Ecole Centrale de Nantes - Universidad de Los Andes

École Doctorale

Sciences et Technologies

de l'Information et de Mathematiques

Année 2009 Provisional Version September 20 2009

Thèse de Doctorat

Diplôme délivré par L'École Centrale de Nantes

Spécialité : Automatique et Traitement du Signal

Présentée et soutenue publiquement par:

Eduardo Mojica-Nava

le 22 septembre 2009

à l'École des Mines de Nantes, Nantes

Titre

A Polynomial Approach for Analysis and Optimal Control of Switched Nonlinear Systems

Jury

Rapporteurs :	Peter Caines - Professor, McGill University (Montreal, Canada)
	Didier Henrion - CNRS Researcher, LAAS (Toulouse, France)
Membre invité du jury :	Pierre Riedinger - Maître de Conférences, CRAN-INPL (Nancy, France)
Examinateurs :	
Directeur France:	Jean Jacques Loiseau - Directeur de Recherche CNRS, IRRCyN (Nantes, France)
Directeur Colombie:	Alain Gauthier - Professor, Univ. de Los Andes (Bogotá, Colombie)
Co-encadrant France:	Naly Rakoto-Ravalontsalama - Maître Assistant, EMN (Nantes, France)
Co-encadrant Colombie:	Nicanor Quijano - Assistant Professor, Univ. de Los Andes (Bogotá, Colombie)

Para mis padres,

Alirio y Cecilia,

por su apoyo incondicional todos estos años

ACKNOWLEDGEMENTS

First and foremost, I would like to express deep gratitude towards my advisors, Alain Gauthier, Naly Rakoto, Jean-Jacques Loiseau and Nicanor Quijano. Alain, for his guidance and support all these years, I will be forever thankful for this opportunity. Naly, for providing a unique working environment that I mostly enjoyed and also for his encouragement during the time at EMN. Jean-Jacques, for let me be part of the research group ACSED under his direction, and also for his many interesting advices. At last but not the least, my greatest appreciation goes to my closest advisor, Nicanor. His advice and friendship lead me through the life of a graduate student, Nick made my Uniandes experience terrific, creating numerous opportunities for me to learn and grow as a researcher.

I also thank the other members of my doctoral committee: Peter Caines, Didier Henrion and Pierre Riedinger, for their encouragement, generosity, and insightful discussions. Further, I'm grateful to Rene Meziat for our long conversations, many of which found their way into the pages of this thesis.

To all my friends at Uniandes: Andrés Pantoja, Iván muñoz, and GIAP group gang, thank you for the Café and Coffee evenings- you have been great. Further, I thank to my closests friend at EMN, Yonatan and Chio, for their noble hearts. I also want to mention Roberto Bustamante and Ángela Cadena, professors at Uniandes, for their good humor and inspiration.

Finally, I'm forever grateful to my parents, Alirio and Cecilia, and the rest of my family for their constant encouragement, confidence, friendship, and unconditional love for all these years. And also, I want to express my endless love to Carito, "mi mitis", who even though was for almost a year 8,960 km away, always waited for me.

Une Approche polynomiale pour l'analyse et la commande optimale des systèmes non-linéaires à commutation

Résumé : Dans cette thèse nous étudions comment la géométrie semi-algébrique convexe et l'optimisation polynomiale globale peuvent être employées pour analyser et concevoir les systèmes non linéaires à commutations. Pour traiter l'analyse de stabilité des systèmes non-linéaires à commutations on montre que la transformation du problème original à commutations en un système polynômial nous permet d'employer l'inégalité de dissipation pour les systèmes polynômiaux. Avec cette méthode et d'un point de vue théorique, nous fournissons une manière alternative de rechercher une fonction commune de Lyapunov pour les systèmes non linéaires à commutations.

L'idée principale derrière l'approche proposée est d'inclure dans l'analyse fonctionnelle les contraintes cachées. Nous devons vérifier le définition semi-négative de dV/dt en ce qui concerne l'ensemble de contraintes. Pour cela, nous employons l'idée de la pénalisation utilisée dans la théorie d'optimisation avec contraintes. Une nouvelle fonction de pénalisation ou multiplicateur de Lagrange est introduite. Cette idée est basée sur des résultats pour les systèmes de commande contraints, où nous pouvons employer le concept d'inégalité de dissipation utilisant des fonctions de stockage et des taux d'approvisionnement. Ainsi nous étendons les résultats à une classe plus générale des systèmes commutés, ceux modélisés par des fonctions élémentaires. Cette classe de fonctions provient des dérivés symboliques explicites, telles que la fonction exponentielle, le logarithme, les fonctions trigonométriques, et les fonctions hyperboliques. Pour ce faire, en utilisant un processus de réécriture, le système est transformé sous une forme polynômiale équivalente. Ainsi les résultats obtenus précedemment sur l'analyse de stabilité des systèmes polynômiaux peuvent être utilisés.

La commande optimale pour les systèmes non-linaires commutés est également étudiée. Nous proposons une approche alternative pour résoudre le problème de commande optimale pour un système non linéaire autonome à commutations, basé sur le principe de maximum généralisé (GMP). L'essentiel de cette méthode est la transformation d'un problème de commande optimale non-linéaire et non-convexe, c'est--dire, le système commuté, en un problème de commande optimale équivalent avec la structure linéaire et convexe, qui permet d'obtenir une formulation convexe équivalente plus appropriée pour être résolu par le calcul numérique plus efficace. En conséquence, nous proposons de convexifier les variables d'état et de commande au moyen de la méthode des moments afin d'obtenir des programmes SDP. Une généralisation pour résoudre le problème de commande optimale des systèmes commutés non-linéaires est ensuite étudiée après à partir du processus réécrit.

En conclusion, nous étudions l'application industrielle obtenue par une approximation linéaire par morceaux de la croissance cellulaire non-linéaire en utilisant des fonctions canoniques linéaires orthonormales par morceaux. Elle est commandée par une stratégie de "probing control". Nous traitons les cellules mammifres BHK (rein de bébé hamster) dans un bio-réacteur. Les résultats de simulation prouvent que cette approximation linéaire par morceaux est bien adaptée pour modeléliser une telle dynamique non-linéaire.

Mots clés : Optimisation convexe, commande optimale, systèmes commutés, analyse de stabilité

A Polynomial Approach for Analysis and Optimal Control of Switched Nonlinear Systems

Abstract: In this dissertation, we investigate how convex semialgebraic geometry and global polynomial optimization can be used to analyze and to design switched nonlinear systems. To deal with stability analysis of switched nonlinear systems it is shown that the representation of the original switched problem into a continuous polynomial system allows us to use the dissipation inequality for polynomial systems. With this method and from a theoretical point of view, we provide an alternative way to search for a common Lyapunov function for switched nonlinear systems.

The main idea behind the proposed approach is to include in the system analysis the hidden constraints. We need to check negative semidefinitness of $\dot{V}$ with respect to the constrained set. In order to do that, we use the idea of penalization used in optimization theory with constraints. For that, we use a function $\lambda(x, s)$, which can be interpreted as a penalization function or a Lagrange multiplier. This idea is based on some results for constrained control systems, where we can use the dissipation inequality concept using storage functions and supply rates. We then extend the results to a more general class of switched systems, those modeled by elementary and nested elementary functions. This class of functions is related with explicit symbolic derivatives, such as exponential, logarithm, power-law, trigonometric, and hyperbolic functions. For this aim, we transform, using a recasting process, the system obtained by the equivalent representation in a system with polynomial form, and then, we use the results of the previous Section for stability analysis.

Besides stability analysis, optimal control problems for switched nonlinear systems are also investigated. We propose an alternative approach for solving effectively the optimal control problem for an autonomous nonlinear switched system based on the

Generalized Maximum Principle (GMP). The essential of this method is the transformation of a nonlinear, non-convex optimal control problem, i.e., the switched system, into an equivalent optimal control problem with linear and convex structure, which allows us to obtain an equivalent convex formulation more appropriate to be solved by high performance numerical computing. Consequently, we propose to convexify the state and control variables by means of the method of moments obtaining SDP programs. A generalization to solve the optimal control problem of nonlinear switched systems is investigated next based on the recasting process.

Finally, we concentrate in the industrial application obtaining a piecewise-linear approximation of nonlinear cellular growth using orthonormal canonical piecewise linear functions, and it is tested by a probing control strategy for the feed rate. We deal with the mammalian cells BHK (Baby Hamster Kidney) in bioreactor in batch, fed-batch and continuous mode operation. Simulation results show that this piecewise linear approximation is well suited for modeling such nonlinear dynamics.

Keywords: Convex optimization, Optimal Control, Polynomial systems, Switched systems, Stability analysis

---

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

1.1 Introductory Remarks and Motivation

Hybrid systems arise in a wide variety of practical systems. We start pointing out that the term *hybrid system* can be understood from several point of views. From the technological point of view, systems that contain analog and digital components, systems that comprise part of different physical natures such as biological, chemical, electrical, electronic, hydraulic, mechanical ones, more generally, settings that involve the use of computers for control propose are termed hybrid systems. From a mathematical modeling point of view, systems described of different forms, such as algebraic equations, difference equations, ordinary differential equations, logical equations, and partial differential equations, are hybrid system. In the opening article of the European Journal of Control [21] (1995), hybrid systems take a more prominent position. Hybrid systems are mentioned among the major open problems in systems and control by several respondents including Lennart Ljung, Peter Caines, and Pravin Varaiya. The thrust of the thinking on the subject can be seen from Vidyasagar's remark:

> *Another interesting question is: 'How can one combine differential/difference equations with logical switches so as to enhance performance?' In some sense, this is the central question of intelligent control.*

It seems therefore that by the mid-nineties, hybrid systems have been clearly identified as a major new research area for systems and control theory, and they still constitute a relatively new and very active area of current research (e.g., [9], [72], [1], [2], [3], [45], [113], [4], [5], [6]). In spite of this currently interest for hybrid systems, they have been with us at least since the days of the relay. The earliest reference we know of is the work of Witsenhausen from MIT, who formulated a class of hybrid-state continuous-time dynamic systems and investigated an optimal control problem [108]. It is worthwhile mention that there are various models for hybrid systems, due to its inherently interdisciplinary nature, the field has attracted the attention of people with diverse backgrounds, primarily computer scientists, applied mathematicians, and engineers [54], [100], [55]. However, we consider continuous-time systems with discrete switching events, that consist of several subsystems and a switching law, which determines the switching times and mode transitions. Such systems are called switched systems and can be viewed as higher-level abstraction of hybrid systems [54]. Switched system modeling of any real-process dealing with physical variables are in agreement with the time continuous and uniqueness principle, i.e., the value of every physical variable changes only continuously in time through every intermediate value (initial and final), and by possessing a unique value at a specific time and space. Any synthesized control should be uniquely defined and continuous in time. Recent efforts in switched systems research have been typically focused on the analysis of dynamic behaviors, such as stability, controllability and observability, and optimal control among others (e.g., [100], [56], [54]).

In this dissertation, we deal with three different problems of switched systems:

- stability analysis under arbitrary switching,
- optimal control problem,
- and piecewise linear model and control of a bioreactor.

The first two problems are related by means of an equivalent polynomial representation. The third problem is an industrial application which uses a class of switched system when the switching law is decided by the partition of the state space. We present a brief introduction of each subject of the dissertation with its corresponding chapter.

1.2 Contributions, Literature Review, and Outline

1.2.1 For Stability Analysis

We deal with the stability analysis of switched non-linear systems under arbitrary switching. Most of the efforts in switched systems research have been typically focused on the analysis of dynamical behavior with respect to switching signals. Several methods have been proposed for stability analysis (see [54], [56], and references therein), but most of them have been focused on switched linear systems. Stability analysis under arbitrary switching is a fundamental problem into the analysis and design of switched systems. For this problem, it is necessary that all the subsystems are asymptotically stable. However, in general, the above stability condition is not sufficient to guarantee stability for the switched system under arbitrary switching. It is well known that if there exists a common Lyapunov function for all the subsystems, then the stability of the switched system is guaranteed under arbitrary switching. Previous attempts for general constructions of a common Lyapunov function for switched non-linear systems have been presented in [30] and [58], using converse Lyapunov theorems. Also in [105], a construction of a common Lyapunov function is presented for the case when the individual systems are handled sequentially rather than simultaneously for a family of pairwise commuting systems. These methodologies are presented in a very general framework, and even though they are mathematically sound, they are too restrictive from a computational point of view, mainly because it is usually hard to check for the set of necessary conditions for a common function over all the subsystems (it could not exist). Also, these constructions are usually iterative, which involves running backward in time for all possible switching signals, being prohibitive when the number of modes increase.

The main contribution of Chapter 2 is twofold. First, we present a reformulation of the switched system as a differential continuous system on a constraint manifold. This representation opens several possibilities of analysis and design of switching systems in a consistent way, and also with numerical efficiency [69], [70], which is possible thanks to some tools developed in the last decade for polynomial differential-algebraic equations analysis [32], [75], [44]. The second contribution is to show an alternative method to search for a common Lyapunov function for switched systems with an efficient numerical method, using results for stability analysis of polynomial based on dissipativity theory [31], [70]. We propose a methodology to construct common Lyapunov functions for switched non-linear systems, which provides a less conservative test for proving stability under arbitrary switching. It has been mentioned in [82] that the sum of squares decomposition, presented only for switched polynomial systems, can sometimes be made for a system with a non-polynomial vector fields. Those cases, where it could be possible, are restricted to subsystems, which after the rendering in polynomial forms using auxiliary variables, preserve all the same dimension. However, to our knowledge this has not been shown in the literature. The methodology that we propose does not have the dimentionality limitation mentioned above. In a previous work [70], we have presented the method only for the case when all the subsystems are in a polynomial form. Later, we extend some preliminary results to a more general non-linear case, and a representative example is presented to show the effectiveness of the methodology by reliable and efficient numerical methods. Basically, this theory is based in terms of an inequality involving a generalized system power input, or, supply rate, and a generalized energy function, or storage function [107]. The interpretation of this storage function establishes the connection between Lyapunov stability and dissipativity. Stability problems can be solved once the dissipativity property is assured, and the storage function becomes a Lyapunov function, which can be used to construct Lyapunov functions for nonlinear dynamical systems. As for a common Lyapunov function, a single storage function for all subsystems is usually difficult to find or may not exist (computational problems arise when a common function needs to be found). However, thanks to the computational tools that have been developed lately, we are able to use dissipativity theory with efficient numerical methods to establish a common Lyapunov function for the equivalent polynomial system.

Alternatively, the authors in [54] propose a Lie algebraic condition for switched LTI systems, which is based on the solvability of the Lie algebra generated by the set of state matrices. The Lie algebraic condition is also extended to switched nonlinear systems to obtain local stability results based on Lyapunov's first method. Most recently global stability properties for switched nonlinear systems is presented in [105], [59], [60], and a Lie algebraic global stability criterion is derived based on Lie brackets of the nonlinear vector fields. This sort of analysis based on algebraic conditions and Lie algebra are not considered in this work.

1.2.2 For the Optimal Control Problem

1.2.2.1 A Brief Literature Review for Optimal Control of Hybrid Systems

The earliest reference we know of optimal control for hybrid systems is the seminal paper of Witsenhausen [108](1966), where an optimal terminal constraint problem is considered on his hybrid systems model. Later in [94], an optimal control for switching systems is presented, followed by the influential work [24], and [23] where the authors compare several algorithms for optimal control and discuss general conditions for the existence of optimal control laws. Eventually, necessary optimality conditions for hybrid systems are derived using general versions of the maximum principle [101], [88], [86] and more recently in [37], and in particular for switching systems in [19] and [106], where the switched system is embedded into a larger family of systems and the optimization problem is formulated. In some recent papers [97], [29] we can find some work related with embedding approach for the linear case. However, they do not provide further insights on how to find the optimal switching strategy. For general hybrid systems, with nonlinear dynamics in each location and with autonomous and controlled switchings, necessary optimality conditions are recently presented in [95], and using these conditions, algorithms based on the hybrid maximum principle are derived. An approach based on the parameterization of the switching instants and the differentiation of the cost function is presented in [109], [110], [111]. The algorithm proposed is based on a two-stage optimization problem. However, the method encounters major computational difficulties when the number of available switches increases.

Lincoln and Rantzer [57] present the method *relaxing dynamic programming* to approximate hybrid optimal control laws and to compute lower and upper bounds of the optimal cost, while the case of piecewise-affine systems is discussed by Rantzer and Johansson [84]. For determining the optimal feedback control law, these techniques require the discretization of the state space in order to solve the corresponding Hamilton-Jacobi-Bellman equations that make it intractable numerically.

