), henceforth called discontinuous ODEs. It is known that Filippov's theory fails to describe some systems nonlinear in the control input. Even though we deal with systems affine in the control input, for the sake of completeness, two other interpretations of solutions to discontinuous ODEs are considered. The first one is due to Utkin's equivalent control method (Utkin 1992), while the other is an extension of Filippov's theory with elements taken from Utkin's method (Aizerman and Pyatnitskii 1974).

For continuous ODEs (those with continuous right-hand sides), the classical theory (Chicone 2006) introduces a solution to

$$\dot{x} = f(t,x) \quad , \quad f : \mathbb{R} \times \mathbb{R}^n \mapsto \mathbb{R}^n \tag{A.3}$$

as a differentiable function $x : \mathbb{R} \mapsto \mathbb{R}^n$, defined on some interval $\mathcal{I} \subseteq \mathbb{R}$. Furthermore, if $f \in C^k$, for some $k = 1, 2, \ldots, \infty$, then also $x \in C^k$.

Existence and uniqueness of solutions of (A.3) are properties usually related to smoothness or Lipschitz continuity, with Lipschitz continuity observed with respect to the second argument of f. Unfortunately, neither smoothness nor Lipschitz continuity are properties that apply to discontinuous ODEs.

Example 1. *Consider the simple relay controlled integrator*

$$\dot{x} = -\operatorname{sign}(x) \tag{A.4}$$

This trivial example of great practical relevance cannot be studied in light of classical ODE theory, since the sign function is discontinuous at $x = 0$.

Instead, we consider f in (A.3) piece-wise continuous, such that its domain of definition can be split into N disjoint open connected subsets $G_j \subset \mathbb{R} \times \mathbb{R}^n$, and the boundary set $S = \bigcup_{j=1}^{N} \partial G_j$ is of measure zero (in Lebesgue's sense). The function f is continuous at every G_j and for each $(t^*, x^*) \in \partial G_j$, any sequence $(t^k, x^k) \in G_j : (t^k, x^k) \to (t^*, x^*)$ is such that $f(t^k, x^k) \to f^j(t^*, x^*)$. Functions $f_j : \mathbb{R} \times \mathbb{R}^n \mapsto \mathbb{R}^n$ are defined according to this limiting process,

$$f^j(t,x) = \lim_{(t^k,x^k)\to(t^*,x^*)} f(t^k, x^k) \quad , \; (t^k, x^k) \in G_j \quad , \quad (t^*, x^*) \in \partial G_j \tag{A.5}$$

A.2.1 Filippov's Definition of Solutions

Filippov's theory of ODEs with discontinuous right-hand side provides an *axiomatic* definition of solutions to (A.3) with f piecewise continuous[1]. For that, equation (A.3) is substituted by the following differential inclusion

$$\dot{x} \in K_f(t,x), \tag{A.6a}$$

$$K_f(t,x) = \begin{cases} \{f(t,x)\} & , \quad (t,x) \in \mathbb{R}^{n+1} \setminus S \\ \operatorname{co}\left(\bigcup\limits_{j \in \mathcal{N}(t,x)} \{f^j(t,x)\} \right) & , \quad (t,x) \in S \end{cases} \tag{A.6b}$$

where $\operatorname{co}(M)$ is the convex closure of a set M, and set-valued index function $\mathcal{N} : \mathbb{R}^{n+1} \mapsto 2^{\{1,2,\ldots,N\}}$ defined on S indicates boundaries ∂G_j which intersect at (t,x),

$$\mathcal{N}(t,x) = \{j \in \{1,2,\ldots,N\} \mid (t,x) \in \partial G_j\} \tag{A.7}$$

Note that, by definition, the set $K_f(t,x)$ is a convex polyhedron for $(t,x) \in S$.

Definition 18 (Filippov 1988, page 50)**.** *An absolutely continuous function $x : \mathcal{I} \mapsto \mathbb{R}^n$ defined on some interval or segment $\mathcal{I}$ is a solution of (A.3) if it satisfies the differential inclusion (A.6) almost everywhere on $\mathcal{I}$*[2].

The intuition behind Filippov's definition is actually simple. The set-valued function K_f determines allowable velocities that a solution of (A.3) might take. For a point $(t,x) \notin S$ that does not belong to the discontinuity set (switching manifold in the case of SMC), $K_f(t,x)$ is a singleton containing only $f(t,x)$, such that $\dot{x} = f(t,x)$. For $(t,x) \in S$ in the discontinuity set, the velocity should belong to the convex combination of every possible velocity vector in a neighborhood $B_\delta \setminus S$ ($\delta \to 0$) of (t,x), that is, an infinitesimal neighborhood of (t,x) that discards the set S of zero measure.