For discrete-time linear hybrid systems, Bemporad and Morari introduce a hybrid modeling framework that, in particular, handles both, internal switches (i.e., caused by the state reaching a particular boundary), and controllable switches [16]. The authors also show how mixed-integer quadratic programming (MIQP) can be used to determine optimal control sequences. On the other hand, it is generally perceived that the best numerical methods available for hybrid optimal control problems involve mixed integer programming (MIP). While great progress has been made in recent years in improving these methods, the MIP is an NP-hard problem so scalability is problematic [106]. They have worked in model predictive control with different problems (e.g., constrained finite time optimal control (CFTOC), constrained infinite time optimal control (CITOC)) [17]. In [71] Morari and Baric present the recent developments to control constrained hybrid systems, in which the control paradigm is focused on MPC, with the emphasis on explicit solution. Nonlinear parametric optimization using cylindrical algebraic decomposition is presented in [34], [35].

For cases where online optimization is not viable, Seatzu *et al* proposes a multi-parametric programming for solving in state-feedback form the infinite-time hybrid optimal control, by showing that the resulting optimal control law is piecewise affine [93]. They consider the optimal control of continuous-time switched affine systems with a piecewise-quadratic cost function by two methods. i) the so-called master-slave procedure (MSP), and ii) the switching table procedure (STP). The drawback of all these approaches is that they take a lot of computing time.

Focus on real-time application, Egerstedt *et al* [33] consider an optimal control problem for switched dynamical systems, where the objective is to minimize a cost functional defined on the state and where the control variable consists of the switching times. A gradient-descent algorithm is proposed based on an especially simple form of the gradient of the cost functional. In [12] and [22] the authors deal with the problems of mode-switching with unknown initial state and the construction of a surface for optimality. Such systems change modes whenever the state intersects certain surfaces that are defined in the state space.

In [52] a control parameterization enhancing transform is presented with pre-specified order of the sequence of subsystems, where the switching instants are included in the cost functional. Both, the switching instants and the control function are to be chosen such that the cost functional is minimized. In [11], [8] an algorithm based on strong variations to handle constraints on both locations and switching instants is proposed for switched nonlinear systems. With the advent of differential inclusion theory, some results using this technique are presented by Vinter in [36], in which the continuous subsystems are modeled as differential inclusions. A distinctive feature of the analysis is that it permits an infinite set of discrete states.

Whereas, the H_∞ control problem for nonlinear switched systems is addressed in [116]. Where, based on multiple Lyapunov functions, a sufficient condition for the problem to be solved is derived in terms of partial differential inequalities. The continuous controllers for each subsystem and the switching law are simultaneously designed.

1.2.2.2 Contributions on Optimal Control of Switched Systems

The main contribution of Chapter 3 is an alternative approach for solving effectively the optimal control problem for an autonomous nonlinear switched system based on the Generalized Maximum Principle (GMP) introduced in [112], and later used in [78], and [73] to establish existence conditions for an infinite-dimensional linear program over a space of measure. At a first stage, we focus our analysis on vector fields and running costs that are of polynomial form. However, it is well known that functions called nested elementary functions, can be recasted exactly in a polynomial systems with a larger state dimension. Therefore, we will use the fact that all system data are polynomial after the recasting process. Then, we apply the *Theory of Moments*, a method previously introduced for global optimization in [46], [47], [49], [61], and for variational calculus in [63] and with some previous results recently introduced for optimal control problems in [51, 49, 80, 62, 40, 50, 69]. The moment approach for global polynomial optimization based on semidefinite programming (SDP) is consistent as it simplifies and/or has better convergence properties when solving convex problems [48]. This approach works properly when the control variable (i.e., the switching signal), and the state variables can be expressed as polynomials. The essential of this method is the transformation of a nonlinear, non-convex optimal control problem (i.e., the switched system), into an equivalent optimal control problem with linear and convex structure, which allows us to obtain an equivalent convex formulation more appropriate to be solved by high performance numerical computing. In other words, we transform a given controllable switched system, into a controllable continuous polynomial system with a linear and convex structure. If we would use a nonlinear, non-convex form of the control variable, we would not be able to use the Hamilton equations of the maximum principle, and nonlinear mathematical programming techniques. That would entail severe difficulties, either in the integration of the Hamilton equations or in the search method of any numerical optimization algorithm. Consequently, we propose to convexify the state and control variables by using the method of moments in the polynomial expression in order to deal with this kind of problems. Finally, we use our previous work, where we have limited our analysis to vector fields and running costs that are of polynomial form, to extend the result to a more general nonlinear switched systems by means of the ideas introduced in [92] that help us to cope with these non-polynomial terms, which are based on the recasting process of a specific kind of non-polynomial functions.

1.2.3 For the Piecewise Linear Model and Control of a Bioreactor

Mammalian cells of Baby Hamster Kidney (BHK) are used in the production of the vaccine against the foot-and-mouth disease. These cells display multiple steady states with widely varying concentrations of cell mass, desired products as well as waste metabolites [74], [65], [64]. It means that for identical input conditions to a fed-batch reactor, the outlet conditions change depending on how the culture is made fed-batch. These multiple states are manifestations of the complex interaction between cells and their environment. What make this process difficult is the additional level of complexity present in biological systems because of the genetic information in living cells. Several nonlinear models have been developed for mammalian cells (see [64] and references therein), but most of them arise in computational problems. Usually, for nonlinear models from the point of view of control design, details about intracellular metabolism are omitted. The models are based on macroscopic mass balance, and include only the more relevant biological reactions. In spite of these attempts to find simple but useful nonlinear models, the modeling of the reaction kinetics are generally represented by rational functions of the state and numerous studies have shown that modeling of the kinetics is a very difficult task [14].

The peculiar features of mammalian cells growth in fed-batch operating condition are addressed. The task of the controller is to determine, in every instant, the best feed substrate, using the compilation of information on-line from the sensor. The determination of an optimal strategy of feed substrate using the nonlinear modeling either if the kinetics are known, is not a straightforward matter and is often further complicated by the presence of constraints imposed on the state variables [14]. All of these difficulties in modeling and control design of a biological process using nonlinear models are driving to the search for new and more efficient tools for both modeling and control design. In this context, hybrid systems, i.e., including both continuous and discrete dynamics, open several possibilities for both modeling and control design. Chapter 4 is related with a modeling class of hybrid systems, piecewise-linear (PWL) systems. The PWL approximation, i.e., systems which are linear or affine on each of the components of a polyhedral partition of the state space [96], have shown advantages on implementation, performance analysis, and calculations [39], [84], [90], [89].

The problem to find a piecewise-linear model given a nonlinear model has been previously treated ([96], [39], [84], [90], and some others), in specific biological systems [43]. More recently, an approximation for modeling of gene regulatory networks is presented in [13]. However, these approaches present many parameters to be provided by the designer, and finding these parameters is a difficult task, even for simple systems. In this work, a canonical piecewise linear approximation over simplicial partitions is used. It provides a partition of the state space into polytopic cells based on value at vertices [42], [89], [38]. This choice is motivated by several facts. First, this class of functions uniformly approximate any continuous nonlinear function defined over a compact domain $\mathbb{R}^n$ (see [42]). Moreover, the canonical expression introduced in [42] uses the minimum and exact number of parameters, and it is the first PWL expression able to represent PWL mappings in arbitrary dimensional domains. As a consequence of this, an efficient characterization is obtained from the viewpoint of memory storage and numerical evaluation [26]. Second, the approximation can be used in real implementations, the points taken from nonlinear model may be replaced for points taken from sensors or data directly from the process, namely, it addresses the problem of finding a PWL approximation of system where a reasonable number of measure samples of the vector field is available (regression set) [98]. Third, this alternative approach deals with an approximation which is easier to handle than the nonlinear model. In fact, it might use many tools developed for hybrid systems, e.g., the MLD model based approach [16], since algorithms for translating MLD systems into PWL systems are available [15], [104]. Finally, this CPWL is used in a model based control termed, probing control in [68], being a first step to develop a hybrid probing control. The task of the controller is to determine, in every instant, the best control action (the best feed substrate) based on the compilation of on-line information of the sensor (or for the nonlinear model). The fact that the probing strategy for feedback control requires a minimum of process knowledge is exploited in [7]. This work refers to a probing control as it is presented in [7] for *E. coli.* Short pulses to the feed rate are added, and taking into account the system response, the pulse is increased or decreased according with the tuning rule. The probing control strategy avoids acetate accumulation while maintaining a high growth rate [7], [103].

The main contribution of Chapter 4 is a hybrid dynamical model using orthonormal high level canonical piecewise linear functions [67], [66]. The approximation model is tested by a recently presented control methodology, the probing control strategy, which was developed in [7] for *E. coli* cultivations. It is implemented in simulations for this mathematical model [68]. The comparative analysis and error approximation between this new biological model and a nonlinear model developed first in [65], [64] are shown. This method is satisfactory for implementation's purposes of a hybrid probing control [68].

CHAPTER II

A POLYNOMIAL APPROACH FOR STABILITY ANALYSIS OF SWITCHED SYSTEMS

The stability analysis of switched non-linear systems, i.e., continuous systems with switching signals under arbitrary switching is treated in this chapter. Stability analysis under arbitrary switching is a fundamental problem into the analysis and design of switched systems. For this problem, it is necessary that all the subsystems are asymptotically stable. However, in general, the above stability condition is not sufficient to guarantee stability for the switched system under arbitrary switching. It is well known that if there exists a common Lyapunov function for all the subsystems, then the stability of the switched system is guaranteed under arbitrary switching.

In this chapter, we present a reformulation of the switched system as a differential continuous system on a constraint manifold. This representation opens several possibilities of analysis and design of switching systems in a consistent way, and also with numerical efficiency [69], [70], which is possible thanks to some tools developed in the last decade for polynomial differential-algebraic equations analysis [32], [75], [44]. Using this representation we develop an alternative method to search for a common Lyapunov function for switched systems with an efficient numerical method, using results for stability analysis of polynomial based on dissipativity theory [31], [70]. We propose a methodology to construct common Lyapunov functions for switched nonlinear systems, which provides a less conservative test for proving stability under arbitrary switching. In Section 2.4, we extend preliminary results to a more general nonlinear case, and a representative example is presented to show the effectiveness of the methodology by reliable and efficient numerical methods. Basically, this theory is based in terms of an inequality involving a generalized system power input, or, supply rate, and a generalized energy function, or storage function [107]. The interpretation of this storage function establishes the connection between Lyapunov stability and dissipativity. Stability problems can be solved once the dissipativity property is assured, and the storage function becomes a Lyapunov function, which can be used to construct Lyapunov functions for nonlinear dynamical systems.

2.1 Definitions and Preliminaries

2.1.1 Basic Concepts

A switched system is a system that consists of several continuous-time systems with discrete switching events. A switched system may be obtained from a hybrid system by neglecting the details of the discrete behavior and instead considering all possible switching patterns. Switched systems have many application examples, such as power electric circuits, automotive controllers, chemical processes, etc [54].

The mathematical model can be described by

$$\dot{x}(t) = f_{\sigma(t)}(x, t), \tag{1}$$

where the state $x \in \mathbb{R}^n$, $f_i : \mathbb{R}^n \times \mathbb{R}^+ \mapsto \mathbb{R}^n$ are vector fields, and $\sigma(t) : [0, t_f] \rightarrow \mathcal{Q} \in \{0, 1, ..., q\}$ is a *piecewise constant* measurable function of time. Every mode of operation corresponds to a specific subsystem $\dot{x}(t) = f_i(x, t)$, for some $i \in \mathcal{Q}$, and the *switching signal* $\sigma(t)$ determines which subsystem is active at each point in time on the time interval $[0, t_f]$, with t_f as the final time. No assumptions about the number of switches nor about the mode sequence are made. In addition, we consider non Zeno behavior, i.e., we exclude an infinite switching accumulation points in time. Finally, we assume that the state does not have jump discontinuities.

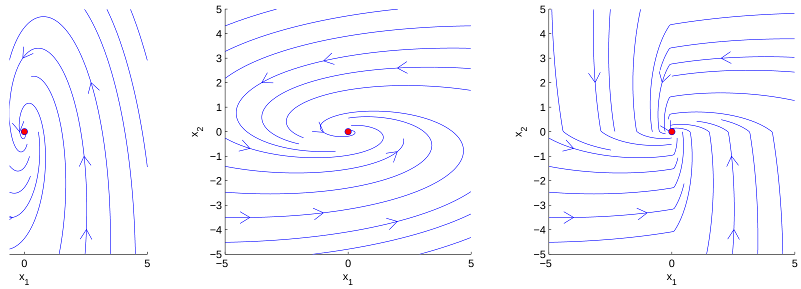

Figure 1: Switching between stable systems becoming stable

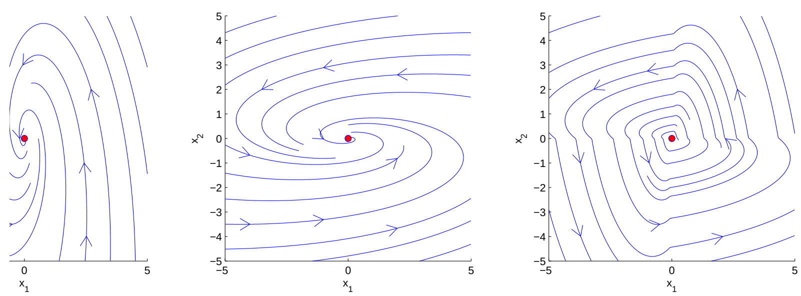

Figure 2: Switching between stable systems becoming unstable

2.1.2 Stability Analysis Under Arbitrary Switching and Dissipativity

The stability problem presents several interesting phenomena. For example, even when all the subsystems are exponentially stable, the switched system may be stable (see Figure 1) or may have divergent trajectories for certain switching signals (see Figure 2). Other fact is also possible, one may carefully switch between unstable subsystems to make the switched system exponentially stable [54]. We can see from these examples that the stability of switched systems depends not only on the dynamics of each subsystem but also on the properties of switching signals. Therefore, the stability study of switched systems can be roughly divided into two kinds of problems. One is the stability analysis of switched systems under given switching signals (maybe arbitrary, slow switching, etc.); the other is the synthesis of stabilizing switching signals for a given collection of dynamical systems. Here, we are dealing with the stability analysis of switched systems under arbitrary switching, i.e., the switched system state goes to zero asymptotically for any switching sequence. If this holds for any initial conditions, we have global uniform asymptotic stability (GUAS) [54], [56]. For this problem, it is necessary to require that all the subsystems be asymptotically stable. However, in general, the above subsystems' stability assumption is not sufficient to assure stability for the switched systems under arbitrary switching, except for some special cases. On the other hand, if there exists a common Lyapunov function for all the subsystems, then the stability of the switched system is guaranteed under arbitrary switching. This provides us with a possible way to solve this problem, and a lot of efforts have been focused on the common quadratic Lyapunov functions [56].

2.1.2.1 Common Lyapunov functions

We are interested in obtaining a Lyapunov condition for GUAS. We proceed using the classic Lyapunov formulation. Given a positive definite continuously differentiable function $V : \mathbb{R}^n \to \mathbb{R}$, we say that it is a common Lyapunov function for the family of systems (1) if there exists a positive definite continuous function $W : \mathbb{R}^n \to \mathbb{R}$ such that we have

$$\frac{\partial V}{\partial x} f_i(x) \leq -W(x) \; \forall x, \quad \forall i \in \mathcal{Q}.$$

Theorem 1 *If all systems in the family (1) share a radially unbounded common Lyapunov function, then the switched system is GUAS.*

Theorem 1 is well known and can be derived in the same way as the standard Lyapunov stability theorem [54]. The main point is that the rate of decrease of V along solutions is not affected by switching, hence asymptotic stability is uniform with respect to σ.

2.1.2.2 A converse Lyapunov Theorem

The question now arises whether the existence of a common Lyapunov function is a more severe requirement than GUAS. A negative answer to this question - and a justification for the common Lyapunov function approach- follows from the converse Lyapunov theorem for switched systems [58], [30], [54] which claims that the GUAS property of a switched system implies the existence of a common Lyapunov function.