For SMC, it is common to consider $f(t,x)$ discontinuous only on a smooth manifold $S = \{x \in \mathbb{R}^n \mid \sigma(x) = 0\}$ computed from the switching function $\sigma : \mathbb{R}^n \mapsto \mathbb{R}$, which separates the state-space into two domains $G^+ = \{x \in \mathbb{R}^n \mid \sigma(x) > 0\}$ and $G^- = \{x \in \mathbb{R}^n \mid \sigma(x) < 0\}$, as shown in fig. A.1.

Let functions f^+ and f^- be defined on G^+ and G^- by the limiting process (A.5), for a constant t. In this case, the set K_f obtained at $(t,x) \in S$ is a line segment

[1] Actually, Filippov's theory covers the case where $f(t,x)$ is not piecewise continuous, but locally measurable. For such systems, the method of Filippov regularization (Filippov 1988, page 85) is used.

[2] The term almost everywhere is used because there might be points $t \in \mathcal{I}$, constituting a set of zero measure, for which x is not differentiable.

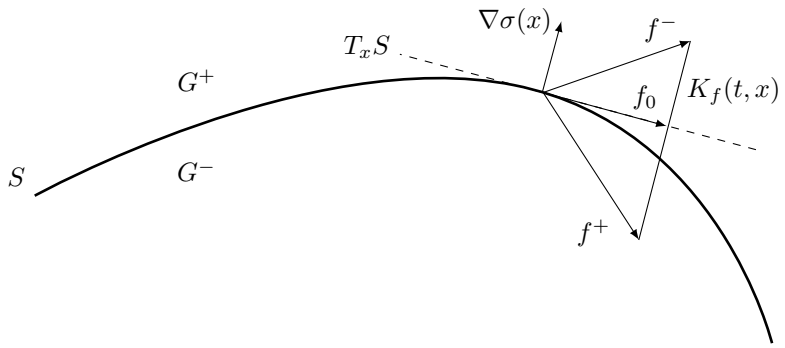


Figure A.1: Geometric illustration of Filippov's definition with f discontinuous on a smooth manifold $S = \{x \in \mathbb{R}^n \mid \sigma(x) = 0\}$.

connecting the end points of vectors f^+ and f^-,

$$K_f(t,x) = \{v \in \mathbb{R}^n \mid v = \alpha f^- + (1-\alpha) f^+\} \quad , \quad (t,x) \in S \quad , \quad \alpha \in [0,1] \tag{A.8}$$

In order to obtain a smooth movement along the manifold, while satisfying the inclusion (A.6), the velocity vector f_0 at $(t,x) \in S$ should belong to the tangent space T_xS of S at (t,x). Since $T_xS = \{v \in \mathbb{R}^n \mid \ <\nabla\sigma(x), v>= 0\}$, the function f_0 is such that $<\nabla\sigma(x), f_0>= 0$, with $f_0 \in K_f(t,x)$ from (A.8). This equation can always be satisfied if $f_N^+ f_N^- < 0$ at (t,x), where $f_N^+ =<\nabla\sigma, f^+>$ and $f_N^- =<\nabla\sigma, f^->$ are the length of the projections of vectors $f_N^+(t,x)$ and $f_N^-(t,x)$ onto $\nabla\sigma(x)$.

Definition 19. *An absolutely continuous function $x : \mathcal{I} \mapsto \mathbb{R}^n$ defined on some interval or segment $\mathcal{I}$ is a solution of (A.3) with f discontinuous only on a smooth manifold $S = \{x \in \mathbb{R}^n \mid \sigma(x) = 0\}$ and $f_N^+ f_N^- < 0$ for every (t,x) if it satisfies*

$$\dot{x} = \begin{cases} f(t,x) & , \quad x \in \mathbb{R}^n \setminus S \\ f_0(t,x) & , \quad x \in S \end{cases} \tag{A.9a}$$

$$f_0 = \frac{f_N^- f^+ - f_N^+ f^-}{f_N^- - f_N^+} \tag{A.9b}$$

almost everywhere on $\mathcal{I}$.

Definition 20. *A solution $x : \mathcal{I} \mapsto \mathbb{R}^n$ of (A.3) in the sense of Definition 19 is said to slide on S, according to the sliding motion equation (A.9b) after it reaches the manifold S.*

Definition 21. *If, for a fixed initial condition x_0, all solutions approach the manifold S and reach it in some finite time $t \in \mathcal{I}$, these solution are said to constitute a sliding family on S.*

Definition 19 assumes $f_N^+ f_N^- < 0$ at (t,x), otherwise $K_f(t,x) \cap T_xS = \emptyset$ and every possible solution satisfying (A.6) will cross S at (t,x). In other words, if on both

sides of the switching manifold S the vector field points to the same side of S, then trajectories necessarily cross the manifold at (t, x).

A.2.2 Nonlinear Systems Affine in the Control Input

In feedback control, it is usual to deal with systems governed by continuous dynamics, but forced by a (possibly) piecewise continuous input. To represent this class of systems, let

$$\dot{x} = f(t, x, u(x)) \tag{A.10}$$

be a modification of (A.3), with $f : \mathbb{R}^{n+1} \times \mathbb{R}^p \mapsto \mathbb{R}^n$ continuous and control input $u : \mathbb{R}^n \mapsto \mathbb{R}^p$ piecewise continuous. Similar to Definition 19, each component u_i of u is assumed to be discontinuous only on a smooth surface $S_i = \{x \in \mathbb{R}^n \mid \sigma_i(x) = 0\}$. Furthermore, (A.10) is simplified to

$$\dot{x} = a(t, x) + b(t, x)u(x) \tag{A.11}$$

with $a : \mathbb{R}^{n+1} \mapsto \mathbb{R}^n$ and $b : \mathbb{R}^{n+1} \mapsto \mathbb{R}^{n \times p}$ continuous. Although simpler than (A.3) and (A.10), system (A.11) covers most practical engineering problems. This modification changes the differential inclusion (A.6) to

$$\dot{x} \in a(t, x) + b(t, x)K_u(x), \tag{A.12a}$$

$$K_u(x) = \begin{bmatrix} K_{u_1}(x) & \dots & K_{u_2}(x) \end{bmatrix}^{\mathsf{T}} \tag{A.12b}$$

$$K_{u_i}(x) = \begin{cases} \{u_i(x)\} & , \quad \sigma_i(x) \neq 0 \\ \alpha u_i^-(x) + (1-\alpha)u_i^+(x) & , \quad \sigma_i(x) = 0 \end{cases} \tag{A.12c}$$

with $\alpha \in [0, 1]$, such that $\alpha u_i^-(x) + (1-\alpha)u_i^+(x)$ is the convex combination of $u_i^+(x)$ and $u_i^-(x)$, defined by the limiting process (A.5).

A.2.3 Nonuniqueness of Solutions

Definition 19 guarantees the existence of solutions of (A.3), but it does not state anything about uniqueness. In fact, by definition of (A.6) on the switching manifold, uniqueness of solutions of (A.3) is not expected. Therefore, for control purposes, a control law should guarantee strong stability properties, since these properties apply for every solution. In this direction, the next section presents strong stability properties based on Lyapunov's stability theory (Lyapunov 1992) and its generalizations for discontinuous systems satisfying (A.6).

A.2.4 Lyapunov Stability and Convergence Properties

Stability is the kernel of most control problems and its modern definition is due to the famous thesis of Lyapunov 1992, originally published in Russian in 1892. Prior to Lyapunov's work, stability analysis consisted of neglecting all terms in f of order higher than k in x[3], usually by applying some linearization procedure, such as Taylor series expansion. One exception to this methodology is the work of Poincare, which inspired some results presented by Lyapunov in his thesis. In contrast to linearization and similar techniques, Lyapunov stability studies how small deviations from the initial condition x_0 of a nominal trajectory (solution) $x^*(t, t_0, x_0)$ impact the resulting motion. That is, what happens to $x^*(t, t_0, x_0')$ when one changes the initial condition to $x_0' \in x_0 + B(\delta)$.

As commonly done in stability analysis, we study stability with respect to the zero solution (origin) $x^*(t, t_0, x_0)$, since any problem can be put in this perspective by an appropriate change of variables $z = x - x^*$. The definitions of Lyapunov stability, asymptotic stability, exponential stability and related properties are not given here, since they are the same as the ones for continuous systems. The reader is directed to (Khalil 2002) for these definitions. Roughly speaking, for an arbitrary vicinity $B(\epsilon)$ of the origin, if one is able to find $B(\delta) : x_0 \in B(\delta)$ for which trajectories do not leave $B(\epsilon)$, then the origin is stable. Additionally, asymptotic stability implies that, on top of stability, trajectories eventually reach the origin, while exponential stability forces the convergence rate to be at least exponential.