Theorem 2 *Assume that the switched system (1) is GUAS, the set $\{f_i(x) : i \in \mathcal{Q}\}$ is bounded for each x, and the function $(x, i) \mapsto f_i(x)$ is locally Lipschitz in x uniformly over i. Then all systems in the family (1) share a radially unbounded smooth common Lyapunov function.*

There is a useful result which we find convenient to state here as a corollary of Theorem 2. It could be shown that if the switched system is GUAS, then all convex combinations of the individuals subsystems from the family (1) must be globally asymptotically stable. These convex combinations are defined by the vector fields

$$f_{p,q,\alpha}(x) := \alpha f_p(x) + (1-\alpha) f_q(x), \quad p, q \in \mathcal{Q}, \;\; \alpha \in [0, 1].$$

Corollary 3 *Under the assumption of Theorem 1, for every* $\alpha \in [0, 1]$ *and all* $p, q \in \mathcal{Q}$ *the system*

$$\dot{x} = f_{p,q,\alpha}(x)$$

is globally asymptotically stable.

2.1.2.3 Dissipativity

A switched system expressed as a polynomial differential-algebraic system allows us to establish an alternative approach for stability analysis. But, instead of searching for a common Lyapunov function in order to provide stability under arbitrary switching using traditional techniques (e.g., search a single Lyapunov function whose derivative along solutions of all systems in the family (1)) satisfies suitable inequalities, which are usually very restrictive techniques based on exhaustive algorithms as it is mentioned in the introduction of this chapter, we look for a Lyapunov function using techniques developed for polynomial continuous systems. It means that we can find a common Lyapunov function using dissipativity inequalities as in [31], or study stability of constrained dynamical systems [81]. With this reformulation, we are dealing with a polynomial differential system on a manifold. Basically, the stability problem of differential-algebraic systems is related to the problem of stability on manifolds, which are defined by the constraints in the system description. From the concept of dissipativity, it could be inferred that storage functions induced by dissipativity are possible Lyapunov functions candidate for stability analysis, which implies that stability and stabilization problems can be solved once the dissipativity property is assured [115]. It is possible to show that if the system is expressed as a purely passive system the origin is an asymptotically unfluctuating equilibrium point, and the storage function V turns into a Lyapunov function. The functionality of stability analysis using dissipativity is that this property is preserved under interconnection [114], [115].

2.2 An Equivalent Polynomial Representation

A polynomial expression able to mimic the behavior of a switched system is developed using a new variable s, which works as a parameter. The starting point is to rewrite (1) as a continuous non-switched control system in its more general case. The approach followed here has had in spirit some counterpart for 0-1 programs (see for instance [47]).

First, we define a drift vector field $\mathbf{F}(x) : \mathbb{R}^n \mapsto \mathbb{R}^n$

$$\mathbf{F}(x) = [f_0(x)\ f_1(x)\ ...\ f_q(x)] \tag{2}$$

where $f_i(x, u)$, $i \in \mathcal{Q}$, is the function for each subsystem of the switched systems given in (1). Let $\mathbf{L}$ be the vector of Lagrange polynomial interpolation quotients [25] defined with the new variable s, i.e.,

$$\mathbf{L}(s) = [l_0(s)\ l_1(s)\ ...\ l_q(s)]^T \tag{3}$$

where

$$l_k(s) = \prod_{\substack{i=0 \\ i \neq k}}^{q} \frac{(s-i)}{(k-i)} \tag{4}$$

We define the set

$$\Gamma = \{s \in \mathbb{R}\,|Q(s) = 0\}$$

where $Q(s)$ is the constraint polynomial so that

$$Q(s) = \prod_{k=0}^{q}(s-k) \tag{5}$$

which is used to constrain s to take only integer values of the original set $\mathcal{Q}$. Notice that this clearly implies that we cannot find a solution if the starting point does not belong to this set. Finally, the solution of this system may be interpreted as an explicit ordinary differential equation (ODE) on the manifold Γ. A related continuous polynomial system of the switched system (1) is constructed in the following theorem.

Theorem 4 *Consider a switched system of the form given in (1) with a drift vector field as is given in (2). Then, there exists a unique polynomial state system with a polynomial state equation $p(x, s)$ of degree q in s, with $s \in \Gamma$ as follows*

$$\dot{x} = p(x, s) = \boldsymbol{F}(x)\boldsymbol{L}(s) = \sum_{k=0}^{q} f_k(x) l_k(s) \tag{6}$$

This polynomial system is an equivalent polynomial representation of the switched system (1).

Proof. Given a set of $q+1$ subsystems $f_0(x), f_1(x), ..., f_q(x)$, using the definition of the interpolation polynomial in the Lagrange form [25], we obtain a linear combination of the Lagrange basis polynomials as follows,

$$p(x, j) = f_j(x), \ \ j = 0, 1, ..., q$$

We use the Lagrange quotients that have the properties that $l_k(s)$ is a polynomial (with degree $q+1$), and

$$l_k(s) = \delta_{ks} \equiv \begin{cases} 1, & s = k \\ 0, & s \neq k \end{cases}$$

where δ_{ks} is the Dirac Delta function supported in ks. With this property, the function $p(x, s)$ can be defined as a polynomial in s with degree at most q, and

$$p(x, j) = \sum_{k=0}^{q} f_k(x) l_k(j) = f_j(x).$$

There can be only one solution to the interpolation problem since the difference of two solutions is a polynomial with degree at most q, and $q+1$ zeros. This is only possible if the difference is identically zero, so $p(x, s)$ is the unique polynomial interpolating the given set of subsystems. From the numerator in Equation (4), we see that $l_k(s)$ is a polynomial of order q having zeros in all subsystems except the k-th ones. The denominator is simply the constant that normalizes its value to 1 at k.

Let $q+1$ be the finite number of subsystems of the switched system (1), i.e., $f_0(x), \ldots, f_q(x)$. Then, the polynomial state equation $p(x, s)$ is unique because the quotients of the Lagrange polynomial interpolation $l_0, ..., l_q$ are unique. Moreover, the solutions of the algebraic equations $Q(s)$ constrain the values of the variable s to be in the set of finite values of the original set $\mathcal{Q}$. Therefore, for any values of the $s \in \Gamma$, the polynomial $p(x, s)$ is equivalent to the switched system (1). ■

For instance, the most simple case arises when $q = 1$. In this case, the system (6) has the same form of the convex combination of two subsystems. When $q = 2$, the polynomial equivalent representation has the form

$$\dot{x} = p(x, s) = \sum_{k=0}^{2} f_k(x) l_k(s)$$
$$= \frac{1}{2} f_0(x)(s-1)(s-2) + f_1(x)(s)(2-s) + \frac{1}{2} f_2(x)(s)(s-1)$$

Notice that the trajectories of the original switched system (1) correspond to piecewise constant controls taking values in the set $\sigma \in \{0, 1, ..., q\}$.

2.3 Results in Stability Analysis for Polynomial Constrained Dynamical Systems

In the previous section, the switched systems are expressed as polynomial differential-algebraic systems or constrained control systems. With this reformulation, we can apply the approach presented recently for constrained polynomial control systems based on dissipation inequalities [32], which is in spirit similar to the approach presented in [81] (however the latter does not consider dissipation inequalities in its analysis). We can show that with some assumptions, both approaches are equivalent from a computational point of view.

The main idea behind the proposed approach is to include in the system analysis the set of constraints, which are represented in this case by the semi-algebraic set Γ. Note that the semi-algebraic set Γ is equivalent to take s as a constrained parameter that takes values on the roots of the polynomial $Q(s)$. We need to check negative semidefinitness of $\dot{V}(x)$ with respect to the constrained set Γ. We use the idea of penalization used in optimization theory with constraints. For that, we use a function

$\lambda(x,s)$, which can be interpreted as a penalization function or a Lagrange multiplier. This idea is based on some results presented in [32] for constrained control systems, where we can use the dissipation inequality concept using storage functions and supply rates [107]. Therefore, a dissipation inequality has the form $\dot{V}(x) \leq a(x, \dot{x}, s)$, where $a(\cdot)$ is an arbitrary scalar-valued function [31]. In the classical point of view, $V(x)$ is considered as the stored energy in the control system, and $a(\cdot)$ as the energy rate supplied into the control system [107]. Note that the stability of general differential-algebraic systems has only been recently presented as a dissipation inequality [31]. In this approach, we take this idea of constrained stability analysis to deal with singular constrained control systems [32].

The following stability theorem is a particular case of the general result presented in [32], and it is used to find a common Lyapunov function for the switched system (1) through the equivalent polynomial representation (6).

Theorem 5 *The equilibrium point $x^* = 0$ of the equivalent polynomial representation (6) of the switched system (1) is stable for any admissible input $s(t)$, if there exist polynomial functions $V : \mathbb{R}^n \mapsto \mathbb{R}$, $\lambda : \mathbb{R}^n \times \Gamma \mapsto \mathbb{R}$, and a constraint polynomial $Q(s) = 0$ such that $V(x)$ is positive definite in a neighborhood of the origin, and $\lambda(x,s) \geq 0$ in $\mathbb{R}^n \times \Gamma$ and the dissipation inequality*

$$\frac{\partial V}{\partial x} p(x,s) \leq Q^2(s)\lambda(x,s) \tag{7}$$

is satisfied for some neighborhood of the origin.

Proof. If the dissipation inequality (7) is satisfied, then the inequality can be integrated in the interval $[0,T)$

$$V(0) - V(T) \geq -\int_0^T \left(Q^2(s)\lambda(x,s)\right) dt$$

$$V(0) - V(T) \geq 0$$

we have used the fact that $Q(s) = 0$, and $s \in \Gamma$. This implies that $(\partial V/\partial x)p(x,s) \leq 0$, for all $t \geq 0$, in some neighborhood of the origin. We consider also that $Q^2(s)$ is positive or zero for all the values of s, in order to check semi-definiteness of $\dot{V}(x)$, and hence to satisfy the inequality above, we make $\lambda(x,s)$ positive. It follows from this Lyapunov inequality and the continuity of the trajectories $x(t)$, that V is not increasing and therefore the equilibrium point $x = 0$ of the system (6) is stable. Due to the equivalence presented in Theorem 1, the equilibrium point $x^* = 0$ is also an equilibrium point for the switched system (1) for any admissible $s \in \Gamma$. ■

Notice that the Lyapunov function $V(x)$ used in Theorem 2 only depends on the state, i.e., it is a common Lyapunov function under arbitrary switching [54].

Remark 6 *If we are interested in establishing asymptotic stability instead of stability, then (7) must be satisfied strictly for all nonzero x in some x-neighborhood, i.e.,*

$$\frac{\partial V}{\partial x}p(x,s) < Q^2(s)\lambda(x,s)$$

In general, it is very difficult to search for a Lyapunov function $V(x)$ and a function $\lambda(x,s)$ for practical problems. However, recently established methods based on semidefinite programming and sum of squares decomposition allow us to verify Lyapunov inequalities of the form (7) very efficiently in the case where $Q(s)$, $V(x)$, and $\lambda(x,s)$ are assumed to be polynomials [31]. Certainly, in our case all of these functions are of polynomial nature. It is impossible to search over all functions $V(x)$, $\lambda(x,s)$. In this approach, it is assumed that $V(x)$ and $\lambda(x,s)$ are polynomials up to certain degrees. Now, we can define the dissipation inequalities for the polynomial representation of the switched system. Since we are studying global uniform asymptotically stable (GUAS) systems, it means that we are searching for a common Lyapunov function regardless of the switching sequence. Therefore, if we try to prove global stability of the system (6), the following polynomial inequalities must be satisfied,

$$V(x) > 0, \text{ and}$$

$$\frac{\partial V}{\partial x} p(x,s) \le Q^2(s)\lambda(x,s)$$

for all $x \in \mathbb{R}^n$ and $s \in \mathbb{R}$. Note that if $V(x)$ is polynomial and positive definite, it implies that it is radially unbounded. To verify such polynomial inequalities is an NP-hard computational problem [31]. However, with the help of the sum of squares decomposition, it is possible to verify such polynomial inequalities very efficiently. On the other hand, this problem coincides with the problem of searching for a common Lyapunov function for the vector field $\mathbf{F}(x) = [f_0(x)\ f_1(x)\ ...\ f_q(x)]$.

For illustration and clarity of exposition, consider the case when $q = 1$. The dissipation inequality (7) becomes

$$\frac{\partial V}{\partial x}\left(f_0(x)(1-s) + f_1(x)s\right) \le \left(s(s-1)\right)^2 \lambda(x,s)$$

Before we state further results, we need to introduce some basic concepts of sum of squares decomposition. A more detailed description can be found in [77], and references therein.

2.3.1 The Sum of Squares Decomposition

In the following, we present some basic concepts of the sum of squares decomposition technique to be used for the system analysis. The sum of squares decomposition is a method to check if a polynomial can be decomposed into a sum of squared polynomials.

Definition 7 *[77] For $x \in \mathbb{R}^n$ a multivariate polynomial $p(x)$ is sum of squares (SOS) if there exist some polynomials $r_i(x)$, $i = 1, ..., M$, such that*

$$p(x) = \sum_{i=1}^{M} r_i^2(x) \tag{8}$$

It is clear that $p(x)$ being SOS naturally implies $p(x) \ge 0$, for all $x \in \mathbb{R}^n$. An equivalent characterization of SOS polynomials is given in the following proposition taken from [77].

Proposition 8 *[77] A polynomial $p(x)$ of degree 2d is an SOS if and only if there exists a positive semidefinite matrix Q and a vector of monomials $Z(x)$ containing monomials in x of degree $\leq d$ such that*

$$p(x) = Z(x)^T Q Z(x)$$

Since we have that $p(x, s)$ is a polynomial vector field, and that we are searching for $V(x)$ that is also a polynomial in x, to solve the testing conditions inequality (7), we can restrict our attention to cases in which the conditions admit SOS decompositions. The only apparent difficulty is the restriction of $V(x)$ to be positive definite, not just positive semidefinite. To deal with this problem we can use the following proposition taken from [77].

Proposition 9 *[77] Given a polynomial $V(x)$ of degree $2d$,*
let $\varphi(x) = \sum_{i=1}^{n} \sum_{j=1}^{d} \epsilon_{i,j} x_i^{2j}$ such that,

$$\sum_{j=1}^{d} \epsilon_{i,j} > \gamma \quad \forall i = 1, ..., n, \tag{9}$$

with γ a positive number, and $\epsilon_{i,j} \geq 0$ for all i and j. Then the condition

$$V(x) - \varphi(x) \quad \textit{is a SOS} \tag{10}$$

guarantees the positive definiteness of $V(x)$.

Using these ideas, we can rewrite inequality (7), and a relaxation of Theorem 2 is stated in the following proposition.

Proposition 10 *For the equivalent polynomial representation system (6) if there exist polynomial functions $V(x)$, $\lambda(x, s)$, and a positive definite function $\varphi(x)$ of the form given in Proposition 6 such that*

$$\begin{aligned} V(x) - \varphi(x) \ &\textit{is SOS} \\ -\tfrac{\partial V}{\partial x} p(x, s) + Q^2(s)\lambda(x, s) \ &\textit{is SOS} \end{aligned} \tag{11}$$

then the polynomials $V(x)$, $\lambda(x,s)$, and the positive definite function $\varphi(x)$ can be computed using SOSTOOLS [76].

The proof follows the same reasoning as in [77]. Therefore, Proposition 7 shows that with the polynomial equivalent representation in (5), we can obtain a common Lyapunov function using numerical tools. This Lyapunov function will be used to prove stability of the switched system (1).