To redefine Lyapunov's function method for systems with discontinuous right-hand side, we first revisit the classical method for f continuous. In this case, let a *continuous function* $V \in C^0$ represent, in some sense, the system energy, with $V(x) > 0$ for $x \neq 0$, and $V(0) = 0$. If, for every solution $x(t)$ of (A.3), the function $V(x(t))$ is weakly decreasing, i.e. it is allowed to remain constant, then the origin of (A.3) is stable. Additionally, if $V(x(t))$ is strictly decreasing and tending to zero as $t \to +\infty$, then the origin is asymptotically stable. For $V \in C^1$ *continuously differentiable* the above properties are rewritten in the widely used forms

$$\dot{V}(x) = \nabla^{\mathsf{T}} V(x) f(t, x) \leq 0 \qquad \text{stable origin} \tag{A.13a}$$

$$\dot{V}(x) = \nabla^{\mathsf{T}} V(x) f(t, x) < 0 \qquad \text{asymptotically stable origin} \tag{A.13b}$$

By analyzing these stability conditions for continuous systems, it is clear what must be tackled to obtain an equivalent method for discontinuous systems:

- Even if V is differentiable with respect to the state x, it will not be continuously differentiable, since $f(t, x)$ is piecewise continuous. Therefore, there is set of

[3] In analyzing stability, it is usual to consider a time-invariant system and augment the state vector x by the fictitious state $x_{n+1} = t$, with $\dot{x}_{n+1} = 1$, if the system is time-variant.

measure zero for which $\dot{V}$ is *discontinuous*.

- When the energy function V is not everywhere differentiable, independent of the continuity or not of f, there is a set of measure zero for which $\dot{V}$ is *undefined*.

In the first case, V is differentiable with respect to x, $\dot{V}$ is discontinuous at points $x \in S$ that belong to the switching manifold. When V is not everywhere differentiable, its derivative $\dot{V}$ is (usually) undefined only at $x = 0$. Therefore, we consider generalized derivatives and generalized gradients, and show how these concepts fit to non-smooth stability analysis[4].

Definition 22. *Let $\varphi : \mathbb{R} \mapsto \mathbb{R}$ be a real-valued function and $\{h_n\}$ an infinite sequence converging to zero. A number*

$$D_{h_n}\varphi(t) = \lim_{n\to+\infty} \frac{\varphi(t+h_n) - \varphi(t)}{h_n} \quad , \quad D_{h_n}\varphi(t) \in \bar{\mathbb{R}} \tag{A.14}$$

is called a derivative number. The set of every possible derivative number

$$D_{\mathbb{K}}\varphi(t) = \bigcup_{\{h_n\}\in\mathbb{K}} D_{h_n}\varphi(t) \quad , \quad D_{\mathbb{K}}\varphi(t) \subset \bar{\mathbb{R}} \tag{A.15}$$

with $\mathbb{K}$ the set of sequences $\{h_n\}$ converging to zero, is called the contingent derivative set.

Naturally, at points t where φ is differentiable, it must follow that $D_{\mathbb{K}}\varphi(t) = \{\dot{\varphi}(t)\}$.

Example 2. *Let $\varphi(t) = |t|$. The contingent derivative at $t = 0$ may be obtained by taking any sequence $\{h_n^-\}$ converging to zero from the left*

$$D_{h_n^-}\varphi(0) = \lim_{n\to+\infty} \frac{\varphi(0+h_n^-) - \varphi(0)}{h_n^-} = \lim_{n\to+\infty} \frac{-h_n^-}{h_n^-} = -1 \tag{A.16}$$

and any sequence $\{h_n^+\}$ converging to zero from the right

$$D_{h_n^+}\varphi(0) = \lim_{n\to+\infty} \frac{\varphi(0+h_n^+) - \varphi(0)}{h_n^+} = \lim_{n\to+\infty} \frac{h_n^+}{h_n^+} = 1 \tag{A.17}$$

such that $D_{\mathbb{K}}\varphi(0) = \{-1, 1\}$.