Remark 11 *Note that in Equation (11) the polynomials are sum of squares in term of x and s.*

2.3.2 Numerical Example of a Polynomial Switched System

We present an illustrative example of a switched nonlinear system reformulated by Theorem 1 as an ordinary differential equation on a manifold. With this example, we illustrate an efficient computational treatment to study stability analysis of switched systems using Theorem 2. Consider the set of systems described by the drift vector field

$$\mathbf{F}(x) = [f_0(x)\ f_1(x)]$$

with

$$f_0(x) = \begin{bmatrix} -\beta_0 x_1 + x_1^2 + x_2^2 - \alpha_0 x_1^3 \\ -\beta 0 x_2 + 2x_1 x_2 - \alpha_0 x_2^3 \end{bmatrix}$$

and

$$f_1(x) = \begin{bmatrix} -\beta_1 x_1 + x_1^2 + x_2^2 - \alpha_1 x_1^3 \\ -\beta 1 x_2 + 2x_1 x_2 - \alpha_1 x_2^3 \end{bmatrix}$$

This system is considered as a homogeneous switched system presented for stability analysis in [53]. In order to prove stability under arbitrary switching, we use the polynomial equivalent representation obtained using Theorem 1

$$\dot{x}(t) = \begin{pmatrix} f_0(x) + (\bar{\beta}x_1 + \bar{\alpha}x_1^3)s \\ f_0(x) + (\bar{\beta}x_2 + \bar{\alpha}x_2^3)s \end{pmatrix}$$
$$s \in \Gamma = \{s \in \mathbb{R} \,|Q(s) = s(s-1) = 0\}$$

with $\bar{\beta} = (\beta_1 - \beta 0)$ and $\bar{\alpha} = (\alpha_1 - \alpha_0)$. In [53] it is shown that $\beta_i > 2$ and $\alpha_i > 4$ for $i = 1, 2$ in order to obtain a set of stable subsystems. We then set $\beta_0 = 10$ and $\alpha_0 = 5$ and to keep $\bar{\beta} > 2$ and $\bar{\alpha} > 4$, we set $\beta_1 = 13$ and $\alpha_1 = 10$. We have obtained a representation of the original system with a polynomial form, so that we can use Proposition 10 to analyze stability. First, we search for a Lyapunov function of the polynomial form $V(x) = \sum\sum_{i,j} a_{i,j}x_ix_j$. We have tried a function of degree 2 and 4, with the latter corresponding to the function that we are looking for. With a degree of $2d = 4$, and $n = 2$, we use a function $\varphi(x) = \epsilon_{11}x_1^2 + \epsilon_{12}x_1^4 + \cdots + \epsilon_{22}x_2^4$, where the ϵ_{ij} are the unknowns to be found by the tool, with a $\gamma = 0.1$, which implies $\sum \epsilon_{ij} \geq \gamma$. For the penalty function, we have assumed a polynomial function of the same degree of the Lyapunov candidate function $V(x)$, but considering also the s variable, i.e., $\lambda(x, s) = b_{11}x_1^2 + b_{12}x_1^4 + b_{21}x_2^2 + \cdots + b_{31}s^2 + b_{32}s^4$. The coefficients b_{ij}'s are again the unknown variables to be found. Using these polynomials and Equation (11), we obtain, using the MATLAB toolbox SOSTOOLS [76], a Lyapunov function of fourth degree, i.e.,

$$V(x) = 0.3x_1^2 + 0.3944x_2^2 + 0.11 \cdot 10^{-3}x_1^4 + 0.11 \cdot 10^{-3}x_2^4 + 0.7 \cdot 10^{-3}x_1x_2$$

which, through Theorem 2, proves that $(0, 0)$ of the homogeneous nonlinear switched system reformulated as a polynomial DAE (6) is a stable equilibrium point. Note that there is not a specific procedure to set the value of the degree of $V(x)$ and the minimum value for γ (it should be noticed also that these functions are not unique). These parameters are chosen by different attempts, we start trying with a degree $d = 1$ and a small value of γ, and then we increase the degree until properties of the Proposition 7 are met, and hence, we obtain a Lyapunov function. It may be interesting to develop an automatic pre-treatment algorithm to choose these values.

2.4 A Generalization for Nonlinear Switched Systems

In the previous sections, we have focused our attention on switched systems of polynomial form, i.e., each subsystem is modeled by a polynomial system. In this section, we extend the results to a more general class of switched systems, those modeled by elementary and nested elementary functions. This class of functions is related with explicit symbolic derivatives, such as exponential, logarithm, power-law, trigonometric, and hyperbolic functions. For this aim, we transform, using a recasting process, the system obtained by the equivalent representation in a system with polynomial form, and then, we use the results of Section 4 for stability analysis.

2.4.1 The Recasting Process For Stability Analysis

We use a recasting process introduced in [92], and later used for stability analysis of nonlinear systems in [75]. It is a procedure with several steps until the system has the expected form. The algorithm is as follows

- **Step 0. Equivalent Representation:** We consider the equivalent representation for the switched system obtained in Theorem 1, and we name it as the original system (i.e., before the recasting process), with $\xi = (\xi_1, \ldots, \xi_n)$ as the state of the original system.

- **Step 1. Original State Equations:** The original system is described by

$$\dot{\xi}_i = \sum_j a_j \prod_k p_{ijk}(\xi, s), \quad i = 1, \ldots, n \tag{12}$$

  here a_js are real numbers, and the factors p_{ijk} are elementary functions, or nested elementary functions of elementary functions.

- **Step 2. Decomposition of Non-Polynomial Functions:** Let $x_i = \xi_i$, for $i = 1, \ldots, n$. For each $p_{ijk}(\xi, s)$ in Equation (12) that is not already a power-law function, replace it with a new variable x_{n+1}. This variable simplifies the differential equation to sums and products of power-law functions. An additional differential equation is generated for each new variable, using the chain rule of differentiation.

- **Step 3. Recasting Process:** When the recasting process leads to some constraints in the new variables, we have to introduce an n-dimensional manifold on which the solutions to the original differential equation lie. The particular choice of initial conditions defines the reference manifold.

- **Step 4. The Polynomial Form:** If the set of equations is in polynomial form, then the recasting process is complete. If not, repeat steps 2-3 until to obtain a system of equations with a rational or polynomial form.

Remark 12 *Notice that the constraints introduced by the definition of new variables, and their initial conditions, restrict the system behavior to a manifold of the same dimension of the original problem.*

As a result of the recasting process we have obtained new variables, which are considered. Suppose that for a switched system consisting of subsystems of non-polynomial form, we apply the equivalent representation and obtain a system,

$$\dot{\xi} = p(\xi, s)$$

The recasted system obtained using the procedure presented above is written as

$$\begin{aligned} \dot{x}_o &= p_o(x_o, x_r, s) \\ \dot{x}_r &= p_r(x_o, x_r, s) \end{aligned} \tag{13}$$

where $x_o = (x_1, ..., x_n) = \xi$ are the state variables of the original system, $x_r = (x_{n+1}, ..., x_{n+m})$ are the new variables introduced in the recasting process, $p_o(x_o, x_r, s)$

and $p_r(x_o, x_r, s)$ have polynomial forms. Previously, in the recasting process we have also obtained new polynomial constraints. Consider the real-valued polynomial $g_k(x_o, x_r, s)$, with $k = 1, ..., m$, where m is the number of polynomial constraints generated in the recasting process. Let

$$\Gamma_r = \{(x_o, x_r, s) \in \mathbb{R}^{n+m+1} \mid g_k(x_o, x_r, s) = 0, \text{ for all } k = 1, ..., m\} \tag{14}$$

be the set of constraints from the recasting process. The following proposition is an extension of the stability Theorem 2, and can be used to prove that the origin of a nonlinear switched system is a stable equilibrium point.

Proposition 13 *The equilibrium point $x^* = 0$ of the equivalent polynomial representation obtained after the recasting process of the nonlinear switched system (1) is stable for any admissible input $s(t)$ if there exist polynomial functions $V : \mathbb{R}^{n+m} \mapsto \mathbb{R}$, $\lambda_o : \mathbb{R}^{n+m} \times \Gamma \mapsto \mathbb{R}$, $\lambda_r : \mathbb{R}^{n+m} \times \Gamma_r \mapsto \mathbb{R}$, and constraint polynomials $Q(s) = 0$, and $g_k(x_o, x_r, s) = 0$, with $k = 1, ..., m$, such that $V(x_o, x_r)$ is positive definite in a neighborhood of the origin, and $\lambda_o(x_o, x_r, s) \geq 0$ in $\mathbb{R}^{n+m} \times \Gamma$, $\lambda_k(x_o, x_r, s) \geq 0$ in $\mathbb{R}^{n+m} \times \Gamma_r$, $k = 1, ..., m$ and the dissipation inequality*

$$\begin{aligned} &\tfrac{\partial V}{\partial x_o} p_o(x_o, x_r, s) + \tfrac{\partial V}{\partial x_r} p_r(x_o, x_r, s) \leq Q^2(s)\lambda_o(x_o, x_r, s) \\ &+ \textstyle\sum_{k=1}^{m} g_k(x_o, x_r, s)\lambda_k(x_o, x_r, s) \end{aligned} \tag{15}$$

is satisfied for some neighborhood of the origin.

The above proposition establishes non-negativity conditions, which can be relaxed to appropriate sum of squares conditions (see Proposition 7), so that we can use the methods based on semidefinite programming and sum of squares decomposition to verify Lyapunov inequalities efficiently. We extend the Proposition 7 as follows.

Proposition 14 *For the equivalent polynomial representation system (13) if there exist polynomial functions $V(x_o, x_r)$, $\lambda_o(x_o, x_r, s)$, $\lambda_k(x_o, x_r, s)$, for $k = 1, ..., m$, and* *a positive definite function $\varphi(x_o, x_r)$ of the form given in Proposition 5 such that*

$$\begin{aligned} & V(x_o, x_r) - \varphi(x_o, x_r) \text{ is SOS} \\ & -\left(\tfrac{\partial V}{\partial x_o} p_o(x_o, x_r, s) + \tfrac{\partial V}{\partial x_r} p_r(x_o, x_r, s)\right) + Q^2(s)\lambda_o(x_o, x_r, s) \\ & + \textstyle\sum_{k=1}^{m} g_k(x_o, x_r, s)\lambda_k(x_o, x_r, s) \text{ is SOS} \end{aligned} \tag{16}$$

then the polynomials $V(x_o, x_r)$, $\lambda_o(x_o, x_r, s)$, $\lambda_k(x_o, x_r, s)$, for $k = 1, ..., m$, and the positive definite function $\varphi(x_o, x_r)$ can be computed using SOSTOOLS [76].

2.4.2 Example of a Non-Polynomial Switched System

In this example, we are dealing with a two-dimensional model of a pendulum, where the acceleration of its pivot is assumed to be the control input. Swinging up and stabilization of the pendulum is usually solved by switching between different laws. We use a damping-pumping strategy as it is proposed in [10]. The normalized model of the pendulum is given by

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= \sin x_1 - (2\sin x_1 + F x_2 \cos x_1)\cos x_1 \end{aligned} \tag{17}$$

where x_1 is the angular position with respect to the origin at the upright position, and x_2 is the velocity. Considering stabilization conditions, it is shown that we can set F as a gain of -1 in some regions, and a gain of 1 in some others, so that the system minimizes the energy consumption all the time. Therefore, we can obtain a switched system depending of the gain F. For $F = -1$ we set $f_0(x)$ and for $F = 1$ we set $f_1(x)$.

We then use Theorem 1 to obtain an equivalent continuous representation of the switched model related to (17) and we begin the recasting process, $\dot{\xi}(t) = f_0(\xi)(1 - s) + f_1(\xi)s$ with $s \in \Omega_1 = \{s \in \mathbb{R}\,|Q(s) = s(s-1) = 0\}$. We obtain the following equivalent continuous system,

$$\begin{aligned} \dot{\xi}_1 &= \xi_2 \\ \dot{\xi}_2 &= \sin\xi_1 - 2\sin\xi_1\cos\xi_1 + (1-2s)\xi_2\cos^2\xi_1 \end{aligned} \tag{18}$$

Now, following the recasting process, it is clear that (18) is in the same form as (12), but in this case the elementary functions are trigonometric functions. Let us follow Step 2 to Step 4 in the recasting process. As a result, we obtain a new set of differential equations given by

$$\begin{aligned}
\dot{x}_1 &= x_2 \\
\dot{x}_2 &= x_3 - 2x_3x_4 + (1-2s)x_2x_4^2 \\
\dot{x}_3 &= x_2x_4 \\
\dot{x}_4 &= -x_2x_3
\end{aligned} \tag{19}$$

As we know by Step 3 in the recasting process, when we introduce the new variables x_3 and x_4, a set of constraints arise. For this case, we have that the manifold on which the solutions to the original system (18) lie is given by $\Omega_2 = \{x_3^2 + x_4^2 - 1 = 0\}$. The resulting system is in a polynomial form so that Proposition 9 can be used to prove stability. Due to the form of the original system, we expect that the Lyapunov function has some trigonometric terms. We are searching for a Lyapunov function of the form $V(x) = a_1x_1^2 + a_2x_2^2 + a_3x_3^2 + a_4x_4^2 + a_5$. This coefficients must satisfy $a_4 + a_5 = 0$ for $V(x)$ to be equal to zero at $(0,0)$. Which is equivalent to $V(\xi) = a_1\xi_1^2 + a_2\xi_2^2 + a_3\sin^2(\xi_1) + a_4\sin^2(\xi_1 - 1)$ in the original variables. The problem now is to search for those $a_i's$ coefficients. In order to guarantee that $V(x)$ is positive definite, the polynomial function $\varphi(x_o, x_r)$ is chosen as,

$$\varphi(x_o, x_r) = \epsilon_1 x1^2 + \epsilon_2 x_2^2 + \epsilon_3 x_3^2 + \epsilon_4(1 - x_3)$$

where $\epsilon' s$ are positive constants. In this particular case, we set all $\epsilon_i \geq 1$ for $i = 1, ..., 4$. In the same way, we define the functions λ_o and λ_1 to be monomials of degree two. Using SOSTOOLS, we find the following common Lyapunov function

$$V = 0.892 + 0.82x_2^2 - 2.596x_4 + 2.596x_3^2 + 1.704x_4^2$$

which in the original variables is

$$V = 0.892 + 0.82\xi_2^2 - 2.596\cos(\xi_1) + 1.704\sin^2(\xi_1) + 1.704\cos^2(\xi_1)$$

Therefore, using Proposition 11, we can show that the origin is asymptotically stable.

CHAPTER III

ON OPTIMAL CONTROL OF SWITCHED SYSTEMS USING A POLYNOMIAL APPROACH

We propose an alternative approach for solving effectively the optimal control problem for an autonomous nonlinear switched system. We are considering a set of several continuous-time subsystems with a discrete switching law. The switching law consists of the switching times and mode transitions. The essential of this method is the transformation of a nonlinear, non-convex optimal control problem into an equivalent optimal control problem with linear and convex structure, a formulation more appropriate to be solved by high performance numerical computing. Therefore, using the Generalized Maximum Principle (GMP), we propose to convexify the state and the control variables by using the method of moments in the polynomial expression to deal with this problem.

At a first stage, we focus our analysis on vector fields and running costs that are of polynomial form. However, it is well known that functions called nested elementary functions, can be recasted exactly in a polynomial systems with a larger state dimension [92]. Therefore, we use the fact that all system data are polynomial after the recasting process, to apply the theory of moments as it has been mentioned above.

3.1 Definitions and Preliminaries

3.1.1 Switched systems and its Optimal Control Problem

The switched system adopted in this chapter has a general mathematical model described by

$$\dot{x}(t) = f_{\sigma(t)}(t, x(t), u(t)), \tag{20}$$

where $x(t)$ is the state, $f_i : \mathbb{R}^+ \times \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}^n$ are vector fields,

$$x(t_0) = x_0,$$

are fixed initial values,

$$u(t) \in \mathcal{U} \subset \mathbb{R}^m$$

is the exogenous input constrained to the convex and compact set $\mathcal{U}$, and

$$\sigma : [t_0, t_f] \to \mathcal{Q} \in \{0, 1, 2, ..., q\}$$

is a *piecewise constant* function of time, with t_0 and t_f as the initial and final times respectively. Every mode of operation corresponds to a specific subsystem $\dot{x}(t) = f_i(t, x(t), u(t))$, for some $i \in \mathcal{Q}$, and the *switching signal* σ determines which subsystem is followed at each point of time, into the interval $[t_0, t_f]$. The control inputs, σ and u, are both measurable functions. In addition, we consider non Zeno behavior, i.e., we exclude an infinite switching accumulation points in time. Finally, we assume that the state does not have jump discontinuities. Moreover, for the interval $[t_0, t_f]$, the control functions must be chosen such that the initial and final conditions are satisfied.

Definition 15 *A control for the switched system in (20) is a triplet consisting of*

(a) a finite sequence of modes,

(b) a finite sequence of switching times such that $t_0 < t_1 < \cdots < t_q = t_f$*,*

(c) a sequence of exogenous control inputs, each control input function being associated with a mode.

Let us define the optimization functional in bolza form to be minimized as,

$$J = \varphi(x(t_f)) + \int_{t_0}^{t_f} L_{\sigma(t)}(t, x(t), u(t))dt, \tag{21}$$

where $\varphi(x(t_f))$ is a real-valued function, and the running switched costs $L_{\sigma(t)} : \mathbb{R}^+ \times \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}$ are continuously differentiable, for each $\sigma \in \mathcal{Q}$.