Example 3. *Let $\varphi(t) = t\sin(1/t)$ for $t \neq 0$ and $\varphi(0) = 0$. To find the contingent derivative at $t = 0$ the process is more involved than that of example 2. Considering*

[4]These concepts are presented in a greater extent in (Polyakov and Fridman 2014).

any sequence $\{h_n\}$ *converging to zero,*

$$D_{h_n}\varphi(0) = \lim_{n\to+\infty} \frac{\varphi(0+h_n)-\varphi(0)}{h_n} = \lim_{n\to+\infty} \frac{\varphi(h_n)}{h_n} = \lim_{n\to+\infty} \sin(1/h_n) \tag{A.18}$$

and with $h_n = 1/(2\pi n + \alpha)$, $\alpha \in [-\pi/2, \pi/2]$

$$D_{h_n}\varphi(0) = \lim_{n\to+\infty} \sin(2\pi n + \alpha) = \sin(\alpha) \tag{A.19}$$

such that the contingent derivative is the interval $D_{\mathbb{K}}\varphi(0) = [-1, 1]$.

Let $\mathcal{I} \subseteq \bar{\mathbb{R}}$ be an interval possibly containing $\pm\infty$. A function $\varphi : \mathcal{I} \mapsto \bar{\mathbb{R}}$ is decreasing on $\mathcal{I}$ if and only if

$$\varphi(t_1) \leq \varphi(t_2) \quad , \quad \forall t_1 < t_2 \tag{A.20}$$

Lemma 9. *Let* $\varphi : \mathbb{R} \mapsto \mathbb{R}$ *be defined on some interval* $\mathcal{I}$. *If* $D_{\mathbb{K}}\varphi(t) \leq 0$ *on* $\mathcal{I}$, *then* φ *is decreasing on* $\mathcal{I}$ *and differentiable almost everywhere on* $\mathcal{I}$.

This lemma provides a background on which to develop a discontinuous Lyapunov function method. Nonetheless, it is more general than needed, since no further assumption is made on ϕ. Since Lyapunov function method is applied to nonnegative functions, Lemma 9 is rewritten considering such functions.

Lemma 10. *Let* $V : \mathbb{R} \mapsto \mathbb{R}$ *be nonnegative on an interval* $\mathcal{I}$,

(1) continuous at any $t \in \mathcal{I} : V(t) = 0$ *and*

(2) its contingent derivative $D_{\mathbb{K}}V(t) \leq 0$ *for* $t \in \mathcal{I} : V(t) \neq 0$,

then $V(t)$ *is decreasing and almost everywhere differentiable on* $\mathcal{I}$.

Basically, Lemma 10 states that $V(t)$ converges asymptotically to zero and it is not allowed to leave the origin. It is clear that function $V(t)$ will be used as a Lyapunov function. For that, one still needs to define its derivative when evaluated along the trajectories of (A.12), i.e. determine a chain rule to differentiate $V(x)$. The first step in this direction is the generalization of direction derivatives.

Definition 23. *Let* $V : \mathbb{R}^n \mapsto \mathbb{R}$ *be a real-valued function defined on an open nonempty set* $\Omega \subseteq \mathbb{R}^n$ *and* $\{v_n\}$ *an infinite sequence of real vectors converging* $d \in \mathbb{R}^n$. *A number*

$$D_{\{h_n\},\{v_n\}}V(x) = \lim_{n\to+\infty} \frac{V(x+h_n v_n)-V(x)}{h_n} \quad , \quad D_{\{h_n\},\{v_n\}}V(x) \in \bar{\mathbb{R}} \tag{A.21}$$

is called a directional derivative number. The set of every possible directional derivative number

$$D_{\mathbb{K},\mathbb{M}(d)}V(x) = \bigcup_{\{h_n\}\in\mathbb{K},\{v_n\}\in\mathbb{M}(d)} D_{\{h_n\},\{v_n\}}V(x) \quad , \quad D_{\mathbb{K},\mathbb{M}(d)}V(x) \subset \bar{\mathbb{R}} \tag{A.22}$$

with $\mathbb{M}(d)$ *the set of sequences* $\{v_n\}$ *converging to* d*, is called the contingent directional derivative set.*

As observed for contingent derivatives, at points x where V is differentiable, it follows that $D_{\mathbb{K},\mathbb{M}(d)}V(x) = \{\nabla V(x)^{\mathsf{T}} d\}$.

Finally, considering the derivative of $V(x)$ along the trajectories of x, described by the differential inclusion (A.12), let

$$D_{K_f(t,x)}V(x) = \bigcup_{d\in K_f(t,x)} D_{\mathbb{K},\mathbb{M}(d)}V(x) \tag{A.23}$$

Lemma 11 (Polyakov and Fridman 2014, corollary 2)**.** *Let the set-valued function* $K_f : \mathbb{R}^{n+1} \mapsto 2^{\mathbb{R}^n}$ *in* (A.6) *be defined and upper-semicontinuous on* $\mathcal{I} \times \Omega$ *and set* $K_f(t,x)$ *be nonempty, compact and convex on this domain, where* $\Omega \subseteq \mathbb{R}^n$ *is open and nonempty. Furthermore, let* $V : \mathbb{R}^n \mapsto \mathbb{R}$ *be a nonnegative function on* Ω*. If the inequality*

$$D_{K_f(t,x)}V(x) \leq 0 \tag{A.24}$$

holds for every $t \in \mathcal{I}$ *and* $x \in \Omega : V(x) \neq 0$*, then function* V *is decreasing along the solutions of* (A.6)*.*

The above lemma closely relates to Lyapunov's stability theory of dynamical systems. Nonetheless, before providing a functional approach to Lyapunov stability two definitions are in order.

Definition 24. *A continuous function* $W : \mathbb{R}^n \mapsto \mathbb{R}$ *is said to be positive definite iff* $W(0) = 0$ *and* $W(x) > 0$ *for every* $x \in \mathbb{R}^n \setminus \{0\}$*.*

Definition 25. *A function* $V : \mathbb{R}^n \mapsto \mathbb{R}$ *defined on a nonempty set* $\Omega \subseteq \mathbb{R}^n : \{0\} \in \text{int}(\Omega)$ *is said to be proper on* Ω *iff*

- *it is continuous at the origin;*
- *it is bounded by a positive definite function* $\underline{V} : \mathbb{R}^n \mapsto \mathbb{R}$ *such that* $\underline{V}(x) \leq V(x)$*.*

Furthermore, if $\Omega = \mathbb{R}^n$ *and* $\underline{V}$ *is radially unbounded,* V *is globally proper.*

Theorem 10 (Polyakov and Fridman 2014, theorem 7)**.** *Let a function $V : \mathbb{R}^n \mapsto \mathbb{R}$ be proper on Ω and*

$$D_{K_f(t,x)}V(x) \leq 0 \quad \text{for} \quad t \in \mathbb{R} \text{ and } x \in \Omega \setminus \{0\} \tag{A.25}$$

Then, the origin of system (A.6) is Lyapunov stable. If V is globally proper, then the origin is globally stable.

Note how Theorem 10 is closely related to usual definitions of Lyapunov stability of continuous systems (Khalil 2002). Theorems for asymptotic, exponential and finite-time stability can be readily stated by following the notion of contingent directional derivatives and these properties continuous counterparts (Polyakov and Fridman 2014).

In the design of sliding-mode control laws, one usually defines a piecewise continuous control input $u : \mathbb{R}^n \mapsto \mathbb{R}^p$, where each component u_i is discontinuous only on a smooth surface $S_i = \{x \in \mathbb{R}^n \mid \sigma_i(x) = 0\}$. This was discussed in page 160. For such systems, considering a Lyapunov function $V : \mathbb{R}^p \mapsto \mathbb{R}$ defined on $\Omega_\sigma \subseteq \mathbb{R}^p$, continuous at the origin and differentiable at $\sigma \in \Omega_\sigma \setminus \{0\}$, it follows that

$$D_{K_f(t,x)}V(\sigma) = \nabla V(\sigma)^\mathsf{T}\dot{\sigma} \quad \text{for} \quad \sigma \in \Omega_\sigma \setminus \{0\} \tag{A.26}$$

Corollary 3. *Consider system (A.12), with control input u discontinuous only on a smooth surface $S_i = \{x \in \mathbb{R}^n \mid \sigma_i(x) = 0\}$. Let σ denote the new coordinate system after the change of variables $\sigma(x)$. Furthermore, consider $V : \mathbb{R}^p \mapsto \mathbb{R}$ defined on $\Omega_\sigma \subseteq \mathbb{R}^p$, positive definite and differentiable at every $\sigma \in \Omega_\sigma \setminus \{0\}$. Then, the origin of the differential inclusion (A.12) is Lyapunov stable iff*

$$\nabla V(\sigma)^\mathsf{T}\dot{\sigma} \leq 0 \quad \text{for} \quad \sigma \in \Omega_\sigma \setminus \{0\} \tag{A.27}$$

If V is radially unbounded and $\Omega_\sigma = \mathbb{R}^p$, then the origin is globally stable.

This corollary establishes circumstances for which a discontinuous system stability may be inferred from traditional Lyapunov stability analysis. This results is implicitly used in most of the SMC literature.

A.2.5 Discussion on Definitions other than Filippov's

Filippov's theory was criticized by many authors, mainly because it apparently yields results not adequate to those observed in real application. This happens when the system is nonlinear in the control input, but it is arguable if these inconsistencies are not a consequence of oversimplified models.