Switched Optimal Control Problem (SOCP)

We want to solve the switched optimal control problem, which can be state in a general form as follows,

Definition 16 *Given the switched system in (20) and a Bolza cost functional J as in (21), the switched optimal control problem (SOCP) is given by*

$$\min_{\sigma(t),u(t)} J(t_0, t_f, x(t_0), x(t_f), x(t), \sigma(t), u(t)) \tag{22}$$

subject to the states $x(\cdot)$ satisfies Equation (20).

The SOCP can have the usual variations of fixed or free initial or terminal state, free terminal time, etc. In [37] it is noted that in this setting, it is not appropriate to choose first a sequence of controls and then determine the trajectory associated to it, because a priori the sequence could not be admissible (in the sense that there could exist no trajectory corresponding to it). This is due to the fact that in every mode $i \in \mathcal{Q}$, it is possible to use only a subset of $\mathcal{U}$, depending on the switching strategy.

3.1.2 Maximum Principle and Necessary Conditions

The *Maximum Principle* gives a necessary condition for a trajectory $x(\cdot)$ to be a solution of the switched optimal control problem (SOCP). The set of variations involves trajectories having the same history of the candidate's optimal one, which is having the same switching strategy (see [101], [37]). Variations of the classical Maximum Principle for general hybrid systems has been presented previously in ([23], [101], [111], [87], [37], [19], [95]). In particular in [101], the principle is presented as an abstract mathematical statement that can be rendered specific in various ways, giving rise to different versions of the principle. We will state in the next section a specific version concerning to our approach. However, we need to introduce first a brief summary of the maximum principle for continuous systems in its basic form.

The Maximum Principle

Consider a dynamical system defined by a differential equation on time the interval $[t_0, t_f]$, i.e.,

$$\dot{x}(t) = f(t, x(t), u(t)), \tag{23}$$

$$x(t_0) = x_0$$

where $x(t) \in \mathbb{R}^n$, the control input $u(t) \in U \subset \mathbb{R}^m$, and the vector field $f : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}^n$. Then, the problem can be stated as follows: choose the initial conditions x_0 and the control input $u(t)$ so that the functional

$$J = \varphi(t_f, x(t_f)) + \int_{t_0}^{t_f} L(t, x(t), u(t))dt \tag{24}$$

is minimized. In order to solve the problem, the following assumptions are made:

(A_1) The control domain U is bounded.

(A_2) The vector fields $f(t, x, u)$ and $L(t, x, u)$ are continuous and continuously differentiable functions with respect to the state variables and the time variable.

(A_3) φ is continuous and continuously differentiable.

Define the Hamiltonian function as

$$H(t, \lambda(t), x(t), u(t)) = L((t, x(t), u(t)) + \lambda^T f(t, x(t), u(t)) \tag{25}$$

and the Hamiltonian system as

$$\dot{x} = \frac{\partial H}{\partial \lambda}(t, \lambda(t), x(t), u(t))$$

$$\dot{\lambda} = \frac{\partial H}{\partial x}(t, \lambda(t), x(t), u(t)), \quad \lambda(t_f) = -\partial_x \phi(t_f, x(t_f)) \tag{26}$$

Theorem 17 *(**Necessary Conditions**) [79]If $(x^*(t), u^*(t))$ is the pair of the corresponding trajectory for the control problem and an admissible optimal control with the assumptions $(A_1) - (A_3)$, then there must exist a function $\lambda(t)$ on $[t_0, t_f]$, such* *that $\lambda^*(x^*(t), u^*(t))$ satisfies the Hamiltonian system (25)-(26) almost everywhere as well as the maximum condition,*

$$H(t, \lambda^*(t), x^*(t), u^*(t)) = \sup_{u \in U} H(t, \lambda^*(t), x^*(t), u(t))$$

The question now is whether necessary conditions of optimality are sufficient. In general, the maximum principle is not a sufficient condition for global optimality of (x^*, u^*). Nevertheless, the maximum principle becomes a sufficient condition for optimality of (x^*, u^*) under some additional assumptions about the control system (23) and the functional (24).

Theorem 18 ***(Sufficiency of Optimality Conditions)*** *[79] Assume that $L(t, x, u)$ and the subset U are convex in (x, u), and $f(t, x, u)$ is linear in (x, u). If the triplet (λ, x, u) satisfies the Hamiltonian system (25) and the maximum condition, then the pair (x^*, u^*) is a global optimal solution of the corresponding optimal control problem.*

Quite often, optimal solutions for control problems cannot be found, either because there are many (or not so many) variables involved, so that it is almost impossible to handle them all by hand, or else because optimality conditions cannot be solved explicitly or it is really cumbersome and tedious to find explicit formulas [79]. In this context, efficient computational tools appear as an important alternative to solve non-linear optimal control problems.

3.1.3 Relaxation and Young Measures

We present the main results concerned with the analysis of optimal control problems governed by ordinary differential equations by means of Young measures. We show a rather general existence theorem for generalized solutions of control problems in the form of Young measures and first order necessary conditions of optimality that this generalized solutions must verify. These conditions come in the form of a generalized maximum principle that, after all, impose constraints on the support of optimal Young measures. We describe the relaxation of optimal control problems, with some hypotheses on the different ingredients of the problem so that, regardless of convexity assumptions, existence of optimal solutions can be achieved. Nonexistence (under coercivity assumption) is always related to oscillatory behavior which in form is induced by a lack of convexity. A classical relaxation theorem establishes, under some technical assumptions, that the infimum of any functional does not change when we replace the integrand by its convexification. The relaxed version of the problem we are using is formulated in terms of Young measures associated with sequences of admissible controls [78].

Consider to minimize the functional defined previously in Equation (24) with initial time $t_0 = 0$, and final time fixed $t_f = T$. Where

(a) the control is assumed measurable and takes values in a given closed set U (not assumed bounded for now),

(b) the state of the system governed by the equation of state (23) is assumed to be measurable on the variable t, continuous in (x, u), continuously differentiable with respect to x, and satisfies a uniform Lipschitz condition with respect to x

$$|f(t, x_1, u) - f(t, x_2, u)| \leq \varepsilon |x_1 - x_2|, \quad \varepsilon > 0,$$

so that problem (23) always has a unique solution,

(c) the running cost L is assumed to be continuous on the pair (x, u), differentiable with respect to x, and measurable on t, it is also assumed to satisfy the coercivity requirement

$$c(|u|^p - 1) \leq L(t, x, u), \quad p > 1, \ c > 0, \tag{27}$$

(d) the function φ is assumed to be continuous and differentiable.

Let us further set

$$h(x,u) = \sup\{|f(t,x,u)| : 0 < t < T\}$$

and

$$L_U(t,x,u) = \begin{cases} L(t,x,u), & u \in U \\ +\infty, & else. \end{cases}$$

Notice that the function h is continuous with respect to x, even Lipschitz continuous. We postulate as well the behavior

$$\lim_{|u|\mapsto\infty} \frac{h(x,u)}{(|u|^p)} = 0 \tag{28}$$

for each $x \in \mathbb{R}^n$. Let $\mathcal{U}$ be the set of admissible controls

$$\mathcal{U} = \{u : u \in L^p(0,T), u(t) \in U\}.$$

The target space for functions in $L^p(0,T)$ is assumed throughout to be $\mathbb{R}^m$ and thus will not be indicated explicitly. The set of Young measures associated to sequences in $\mathcal{U}$ is [78]

$$\widetilde{\mathcal{U}} = \left\{\nu = \{\nu_t\}_{t\in(0,T)} : supp(\nu_t) \subset U,\ a.e.,\ t \in (0,T),\ \int_0^T \int_U |\eta|^p \, d\nu_t(\eta)dt < \infty\right\}.$$

where ν is a probability measure supported in U. The extended functional $\widetilde{J}$ defined on $\widetilde{\mathcal{U}}$ is given by

$$\widetilde{J}(\nu) = \varphi(x(T)) + \int_0^T \int_U L(t,x(t),\eta)d\nu_t(\eta)dt$$

where $x(t)$ is the solution of

$$\dot{x}(t) = \int_U f(t,x(t),\eta)d\nu_t(\eta), \quad x(0) = x_0,$$

This initial value problem is well-posed because this function satisfies the Lipschitz condition in x necessary to ensure a unique solution [73], [78].

Theorem 19 *If in addition to the coercivity condition (27) and the behavior indicated in (28) (no convexity in L_U is assumed), we have the upper bound*

$$L(t, x, u) \leq k(x)(1 + |u|^p)$$

where $k \in L^\infty_{loc}(\mathbb{R}^n)$, then $\inf\{J(u) : u \in \mathcal{U}\} = \min\{\widetilde{J} : \nu \in \widetilde{\mathcal{U}}\}$.

Notice the use of min in the generalized problem to emphasize the existence of a solution. We state the generalized maximum principle as in [112], [78], [73]. Before stating the generalized maximum principle, we introduce the generalized Hamiltonian by

$$H(t, x, \mu, \lambda) = \langle H(t, x, \cdot, \lambda), \mu \rangle,$$

where

$$H(t, x, u, \lambda) = L(t, x, u) + \lambda f(t, x, u)$$

and μ is a probability measure supported in U such that

$$\int_U |\eta|^p d\mu(\eta) < \infty$$

Theorem 20 *If $\nu = \{\nu_t\}_{t \in (0,T)}$ is a minimizer for $\widetilde{J}$ in $\widetilde{\mathcal{U}}$, and the assumptions (a)-(d) above hold, then there exists a function $\lambda(t)$, $t \in (0, T)$ such that for a.e. $t \in (0, T)$,*

$$\dot{\lambda}(t) = -\int_U \frac{\partial H}{\partial x}(t, x, \eta, \lambda) d\nu_t(\eta), \quad \lambda(T) = \nabla \varphi(x(T)),$$

with the generalized Hamiltonian condition

$$\int_U H(t, x, \eta, \lambda) d\nu_t(\eta) = \inf\{H(t, x, \eta, \lambda) : \eta \in \widetilde{U}\},$$

where the initial value problem is

$$\dot{x}(t) = \int_U f(t, x, \eta) d\nu_t(\eta), \quad x(0) = x_0.$$

Notice that $\partial H/\partial x$ is a continuous function of u, because both L and f are continuous functions of (x, u) so that the integral that appears in $\dot{\lambda}$ above is well-defined. The conclusion of this theorem is a direct generalization of the classical Pontryagin principle, and its proof does not involve any particular difficulty. There are some other ways of obtaining this type of optimality conditions based on sliding mode variation of relaxed controls as it is mentioned in [73].

3.2 An Equivalent Polynomial Optimal Control Problem

3.2.1 Equivalent Representations

The starting point is to rewrite (20) as a continuous non-switched control system as it has been shown in Chapter 2 for the stability analysis. The polynomial expression in the control variable that is able to mimic the behavior of the switched system is developed using a variable s, which works as a control variable.

Let Lagrange polynomial interpolation quotients [25] be defined with the new variable s, i.e.,

$$l_k(s) = \prod_{\substack{i=0 \\ i \neq k}}^{q} \frac{(s-i)}{(k-i)} \tag{29}$$

We define the set of polynomial constraints which is used to constrain s to take only integer values of the original set $\mathcal{Q}$ as,

$$\Omega = \{s \in \mathbb{R} \,|Q(s) = 0\}$$

where $Q(s)$ is the constraint polynomial defined by

$$Q(s) = \prod_{k=0}^{q} (s-k). \tag{30}$$

The solution of this system may be interpreted as an explicit ODE on the manifold Ω. A related continuous polynomial system of the switched system (20) is constructed in the following theorem.

Theorem 21 *Consider a switched system of the form given in (20). Then, there exists a unique continuous state system with polynomial dependence in the control variable s, $p(x, s)$ of degree q in s, with $s \in \Omega$ as follows*

$$\dot{x} = p(x, s) = \sum_{k=0}^{q} f_k(x) l_k(s) \tag{31}$$

This polynomial system is an equivalent polynomial representation of the switched system (20).

Proof. The proof of this proposition is presented in Theorem 4 of Chapter 2. ■

Note that the trajectories of the original switched system (20), correspond to piecewise constant controls taking values in the set $\sigma \in \{0, 1, ..., q\}$.

Similarly, we define a polynomial equivalent representation for the running cost $L_{\sigma(t)}$ using the Lagrange's quotients as follows,

Proposition 22 *Consider a switched running cost of the form given in (21). There exists a unique polynomial running cost equation $\mathcal{L}(x, s)$ of degree q in s, with $s \in \Omega$ as follows*

$$\mathcal{L}(x, s) = \sum_{k=0}^{q} L_k(x) l_k(s) \tag{32}$$

with $l_k(s)$ defined in (29). This polynomial system is an equivalent polynomial representation of the switched running cost in (21).

Proof. The proof of this proposition uses the same ideas as Theorem 4 of Chapter 2. ■

3.2.2 Equivalent Optimal Control Problem

Based on the reformulation presented in the previous section, we define an optimal control problem based on these equivalent polynomial representations. Consider the equivalent optimal control problem (EOCP),

The functional using Equation (32) is defined by

$$J = \varphi(x(T)) + \int_0^T \mathcal{L}(x, s) dt$$

subject to the system

$$\dot{x}(t) = p(x,s) = \left(\sum_{k=0}^{q} f_k(x) l_k(s) \right)$$

$$\text{with} \quad x \in \mathbb{R}^n, \quad s \in \Omega, \quad x(0) = x_0$$

where $l_k(s)$, Ω, and $\mathcal{L}$ are defined as above. Note that this control problem is a continuous polynomial system with the input constrained by a polynomial $Q(s)$. Notice also that we are using s as a control variable. In order to develop a methodology based on Young measures and the theory of moments we have the following assumption, which will be omitted in the next section for a generalization of this approach.

Assumption 23 *All functions in the drift vector field* $\boldsymbol{F}(x) = [f_0(x) f_1(x) \cdots f_q(x)]$, *and running cost functions* $L_0(x), ..., L_q(x)$, *are polynomials.*

we rewrite the EOCP using this polynomial dependence. Let $\mathbb{R}[x,s] = [x_1, ..., x_n, s]$ denote the ring of polynomials in the variables x and s, with bases ordered lexicographically as it is shown in the Appendix 1,

$$1, x_1, x_2, \ldots, x_n, x_1^2, x_1x_2, \ldots, x_1x_n, x_2^2, \cdots, x_2x_3, \ldots, x_n^2, \ldots, x_1^r, \ldots, x_n^r, s, s^2, \cdots, s^q$$

for the vector space of real-valued polynomials of degree at most r in x, and the scalar polynomial variable s of degree at most q. Then, an r-degree polynomial running cost $\mathcal{L}(x,s) : \mathbb{R}^{n+1} \mapsto \mathbb{R}$ is written as

$$\mathcal{L}(x,s) = \sum_{\eta,k} a_{\eta k} x^{\eta} s^k$$

where η is the biggest degree of the polynomial $\mathcal{L}$ in the x variable, and a polynomial map $p(x,s) : \mathbb{R}^{n+1} \mapsto \mathbb{R}^n$

$$p(x,s) = \sum_{\gamma,k} p_{\gamma k} x^{\gamma} s^k$$

where γ is the biggest degree of the polynomial p in the x variable, and some coefficients $a_{k\eta}$ and $p_{k\gamma}$, that depend on $\mathcal{L}(x,s)$ and $p(x,s)$, respectively. We define

$\beta = max\{\eta, \gamma\}$ as the biggest degree of the polynomials in x, and use a canonical basis on that, putting zeros where is necessary. Then, we use a unified exponent and introduce the Hamiltonian as

$$H(x, \lambda, s) = \mathcal{L}(x, s) + \lambda^T p(x, s) = \sum_{k,\beta} a_{k\beta} x^\beta s^k + \sum_{k,\beta} \lambda^T p_{k\beta} x^\beta s^k \tag{33}$$

where λ is the vector of co-states. With this polynomial Hamiltonian, we can establish a polynomial Hamiltonian problem as follows.

The function of co-state $\lambda(t)$ can be expressed as polynomial in the form:

(a) The co-state equation

$$\dot{\lambda}(t) = -\frac{\partial H}{\partial x} = -\sum_{k,\beta} a_{k\beta} \left(\nabla_x x^\beta\right)^\top s^k - \sum_{k,\beta} p_{k\beta} \left(\nabla_x x^\beta\right)^\top \lambda s^k \tag{34}$$

with

$$\lambda(T) = \nabla_x \varphi(x(T))$$

(b) The minimum condition over s

$$\min_s H = \min_s \left(\sum_{k,\beta} a_{k\beta} x^\beta s^k + \sum_{k,\beta} \lambda^T p_{k\beta} x^\beta s^k \right) = \min_s \left(\sum_{k,\beta} \alpha_{k\beta}(\lambda) x^\beta s^k \right) \tag{35}$$

Notice that due to the linear relation of λ^T, we can obtain a function $\alpha_{\beta k}(\lambda)$, which is linear and depends on the constant coefficients $a_{\beta k}$ and $p_{\beta k}$.

(c) The state equation

$$\dot{x}(t) = \sum_{k,\gamma} p_{k\gamma} x^\gamma s^k \tag{36}$$

Thanks to this polynomial form of the Hamiltonian system, we can use some concepts introduced in [78] for relaxation of functions and use the generalized maximum principle presented in Theorem 20. This relaxation is mainly based on parametrized measures as it is shown in the next Section.

3.3 Relaxation of the Equivalent Optimal Polynomial Problem

In the previous section, we have introduced the basic concepts about relaxation and Young measures, which are used in this section to obtain, in regard of the special structure, a polynomial form of the optimal control problem, and hence, a relaxation of this equivalent optimal control problem as it is in Equations (34)-(36). Due to the polynomial dependence on x and s, we are concerned with moments of such probability measures. We obtain a convexification of the state x and the control s by using moment variables, which allows us to obtain an equivalent convex formulation more appropriate to be solved by high performance numerical computing. To solve non-convex polynomial programs as (35) subject to (34) and (36), we use the convex hull of the graph of the polynomial (35), once it has been provided the coercivity requirement, which is assumed if $\alpha_{\beta k} > 0$, with β and k be even.

Let Ω be the set of admissible controls $s(t)$ up to time T, $\Omega = \{s(t) \in \mathbb{R} | Q(s) = 0\}$ with $Q(s)$ as is defined in Equation (30). The set of Young measures associated to admissible state-control in $\mathcal{S}$ is

$$\widetilde{\Omega} = \left\{ \mu = \{\mu_t\}_{t \in (0,T)} : supp(\mu_t) \subset \Omega, a.e., t \in (0,T), \quad \int_0^T \int_{\mathcal{S}} |\eta|^p d\mu_t(\eta) dt < \infty \right\}$$

where μ is a probability measure supported in Ω. The extended functional $\widetilde{J}(x,s)$ defined on Ω is now given by

$$\widetilde{J}(x,s) = \varphi(x(T)) + \int_0^T \int_\Omega \mathcal{L}(t, x(t), \eta) d\mu_t(\eta) dt$$

where $x(t)$ is the solution of

$$\dot{x}(t) = \int_\Omega p(x,\eta) d\mu_t(\eta), \quad x(0) = x_0,$$

This initial value problem is well-posed because this function satisfies the Lipschitz condition in x necessary to ensure a unique solution [73], [78]. We can restate the generalized maximum principle presented in Theorem 20 as a generalized maximum principle as follows

Theorem 24 ***(Generalized Maximum Principle - GMP)***

If $\lambda = \{\lambda_t\}_{t\in(0,T)}$ is a minimizer for $\widetilde{J}$ in $\widetilde{\Omega}$, and the assumption ((a)-(d)) in Section 3.1.3 hold, then there exists a function $\lambda(t) \in \mathbb{R}^n$ such that for a.e. $t \in (0,T)$,

$$\dot{\lambda}(t) = -\int_{\Omega} \frac{\partial H}{\partial x}(t,x,\eta,\lambda)d\mu_t(\eta), \quad \lambda(T) = \nabla\varphi(x(T)) \tag{37}$$

the generalized Hamiltonian minimum condition is now written as

$$\int_{\Omega} H(t,x,\eta,\lambda)d\mu_t(\eta) = \inf\{H(t,x,\eta,\lambda) : \eta \in \Omega\} \tag{38}$$

and the initial value problem is

$$\dot{x}(t) = \int_{\Omega} p(x,\eta)d\mu_t(\eta), \quad x(0) = x_0 \tag{39}$$

with this generalized maximum principle, we have obtained an infinite dimensional linear program. Note that (35) is feasible whenever there exists an admissible control. This linear program is a rephrasing of the polynomial Hamiltonian system (34-36) in terms of the Young measures of its trajectories (x, s). In order to obtain a semidefinite program (SDP), which is a relaxation of the GMP, we use the fact that all functions of the EOCP are polynomials. Indeed, this is precisely the relaxation in moments of the global optimization of the polynomial Hamiltonian $H(x, \lambda, s)$, when the state and the variable s are transformed into a vector m. Thus, every minimizer of the convex formulation (37) attains the minimum value of the equivalent polynomial optimal control problem (34-36), therefore the minimizers attain the minimum value of the switched optimal control problem (21). Since this method is a relaxation, H always produces a lower bound for the optimal value H^*.

If we consider the global minimization of H as a global optimization problem, it is well known that the main objective is to find the global minima of a function H

defined on a subset $\mathcal{S}$. In other words, we are interested in solving a mathematical program given in the general form,

$$\min_{s\in\Omega} H(x,s)$$

where the objective function $H(x,s)$ is a linear combination of simpler functions.

One approach to tackle this problem comes from convex analysis, since we can use the convex envelope of the function H in order to locate its global minima. As we have shown, every convex combination of points in Ω can be described as a discrete probability distribution μ supported in Ω such that every integral

$$\int_\Omega H(x,s)d\mu(\eta)$$

represents one point over the convex envelope of the function H. For this reason, we study the relaxed problem

$$\min_{\mu\in P(\Omega)} \int_\Omega H(x,s)d\mu(\eta) \tag{40}$$

in order to find the global minima of the objective function H in Ω. The relaxed problem (40) contains information about all the global minima of the function H in $\mathcal{S}$. The solution of this problem is the family of all probability measures supported in $\arg\min(H)$ (see [46], [62]). However, it cannot be solved easily in practice (consider for instance, the difficulty of describing all possible convex combinations of points in $\mathcal{S}$). The linear program (40) is infinite dimensional, and thus, not tractable as it stands. Therefore, we present a relaxation scheme that provides a sequence of semidefinite programs, or linear matrix inequality relaxations, each with finitely many constraints and variables.

As we know, the function H is a polynomial, and hence, it can be expressed as a linear combination of simpler functions. In this case, the simpler functions are the algebraic system of integer exponents. With these considerations, we can deal with polynomial optimization problems using the method of moments.

The method of moments is a general method for treating non convex optimization problems. It takes a proper formulation in probability measures of a non convex optimization problem (in Appendix 1 we present the main ideas behind this method). Therefore, when the problem can be stated in terms of polynomial expressions, we can transform the measures into algebraic moments to obtain a new convex program defined in a new set of variables that represent the moments of every measure [61]. We can express the linear combination as,

$$H = \sum_{\beta,k} \alpha_{\beta k} \psi_{\beta k}, \quad \beta \in \mathbb{N}^r, \quad k \in \mathbb{N}^q$$

where the function basis $\{\psi_i\}$ is the algebraic system $\psi_{\beta,k} = x^\beta s^k$. Then, we are dealing with the algebraic system in the form

$$H(x, \lambda, s) = \sum_{\beta,k} \alpha_{\beta k}(\lambda) x^\beta s^k, \tag{41}$$

We show that it is possible to determine the global minima for this algebraic polynomials. For that, consider the following optimization problem

$$\min_{x,s \in \mathcal{S}} H(x, s, \lambda)$$

We apply the convexification of this function, obtaining the envelope of the function H

$$\min_{x,s \in \mathcal{S}} H(x, \lambda, s) \rightarrow \min_{\mu} \int_{\mathcal{S}} H(x, \lambda, s) d\mu(x, s) = \min_{\mu} I$$

$$\min_{\mu} I = \min_{\mu} \int_{\mathcal{S}} \left(\sum_{\beta,k} \alpha_{\beta k} x^\beta s^k \right) d\mu(x, s)$$

$$\min_{\mu} I = \min_{\mu} \sum_{\beta,k} \alpha_{\beta k} \underbrace{\int_{\mathcal{S}} x^\beta s^k d\mu(x, s)}_{\text{K-truncated moments}}$$

$$\min_{m} I = \min_{m_{\beta,k}} \sum_{\beta,k} \alpha_{\beta k} m_{\beta,k} \tag{42}$$

In a similar way, we obtain the convexification of the state

$$\dot{m}_{\beta 0}(t) = \int_{\mathcal{S}} \sum_{\beta,k} p_{\beta k} x^\beta s^k d\mu(x, s) = \sum_{\beta,k} p_{k\beta} m_{\beta,k}(t) \tag{43}$$

We have that $\mathcal{S} \subset \mathbb{R}^{n+1}$ is a semialgebraic set. However, we recall the fact that the vector variable m should be moments of a measure μ with support contained in $\mathcal{S}$. We invoke recent results of real algebraic geometry on the representation of positive polynomials on a compact set, and obtain necessary and sufficient conditions on the variables $m_{\beta k}$ to be, indeed, moments of a measure μ with appropriate support. Therefore, a sequence m has a representing measure μ supported on $\mathcal{S}$ only if these moments are restricted to be the entries on a positive semidefinite moment matrix $M_n(m)$ with $m_0 = 1$ and a localizing matrix defined as follows [28], [47], [46] (More details about moment and localizing matrices can be found in the Appendix 1 and references therein).

Definition 25 ***Moment matrix:*** *For a given real sequence $m = \{m_\gamma\}_{\gamma \in \mathbb{N}^n \times \mathbb{N}^q}$ of real numbers, the moment matrix $M_r(m)$ of order r associated with m, has its rows and columns indexed in the canonical basis $\{x^\beta, s^k\}$, and is defined by*

$$M_r(m)(\gamma, \alpha) = m_{\gamma+\alpha}, \quad \gamma, \alpha \in \mathbb{N}^n \times \mathbb{N}^q, \ |\gamma|, |\alpha| \leq r, \tag{44}$$

where $|\gamma| := \sum_j \gamma_j$.

$M_r(m)$ is symmetric nonnegative (denote $M_r(m) \succeq 0$, for every r). We define the localizing matrix $M_r(\theta m)$ whose positivity is directly related to the existence of a representing measure for m with support in $\mathbb{K} = \{(x, s) \in \mathbb{R}[x, s] : \theta(x, s) \geq 0\}$.

Definition 26 ***Localizing matrix:*** *For a given polynomial $\theta \in \mathbb{R}[x, s]$, written as*

$$\theta(x, s) = \sum_{\beta,k} \theta_{\beta,k} x^\beta s^k,$$

We define the localizing matrix $M_r(\theta m)$ associated with m, θ, and with rows and columns also indexed in the canonical basis of $\mathbb{R}[x, s]$, by

$$M_r(\theta m)(\gamma, \alpha) = \sum_{\beta,k} \theta_{\beta,k} m_{(\theta,k)+\gamma+\alpha}, \quad \gamma, \alpha \in \mathbb{N}^n \times \mathbb{N}^q, \quad |\gamma|, |\alpha| \leq r. \tag{45}$$

$M_r(\theta m)$ is also symmetric nonnegative, $M_r(\theta m) \succeq 0$ for every r. The $\mathbb{K}$-moment problem identifies those sequences m that are moments-sequences of a measure with support contained in the semialgebraic set $\mathcal{S}$.

The important property of all the above conditions is that when it is stated for all polynomials of degree less than r, they translate into linear matrix inequalities (LMI) conditions on m, via moment and localizing matrices associated with m and $\mathcal{S}$. It is shown that the global minima of the optimization problem are equivalent to the minima of the programming problem (42). It can be shown that the optimization problem given in (42) is a semidefinite program (SDP) because of the symmetry on the moments matrices. Consequently, we have now the necessary elements to state a linear convex program related with the original problem (22).

3.3.1 SDP Relaxation for the Optimal Control Problem

We present the semidefinite relaxation of the optimal control problem obtained from the polynomial EOCP using the theory of moments. Consider the polynomial Hamiltonian system defined by (35) and subject to (36) and (34), where we have the set $\mathcal{S} = \{\mathbb{R}^n \times \Omega\}$ of admissible control values. Recall that originally we have defined the set Ω as the constraint polynomial $\Omega = \{s \in \mathbb{R} \mid Q(s) = 0\}$ with $Q(s)$ defined as in Equation (30). Now, to be coherent with the definitions of localizing matrix and the representation results of the Appendix 1, we treat the polynomial $Q(s)$ as two opposite inequalities $Q_1(s) = Q(s) \geq 0$, and $Q_2(s) = -Q(s) \geq 0$, and we redefine the compact set Ω to be

$$\Omega = \{Q_i \geq 0, \;\; i = 1, 2\}$$

Define the space of moments as

$$\Gamma = \Big\{ m = \{m_{\beta k}\} : \; m_{\beta k} = \int_{\mathcal{S}} x^\beta s^k d\mu(x, s), \;\; \mu \in P(\mathcal{S}), \cdots$$

$$\cdots M_i(m) \succeq 0, \; M_{i-d_i}(Q_1 m) \succeq 0, \; M_{i-d_i}(Q_2 m) \succeq 0 \}$$

where μ is a probability measure supported in $P(\mathcal{S})$, M_i is a moment matrix associated to the vector of moments m, $M_{i-d_i}(Q_1 m)$, and $M_{i-d_i}(Q_2 m)$ are localizing matrices related to the vector of moments constrained to the set Ω, and $d_i = \lceil deg(Q_1)/2 \rceil$. We easily see from (30) that $deg(Q_1) = deg(Q_2) = q+1$, where $q+1$ is the number of modes of the switched system, so that $d_i = \lceil (q+1)/2 \rceil$. Since the mapping $\mu \in P(\mathcal{S}) \mapsto \Gamma$ is linear, we conclude that Γ is a convex set of vectors [80].

We can take advantage of the moment structure of the Hamiltonian and the state equation to rewrite the relaxed formulation obtained in Theorem 24 as a SDP. For $i \geq \max[deg(H), \max_i deg(Q_i)]$ consider the positive semidefinite programs (LP_i)

Semidefinite programs-LP_i:

LP_i: $\begin{cases} \text{Minimize the Hamiltonian defined in moments} \\ H_i^*(m) = \min_{m_{\beta k} \in \Gamma} \sum_{\beta,k} (a_{\beta k} + \lambda^T p_{\beta k}) m_{\beta,k} \\ \text{subject to the adjoint equation} \\ \dot{\lambda} = -\sum_{\beta,k} \frac{\partial}{\partial m_{\beta,0}} (a_{\beta k} + \lambda^T p_{\beta k}) m_{\beta,k}, \quad \lambda(T) = \nabla_{m_{\beta,0}} \varphi(T) \\ \text{and the state equation in moment variables} \\ \dot{m}_{\beta 0} = \sum_{\beta,k} p_{\beta k} m_{\beta k}, \quad m_{\beta 0} = x(0) = x_0, \\ \text{and the corresponding moment and localizing matrices} \\ \text{related to the space of moments } \Gamma \\ M_i(m) \succeq 0,\ M_{i-d_i}(Q_1 m) \succeq 0,\ M_{i-d_i}(Q_2 m) \succeq 0. \end{cases}$

Note that it is given a sequence $m = \{m_{\beta k}\}$ indexed in the basis of $\mathbb{R}[x, s]$, we denote $\{m_{\beta 0}\}_{\beta \in \mathbb{N}^n}$ the marginal with respect to the variable x, and $\{m_{0k}\}_{k \in \mathbb{N}^q}$ the marginal with respect to the control variable s. These sequences are indexed in the canonical basis of $\mathbb{R}[x]$ and $\mathbb{R}[s]$ respectively. Also note that the optimum H_i^* is a lower-bound on the global optimum H^* of the original problem (41), since any feasible solution (x, s) yields a feasible solution m of LP_i through Equation (42). Moreover, $H_i^* \leq H_{i'}^*$, when $i \geq i'$. We refer to problem LP_i as the semidefinite program relaxation of order i of (41). If any feasible point of the relaxation of order i is bounded then $H_i^* \mapsto H^*$ as $i \mapsto \infty$. The LMI constraints of LP_i state necessary conditions for m to be the vector of moments up to order $2i$, of some probability measure μ with support constrained in $\mathcal{S}$. This implies that $\inf LP_i \leq H^*$, as the vector of moments of the Dirac measure at a feasible point of (41), is feasible for LP_i. Since the convex relaxation of the polynomial optimal control problem has convex structure in the state x, and in the control variable s, one may suppose, under mild assumptions, that the problem has a minimizer m^*. Hence, we can obtain minimizers of the polynomial problem, and then to obtain minimizers of the switched optimal problem (21). We now can state an important result in the following theorem

Theorem 27 *Consider the problem defined in (34)-(36) and let* $\deg(Q_1) = q + 1$. *Then for every* $i \geq n + q + 1$.