Utkin provides an alternate definition, based on the concept of equivalent con-

trol (Utkin 1992). To apply the equivalent control method, one considers system (A.10) and computes a equivalent control input u_{eq} such that trajectories stay on the switching manifold. This is achieved by finding u_{eq} that solves

$$< \nabla\sigma(x), f(t, x, u_{eq}) > = 0 \tag{A.28}$$

This procedure is can be written in terms of the differential inclusion

$$\dot{x} \in f(t, x, K_u(x)) \tag{A.29}$$

with solution defined as in Definition 19 and $K_u(x)$ from (A.12b)-(A.12c). It is clear that, in general, Utkin's equivalent control method allows $\dot{x}$ to take values from a non-convex set $f(t, x, K_u(x))$. a fact that complicates the analysis of the method (Polyakov and Fridman 2014).

A definition that, in a way, merges both Utkin's and Filippov's definitions is given by Aizerman and Pyatnitskii 1974. In short, it consists of taking the convex closure of $f(t, x, K_u(x))$, such that

$$\dot{x} \in \text{co}\left\{f_0(t, x), f(t, x, u_{eq})\right\} \tag{A.30}$$

An important aspect of this definition is that, if some stability property is obtained for (A.30), than the same property holds for both Filippov's and Utkin's definitions (Polyakov and Fridman 2014).

It is not hard to show that the three definitions coincide when f is affine in the control input and the switching manifold S depends solely on x, in fact, they yield the exact same differential inclusions.

Appendix B

Raman Amplifier Power Dynamics and Numerical Coefficients

Consider a simpler average field power model (Bromage 2004) which suffices for Raman amplifier fiber spans. This model is described by N first-order nonlinear partial differential equations (transport PDEs): $\tau_i \frac{\partial P_i}{\partial t} + \mu_i \frac{\partial P_i}{\partial z} = -\alpha_i P_i + \sum_{j=1\ (j\neq i)}^{N} c_{ij} P_i P_j$, where $P_i(t,z)$ is the power (in W) of a propagating signal (data or pump) at time t and distance z from the upstream end of the span (measured along the fiber), corresponding to the wavelength λ_i (data or pump). The constants τ_i, μ_i and α_i denote, respectively, the propagation delays per unit length (μs/km), propagation directions (dimensionless), losses per unit length (1/km). The constant c_{ij} denotes the coupling coefficient per unit length (1/W km) from the signal with wavelength λ_j to the signal with wavelength λ_i propagating in the fiber. This quantities $\tau_i(\lambda_i)$, $\alpha_i(\lambda_i)$ and $c_{ij}(\lambda_i, \lambda_j)$ are functions of the wavelengths propagating in the fiber. In this Thesis, we consider only counter-propagating pump signals, the data signals are co-propagating by default. The propagation direction μ_i $(i = n_p + 1, \ldots, N)$ is equal to -1 for pump signals, and it is equal to $+1$ for data signals.

For a compact form for the power dynamics, consider the following uncertain matrices: $\tau := \operatorname{diag}\left(\begin{bmatrix} \tau_1 & \ldots & \tau_N \end{bmatrix}\right)$, $\mu := \operatorname{diag}\left(\begin{bmatrix} \mu_1 & \ldots & \mu_N \end{bmatrix}\right)$, $A := \operatorname{diag}\left(\begin{bmatrix} \alpha_1 & \ldots & \alpha_N \end{bmatrix}\right)$ and $C \in \mathbb{R}^{N\times N}$, where its ijth entry is given by c_{ij}. The power dynamics can be rewritten as (Dower et al. 2008):

$$\tau \frac{\partial P}{\partial t} + \mu \frac{\partial P}{\partial z} = -AP + \operatorname{diag}(P)CP$$

with $P = \begin{bmatrix} P_1 & \ldots & P_N \end{bmatrix}^{\mathsf{T}}$ partitioned in the form $P = P(t,z) := \begin{bmatrix} P_p(t,z) \\ P_s(t,z) \end{bmatrix} \in \mathbb{R}^N$. These PDEs require the specification of one spatial and one temporal boundary condition per entry in P. The temporal initial condition is specified by an initial

spatial power distribution $P(0,z)$. The spatial boundary conditions are specified by the vector of backward propagating pump powers and the vector of forward propagating data signal powers (Krstić and Smyshlyaev 2008; Krstić 2009).