(a) LP_i *is solvable with* $H^* = \min LP_i$ *and to every optimal solution* (x^*, s^*) *of (34)-(36) corresponds the optimal solution*

$$m_{\beta,0}^* = (x_1^*, ..., x_n^*, (x_1^*)^2, x_1^* x_2^*, ..., (x_1^*)^{2i}) \tag{46}$$

and

$$m_{0,k}^* = (s^*, ..., (s^*)^{2i}) \tag{47}$$

of LP_i*;*

(b) *every optimal solution* m^* *of* LP_i *is the finite vector of moments of a probability measure finitely supported on* v *optimal solutions of (34)-(36), with* $v = rankM_i(m) = rankM_n(m)$.

In many cases, low order relaxation (that is with $i << n$) will provide the optimal value H^*. We provide a criterion based on the work presented in [47], to detect whether some relaxation LP_i achieves the optimal value H^*. One way is to determine by inspection whether an optimal solution m of LP_i is a moment vector. This will be the case if, for instance, $rankM_r(m) = 1$. However, in case Equations (34)-(36) have multiple optimal solutions, it can happen that m is a convex combination of moments of Dirac measures supported on the optimal solutions, which in general is not easy to detect. The next criterion let us to test if the relaxation LP_i achieves the optimal value H^* [47].

Theorem 28 *Consider the problem defined in (34)-(36) and let* m^* *be an optimal solution of* LP_i *with* $i < n + q + 1$. *If*

$$rankM_{i-q}(m^*) = rankM_{i-q-1}(m^*),$$

then $\min LP_i = H^*$ *and* m^* *is the vector of moments of a probability measure supported on* $v = rankM_i(m^*) = rankM_{i-q-1}(m^*)$ *optimal solutions of (34)-(36).*

Considering the result presented in Theorem 27, we can set the correspondence between the minimizer of (35) and the minimizers of LP_i, and then set a correspondence with the minimizers of the original optimal switched problem (21).

Theorem 29 *If the moment sequence* $m^* = \{m^*_{\beta,k}\}$ *is the unique minimizer of the semidefinite program* LP_i, *then the problem (34)-(36) admits a unique minimizer*

$$(x^*, s^*) = m^*_{\beta 1} \tag{48}$$

where we can state, using the equivalent representation, that the autonomous optimal switching is given by

$$\sigma^*(t) = s^*(t) = m^*_{0,1}(t) \tag{49}$$

and the optimal trajectory is given by the first n-terms of the moments sequence

$$x^*(t) = m^*_{\beta,0}(t), \tag{50}$$

This result is particularly convenient to obtain an algorithm for switching control law, as it is presented in the next section.

3.3.2 Switched Optimization Algorithm

In this section we sketch the algorithm we use to numerically solve the optimal control problem formulated in Definition 16. The algorithm is mainly based on the interrelationship of the following:

(a) The equivalent optimal control problem EOCP

The EOCP is formulated in Section 3.1.1 where it is used the equivalent representation of the switched system and the running cost to obtain a polynomial Hamiltonian system

(b) The relaxation of the EOCP

The relaxation let us to obtain a generalized maximum principle which transforms the problem in a suitable form to apply the theory of moments. Therefore, we obtain an equivalent linear convex formulation.

(c) SDP relaxation algorithm

With the theory of moments we obtain a semidefinite program, which can be solved efficiently by numerical algorithm and then we can apply Theorem 27 and Theorem 29 to obtain an optimal switching law.

we propose the following control algorithm.

Algorithm SOCP

We start by partitioning the time interval $[t_0, t_f]$ into N subintervals with points $t_0 < t_1 < \cdots < t_N = t_f$.

Step 1. Obtain the equivalent representation of the optimal switched problem using Theorem 21 and Proposition 22.

Step 2. Apply the SDP relaxation over the equivalent polynomial optimal control problem obtaining the new set of moment variables and the LP_i programs.

Step 3. Set the initial conditions:

- $m_{00} = 1$
- the switching signal $m_{0k} = 0$ for all $1 \leq k \leq q$ (the switching system starts with the subsystem f_0 at t_0, i.e., $\sigma(0) = 0$)
- the state variables in the moment variables: $m_{\beta 0} = x(0) = x_0$

Step 4. Solve the LP_i program for the corresponding point of time and calculate the optimal value of H^* using Theorem 28. In order to solve the LP_i program we use an indirect shooting method.

Step 5. Calculate the moments corresponding to the next point of time, use Theorem 29 to set the switching signal and the state variables for the corresponding point of time.

Step 6. If terminal conditions are not fulfilled goto Step 4.

In the next section we present a numerical example to show the results presented in this section.

3.3.3 Numerical Example: The Artstein's Circle

We present an illustrative example of a switched nonlinear optimal control problem reformulated by Theorem 1 and Proposition 2 as a polynomial optimal control problem and solved by the Algorithm SOCP proposed using the theory of moments. We illustrate an efficient computational treatment to study the optimal control problem of a switched system reformulated as a polynomial expression. Consider the so-called Artstein's circle, considered as a polynomial switched system presented for stability analysis in [83], [27]. It is established that this system is asymptotically controllable and asymptotically stabilizable. It is described by the equation

$$f(x) = \begin{bmatrix} (-x_1^2 + x_2^2)u \\ -2x_1x_2u \end{bmatrix} \tag{51}$$

where $u < 0$ generates clockwise movement, $u > 0$ generates counterclockwise movement, and $u = 0$ induces an equilibrium point. For this system, a continuous feedback law that asymptotically stabilizes the origin does not exist. One characteristic of this problem is that if initially the state is on the circle

$$x_1^2 + (x_2 - c)^2 = c^2$$

then any resulting trajectory remains on that circle regardless of the input. We impose some state constraints $\{x \in \mathbb{R}^2,\ x_1 \le c\}$ with $c \in (0, 1)$. If we set $c = 0.577$, and $x_1 = c = 0.577$, we obtain $x_2 = 1.1143$. It is shown in Figure 3 a phase plane for two different initial conditions. The system shows a stable behavior for both cases. We propose to use the SOCP to stabilize the origin of system (51) respecting the input and state constraints. Consider the set of systems described by the drift vector field

$$\mathbf{F}(x) = [f_0(x)\ f_1(x)]$$

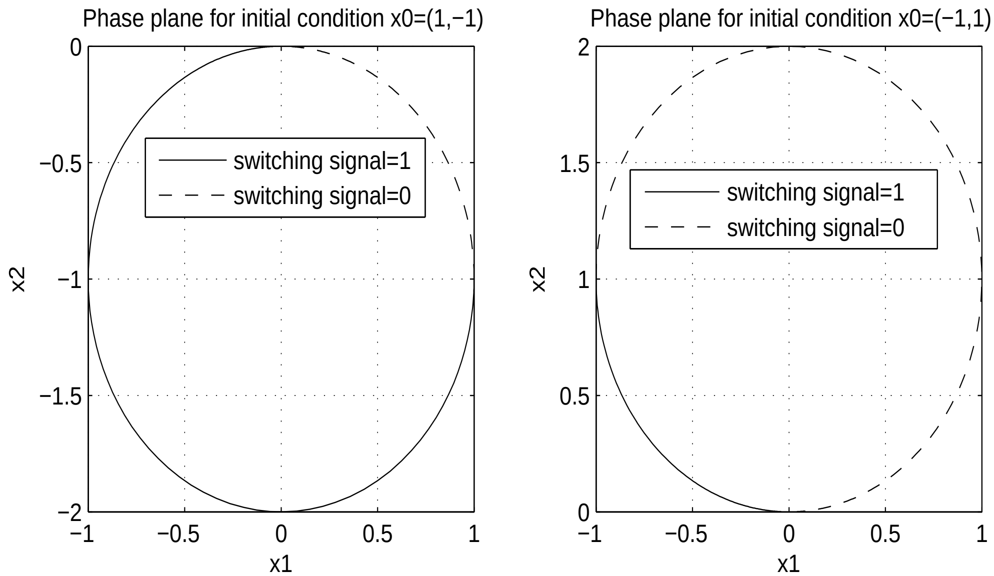

Figure 3: Phase plane for two different initial conditions

with

$$f_0(x) = \begin{bmatrix} -x_1^2 + x_2^2 \\ -2x_1x_2 \end{bmatrix}$$

and

$$f_1(x) = \begin{bmatrix} x_1^2 - x_2^2 \\ 2x_1x_2 \end{bmatrix}$$

Note that $f_0(x)$ and $f_1(x)$ are obtained from (51) setting $u = 1$ and respectively $u = -1$. We use the polynomial equivalent representation to obtain the polynomial optimal control problem,

$$\begin{array}{cl} \min_{x,s} & \int_0^{t_f} \left(x^T \mathbf{R}_0 x(1-s) + x^T \mathbf{R}_1 x s \right) dt \\ \text{s.t.} & \begin{array}{l} \dot{x}(t) = f_0(x)(1-s) + f_1(x)s \\ x \in \mathbb{R}^n, s \in \Omega, x(0) = x_0 \end{array} \end{array}$$

With this polynomial problem we obtain a polynomial Hamiltonian system as in (34)-(36). Hence, we are ready to apply the SOCP Algorithm using the SDP relaxation in moments. Consider a regulator problem, we want to stabilize the system minimizing the control energy, in this case, the switching between the subsystems(i.e., $\sigma \in \mathcal{Q} =$

$\{0, 1\}$. We use matrices $\mathbf{R}_0 = \mathbf{R}_1 = \mathbf{I}_{2\times 2}$ to set the running cost for both subsystem, $L_0 = L_1 = x^T \mathbf{I}_{2\times 2} x$ with initial time $t_0 = 0$, and final time $t_f = 10$. The degree of the polynomial equivalent system is the biggest degree of the field and the running cost, i.e., $r = 2$. The number of variables is the number of states plus s, i.e., $n = 3$ because $\mathbb{R}^2 \times \mathbb{R}$. We obtain with this data a basis in a lexicographical order, i.e., $1, x_1, x_2, s, x_1^2, x_1x_2, x_1s, x_2^2, x_2s, s^2$. We recall that moment and localizing matrices have the rows and columns indexed in the previous basis of polynomials. Define the sets

$$\Omega = \{s \in \mathbb{R} |\ Q_1 = s(s-1) \geq 0,\ Q_2 = s(1-s) \geq 0\}, \quad \mathcal{S} = \{\Omega \times \mathbb{R}^3\},$$

the moment matrix with $i \geq \max\deg = 2$, $M_2(m)$, and the localizing matrices with $d_i = (q+1)/2 = 1$, $M_1(Q_1 m)$ and $M_1(Q_2 m)$. Using the set $\mathcal{S}$ and moment and localizing matrices we set the problem in moment variables obtaining the positive semidefinite programs (LP_i). Solving the (LP_i) programs in time, with a relaxation order $i = 2$ and the algorithm SOCP, we obtain an optimal value of $H^* = -8.547$, and the moment sequence which allows us to calculate the switching signal and the trajectories. Figure 4 shows the trajectories, the co-state, and the switching signal obtained for a relaxation of order $i = 2$, solution of the corresponding optimal control problem. In figure 5 it is shown a phase plane of the system trajectories with the switching signal. It is clear the pints where the two subsystems switch between them, and eventually the switched system reach the stable point $(0, 0)$. This numerical example allow us to confirm that the first moment, i.e., m_{001}, is equivalent to the polynomial variable s, and therefore with the switching signal $\sigma(t)$ as it is stated in Theorem 27.

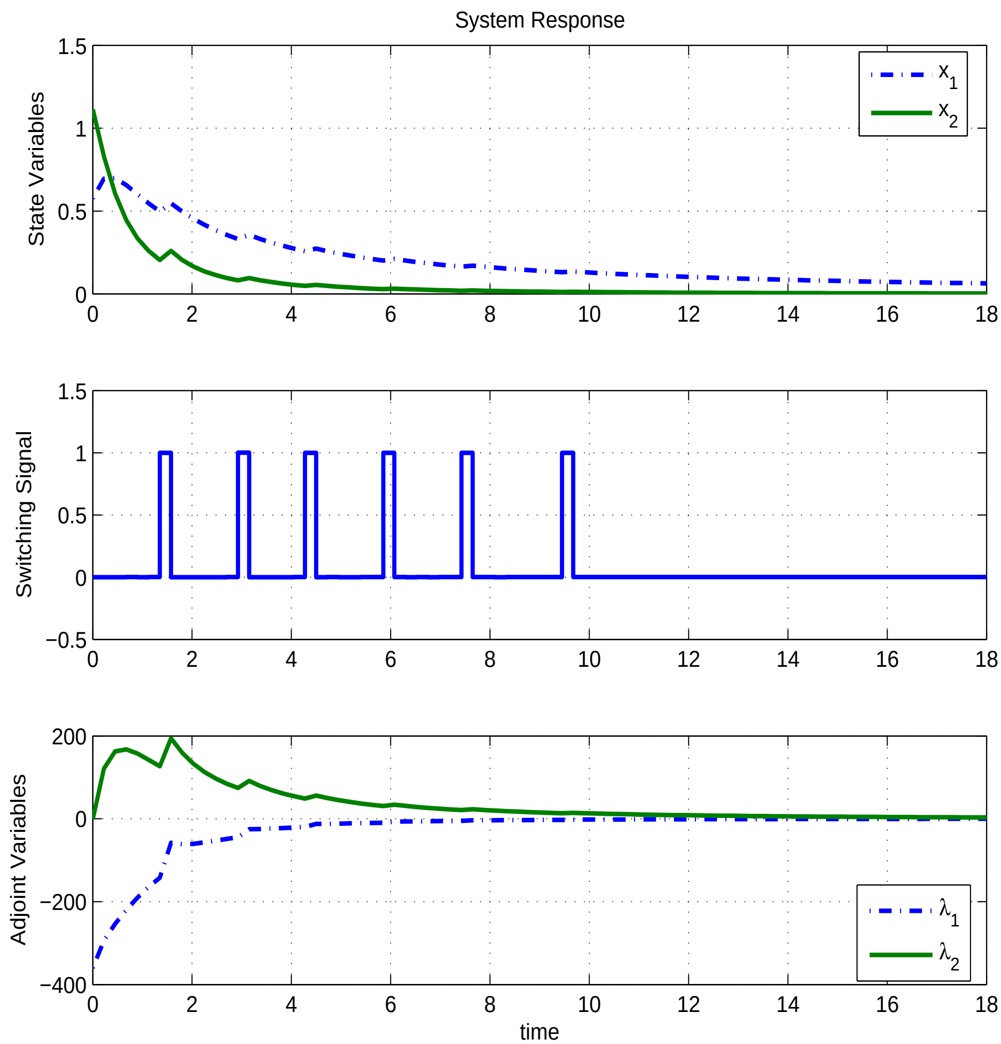

Figure 4: States, co-states, and switching signal for the Arstein's circle example