For the case where 4 data signals and 2 pump signals are considered, the physical parameters τ, A, and C were experimentally characterized by the company *OFS Fitel Denmark Ap.*, consider the following wavelengths (em nm) presented in the fiber: $\Lambda_p = \begin{bmatrix} 1442 & 1480 \end{bmatrix}^\mathsf{T}$ e $\Lambda_s = \begin{bmatrix} 1530 & 1550 & 1570 & 1590 \end{bmatrix}^\mathsf{T}$. The physical parameters are: $10^6\tau$=diag $\left(\begin{bmatrix} 4.876 & 4.877 & 4.877 & 4.878 & 4.878 & 4.879 \end{bmatrix}\right)$, A=diag $\left(\begin{bmatrix} 0.058 & 0.051 & 0.047 & 0.045 & 0.045 & 0.045 \end{bmatrix}\right)$,

$$C = \begin{bmatrix} 0 & -0.23 & -0.59 & -0.62 & -0.18 & -0.10 \\ 0.22 & 0 & -0.17 & -0.25 & -0.39 & -0.57 \\ 0.55 & 0.16 & 0 & -0.12 & -0.15 & -0.21 \\ 0.58 & 0.24 & 0.12 & 0 & -0.11 & -0.14 \\ 0.17 & 0.37 & 0.15 & 0.11 & 0 & -0.11 \\ 0.09 & 0.53 & 0.20 & 0.14 & 0.11 & 0 \end{bmatrix}$$

and $\mu = \text{diag}\left(\begin{bmatrix} -1 & -1 & 1 & 1 & 1 & 1 \end{bmatrix}\right)$.

For the case of 32 data signals, we have chosen 16 wavelengths equally spaced by 100 GHz from 1530 nm to 1541.8 nm for the first group and 16 other wavelengths equally spaced from 1570 nm to 1582.4 nm for the second group. We have also changed the second pump wavelength to 1490 nm. All the coefficient matrices were approximated via interpolation of experimental data, as in (Dower et al. 2008). In particular, the entries of the matrices τ, A and C were determined numerically as follows:

$$\tau_i = f_\tau(\lambda_i)\,, \quad A_i = f_a(\lambda_i)\,, \quad C_{ij} = f_c(\lambda_i, \lambda_j)$$

where

$$f_\tau(\lambda) = \sum_{i=1}^{4} \hat{\eta}_i \lambda^i\,, \quad f_a(\lambda) = \sum_{i=1}^{4} \hat{\alpha}_i \lambda^{i-1}$$

with

$$\hat{\eta} = \begin{bmatrix} 2.48 \times 10^{-4} & -1.03 & 0.022 & 0.022 \end{bmatrix} \times 10^{-14}$$

and

$$\hat{\alpha} = \begin{bmatrix} 4.9 \times 10^{-9} & -1.9 \times 10^{-5} & 2.33 \times 10^{-2} & -8.5 \end{bmatrix}$$

obtained via least square to fit the case of 4 data signals. By using the available Raman gain efficiency, experimentally characterized by the company *OFS Fittel Denmark Ap.*, as a function of the frequency difference between the signal and the pump with wavelength of 1442 nm, it is possible to determine the coefficient of the

matrix C via interpolation by using

$$f_c(\lambda, \eta) = \gamma(\lambda, \eta)\, \Gamma(\text{sgn}(\lambda - \eta)\rho(\lambda, \eta)) \tag{B.1}$$

where

$$\gamma(\lambda, \eta) = \begin{cases} \frac{1}{4} & , \text{ if } \lambda > \eta \\ 0 & , \text{ if } \lambda = \eta \\ \frac{-\eta}{4\lambda} & , \text{ if } \lambda < \eta \end{cases}$$

$\Gamma(\zeta) = \sum_{i=1}^{3} \hat{\kappa}_i \exp\left(-\hat{\beta}(\zeta - \hat{\zeta}_i)^2\right)$, with

$$\hat{\kappa} = \begin{bmatrix} 1.813 \times 10^{-1} & 5.652 \times 10^{-1} & -3.117 \times 10^{-2} \end{bmatrix}$$

and

$$\hat{\zeta} = \begin{bmatrix} 4.250 & 1.282 \times 10^{1} & 4.176 \times 10^{-1} \end{bmatrix}$$

$\rho(\lambda, \eta) = 2.9979 \times 10^5 \left(\frac{1}{\eta} - \frac{1}{\lambda}\right)$ and $\hat{\beta} = 5.1509 \times 10^{-2}$. This choice of the interpolates came from (Dower et al. 2008), which is commonly used in optical fibers.