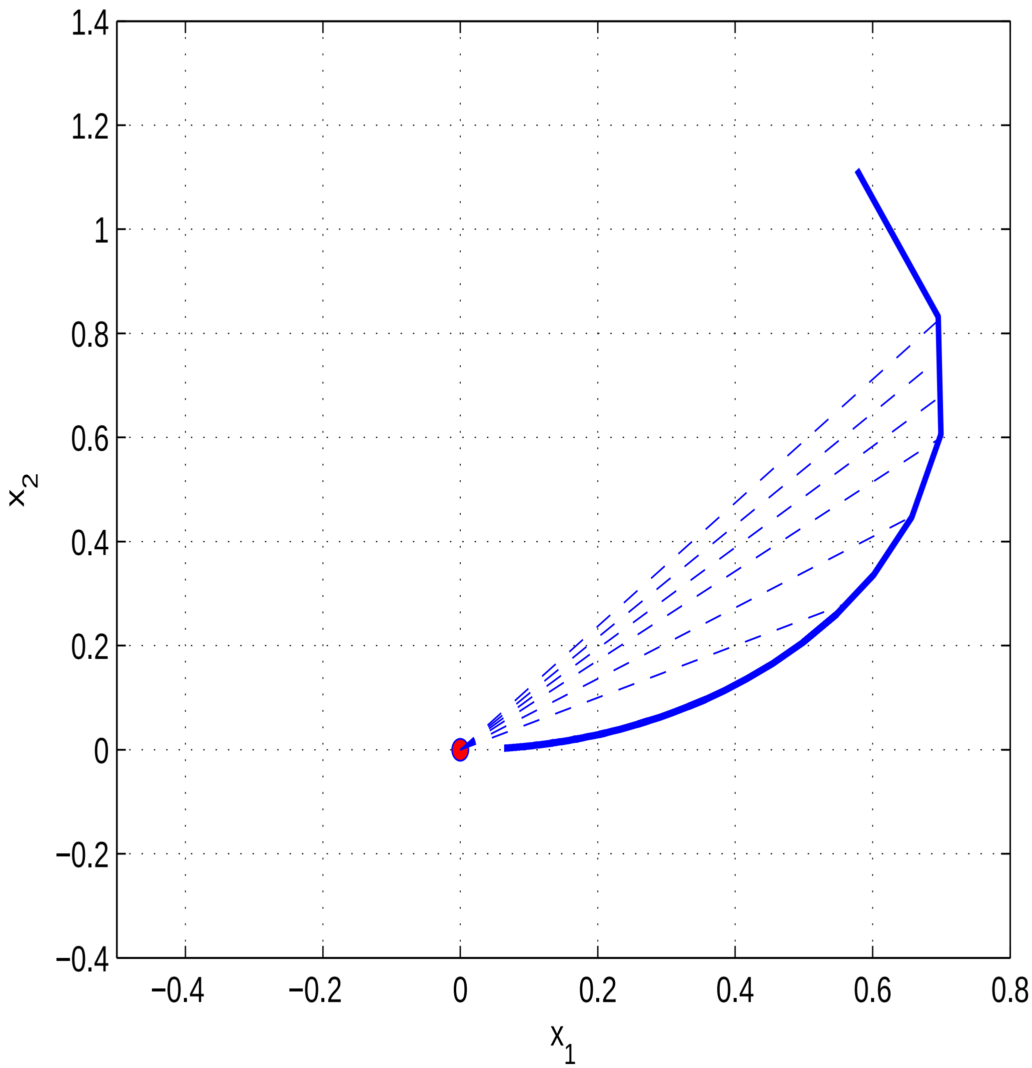

Figure 5: Phase plane of the system response for the Arstein's circle example

3.4 Extension Results to More General Nonlinear Optimal Control Problems

In this section, we omit the Assumption 23, and let the state equations and the running costs be of a more general class. Namely, we deal with elementary and nested elementary functions. This class of functions is related with explicit symbolic derivatives, for instance, exponential, logarithm, power, trigonometric, and hyperbolic functions, among others. We have mentioned that a very large class of non-polynomial nonlinearities can be converted into polynomial systems. This process introduced in [92], and later used for stability analysis in [75] and in Chapter 2, is based on the recasting process of elementary and nested elementary functions. Next, we describe the recasting process and how it can be used in optimal control problems of switched systems.

3.4.1 The Recasting Process

The recasting process, as it is presented in Chapter 2 for stability analysis and presented here for convenience in the reading, is presented as a procedure with several steps until the system has the expected form. The algorithm is as follows.

- **Step 0. Equivalent Representation:** We consider the equivalent representation for the switched system obtained in Theorem 1, and we name it as the original system (i.e., before the recasting process), with $\xi = (\xi_1, \ldots, \xi_n)$ as the state variables and s as the control variable.

- **Step 1. Original State Equations:** The original system is described by

$$\dot{\xi}_i = \sum_j a_j \prod_k p_{ijk}(\xi, s), \quad i = 1, \ldots, n \tag{52}$$

  here a_js are real numbers, and the factors p_{ijk} are elementary functions, or nested elementary functions of elementary functions.

- **Step 2. Decomposition of Non-Polynomial Functions:** Let $x_i = \xi_i$, for $i = 1, \ldots, n$. For each $p_{ijk}(\xi, s)$ in equation (52) that is not already a power-law function, replace it with a new variable x_{n+1}. This variable simplifies the differential equation to sums and products of power-law functions. An additional differential equation is generated for each new variable, using the chain rule of differentiation.

- **Step 3. Recasting Process:** When the recasting process leads to some constraints in the new variables, we have to introduce an n-dimensional manifold on which the solutions to the original differential equation lie. The particular choice of initial conditions defines the reference manifold.

- **Step 4. The Polynomial Form:** If the set of equations is in polynomial form, then the recasting process is complete. If not, repeat steps 2-3 until to obtain a system of equations with a rational or polynomial form.

Remark 30 *Notice that the constraints introduced by the definition of new variables, and their initial conditions, restrict the system behavior to a manifold of the same dimension of the original problem.*

Let first define the set of polynomial constraints as it has been done in the previous section, we treat the polynomial $Q(s)$ as two opposite inequalities $Q_1(s) = Q(s) \geq 0$, and $Q_2(s) = -Q(s) \geq 0$, and we redefine the compact set Ω to be

$$\Omega_1 = \{Q_i \geq 0, \;\; i = 1, 2\}$$

As a result of the recasting process we have obtained new variables, so that we take them into account. Suppose that for a switched system consisting of subsystems of non-polynomial form, we apply the equivalent representation and obtain a system,

$$\dot{\xi} = p(\xi, s)$$

The recasted system obtained using the procedure presented above is written as

$$\begin{aligned} \dot{x}_o &= p_o(x_o, x_r, s) \\ \dot{x}_r &= p_r(x_o, x_r, s) \end{aligned} \tag{53}$$

where $x_o = (x_1, ..., x_n) = \xi$ are the state variables of the original system, $x_r = (x_{n+1}, ..., x_{n+m})$ are the new variables introduced in the recasting process, $p_o(x_o, x_r, s)$ and $p_r(x_o, x_r, s)$ have polynomial forms. Previously, in the recasting process we have also obtained new polynomial constraints. Consider the real-valued polynomial $g_k(x_o, x_r, s)$, with $k = 1, ..., m$, where m is the number of polynomial constraints generated in the recasting process. Let Ω_2 be the set of polynomial constraints defined by the real-valued polynomial that we obtain from the recasting process. Again, we treat the polynomials $g_k(x_o, x_r, s)$ as two opposite inequalities $g_k(x_o, x_r, s) = g_k(x_o, x_r, s) \geq 0$, and $g_{k+m}(x_o, x_r, s) = -g_k(x_o, x_r, s) \geq 0$ so that we define the following set. let

$$\Omega_2 = \{(x_o, x_r, s) \in \mathbb{R}^{n+2m} \times \Omega_1 \mid g_k(x_o, x_r, s) \geq 0, g_{k+m}(x_o, x_r, s) \geq 0, \quad k = 1, ..., m\}$$

be the set of constraints generated by the recasting process. Finished the recasting process, we have to redefine the set of constraints of the equivalent polynomial representations. Let

$$\mathcal{D} = \{(x, s) \in \mathbb{R}^{n+m} \times \Omega_1 \times \Omega_2\} \tag{54}$$

be the compact semi-algebraic set, where Ω_1 and Ω_2 are defined as above. Using this polynomial dependence, we rewrite the EOCP.

Let $\mathbb{R}[x_o, x_r, s] = [x_1, ..., x_n, x_{n+1}, ..., x_{n+m}, s]$ denote the ring of real-valued polynomials in the variables x_o, x_r, s, with bases ordered lexicographically as it is shown in the previous section, for the vector space of real-valued polynomials of degree at most r in x_o, x_r, and the scalar polynomial variable s of degree at most q. Then, an r-degree polynomial running cost $\mathcal{L}(x_o, x_r, s) : \mathbb{R}^{n+m+1} \mapsto \mathbb{R}$ is written as $\mathcal{L}(x_o, x_r, s) = \sum_{\eta,k} a_{\eta k} x^\eta s^k$ and a polynomial map $p(x_o, x_r, s) : \mathbb{R}^{n+m+1} \mapsto$

$\mathbb{R}^{n+m+1}$, $p(x_o, x_r, s) = \sum_{\gamma,k} p_{\gamma k} x^{\gamma} s^k$ for some coefficients $a_{k\eta}$ and $p_{k\gamma}$, that depend on $\mathcal{L}(x_o, x_r, s)$ and $p(x_o, x_r, s)$, respectively for every η, $k \in \mathbb{N}^r \times \mathbb{N}$. We define $\beta = \max\{\eta, \gamma\}$ as the biggest degree of the polynomials, and use a canonical basis on that, putting zeros where it is needed. Then, we use an unified exponent and introduce the Hamiltonian as in (33).

3.4.2 SDP Relaxation

Consider the polynomial form of the Hamiltonian, we use the same ideas of the previous section for relaxation based on Young measures and the theory of moments.

Let $\mathcal{D}$ be the set of admissible trajectories $\mathcal{D} = \{(x, s) : (x(t), s(t)) \in \Omega_1 \times \Omega_2\}$. The set of Young measures associated to trajectories in $\mathcal{D}$ is

$$\widetilde{\mathcal{D}} = \left\{ \mu : supp(\mu) \subset \mathcal{D},\ a.e., t \in (0, T),\ \int_0^T \int_{\mathcal{D}} x^{\alpha} s^{\nu} d\mu(x, s) dt < \infty \right\},$$

where μ is a probability measure supported in $\mathcal{D}$. The extended functional $\widetilde{J}(x, s)$ defined on $\mathcal{D}$ is given by

$$\widetilde{J}(x, s) = \varphi(x(T)) + \int_0^T \int_{\mathcal{D}} \mathcal{L}(x, s) d\mu(x, s) dt$$

where $x(t)$ is the solution of

$$\dot{x}(t) = \int_{\mathcal{D}} p(x, s) d\mu(x, s), \quad x(0) = x_0,$$

This initial value problem is well-posed because this function satisfies the Lipschitz condition in x necessary to ensure a unique solution.

Let $m = \{m_{\beta,k}\}$ be a sequence of real numbers, the moment and localizing matrices

$$M_i(m) \succeq 0, \ \ M_{i-d_i}(Q_1 m) \succeq 0, \ \ M_{i-d_i}(Q_2 m) \succeq 0,$$

where $M_i(m)$ is a moment matrix associated to the vector of moments m, $M_{i-d_i}(Q_1 m)$, and $M_{i-d_i}(Q_2 m)$ are localizing matrices related to the vector of moments constrained to the set Ω_1, and $d_i = \lceil deg(Q_1/2) \rceil$. We easily see from Equation (30) that $deg(Q_1) =$

$deg(Q_2) = q+1$, where $q+1$ is the number of modes of the switched system, so that $d_i = \lceil (q+1)/2 \rceil$. Let $w_k = 2v_k$ or $w_k = 2v_k - 1$ be the degree of the polynomial $g_k(x_o, x_r, s)$ depending on its parity, consider the localizing matrices related to the polynomial constraints obtained by the recasting process, for $i \geq \max_k v_k$

$$M_{i-v_k}(g_k m) \succeq 0, \quad \forall k = 1, ..., 2m$$

then m has a representing measure with support contained in $\mathcal{D}$. As in the previous section, we define a space of moments as

$$\Lambda = \Big\{ m = \{m_{\beta k}\} : \ m_{\beta k} = \int_{\mathcal{D}} x^{\beta} s^{k} d\mu(x,s), \quad \mu \in P(\mathcal{D}), \cdots$$

$$\cdots M_i(m) \succeq 0, \ M_{i-d_i}(Q_1 m) \succeq 0, \ M_{i-d_i}(Q_2 m) \succeq 0, \ M_{i-v_k}(g_k m) \succeq 0 \}$$

where μ is a probability measure supported in $P(\mathcal{D})$. It is shown in the previous section that we can take advantage of the moment structure of the Hamiltonian and the state equation to rewrite the relaxed formulation obtained in Theorem 24 as a SDP.

For $i \geq \max[deg(H), \max_i deg(Q_i), \max_k v_k]$ consider the positive semidefinite programs (GLP_i)

Semidefinite programs-GLP_i:

$$
LPG_i: \begin{cases}
\text{Minimize the Hamiltonian defined in moments} \\
H_i^*(m) = \min_{m_{\beta k}\in\Lambda} \sum_{\beta,k}(a_{\beta k} + \lambda^T p_{\beta k})m_{\beta,k} \\
\text{subject to the adjoint equation} \\
\dot{\lambda} = -\sum_{\beta,k} \frac{\partial}{\partial m_{\beta,0}}(a_{\beta k} + \lambda^T p_{\beta k})m_{\beta,k}, \quad \lambda(T) = \nabla_{m_{\beta,0}}\varphi(T) \\
\text{the state equation in moment variables} \\
\dot{m}_{\beta 0} = \sum_{\beta,k} p_{\beta k}m_{\beta k}, \quad m_{\beta 0} = x(0) = x_0, \\
\text{the corresponding moment and localizing matrices} \\
M_i(m) \succeq 0,\ M_{i-d_i}(Q_1 m) \succeq 0,\ M_{i-d_i}(Q_2 m) \succeq 0, \\
\text{and the localizing matrices related to the polynomials} \\
\text{obtained in the recasting process} \\
M_{i-v_k}(g_k m) \succeq 0, \forall k = 1, ..., 2m.
\end{cases}
$$

The GLP_i programs can be used in a slightly variant of the Algorithm SOCP, taking into account the same considerations stated in the previous section for the LP_i programs. Basically, in Step 4 of the Algorithm SOCP we replace LP_i by GLP_i programs. In the next section we present a numerical example to illustrate the effectiveness of the approach presented in this section.

3.4.3 Numerical Example: Swing-up a Pendulum

In this example, we are dealing with a two-dimensional model of the pendulum and thus, the acceleration of its pivot is assumed to be the control input. Swinging up and stabilization of the pendulum is usually solved by switching between different laws. We use a damping-pumping strategy as it is proposed in [10]. The normalized model of the pendulum, when the control input is the acceleration of the pivot, is given by

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= \sin x_1 - (2\sin x_1 + F x_2 \cos x_1)\cos x_1 \end{aligned} \tag{55}$$

where x_1 is the angular position with respect to the origin, at the upright position, and x_2 is the velocity. Considering stabilization conditions, it is shown that we have to take F negative in some regions, and positive in some others, so that the system minimizes the energy consumption all the time. Therefore, this is equivalent to a switched system, where f_0 is defined by equation (55) when F is negative, and f_1 when it is positive.

First, we use Theorem 1 to obtain an equivalent continuous representation of the switched model related to (55) for the pendulum, $\dot{\xi}(t) = f_0(\xi)(1-s) + f_1(\xi)s$ with $s \in \Omega_1 = \{s \in \mathbb{R} \,|\, Q(s) = s(s-1)\}$. We obtain the following equivalent continuous system,

$$\begin{aligned} \dot{\xi}_1 &= \xi_2 \\ \dot{\xi}_2 &= \sin\xi_1 - 2\sin\xi_1\cos\xi_1 + (1-2s)\xi_2\cos^2\xi_1 \end{aligned} \tag{56}$$

Now, following the recasting process, it is clear that (56) is in the same form as (52), but in this case the elementary functions are trigonometric functions. Let us follow Step 2 to Step 4 in the recasting process. As a result, we obtain a new set of differential equations given by

$$\begin{aligned} \dot{x}_1 &= \dot{\xi}_1 = \xi_2 = x_2 \\ \dot{x}_2 &= \dot{\xi}_2 = x_3 - 2x_3x_4 + (1-2s)x_2x_4^2 \\ \dot{x}_3 &= \dot{x}_1\cos x_1 = x_2x_4 \\ \dot{x}_4 &= -\dot{x}_1\sin x_1 = -x_2x_3 \end{aligned} \tag{57}$$

As we know by Step 3 in the recasting process, when we introduce the new variables x_3 and x_4, a set of constraints arise. For this case, we have that the manifold on which the solutions to the original system (56) lie is given by $\Omega_2 = \{x_3^2 + x_4^2 - 1 = 0\}$. Using this reformulation we recast the nonlinear optimal control problem in an equivalent polynomial problem. Then, this reformulation allows us to apply the positive semidefinite relaxation using the theory of moments. We want to design a control law in such a way that the closed-loop energy presents a minimum at the desired position, and the energy controller is globally defined. Since the chosen target energy has other minima different from the desired equilibrium, a combination of energy dissipation (damping) and injection (pumping) is needed in order to globally stabilize the origin. Consider that we start the pendulum at the position $x_0 = (-\pi, 0.5)$. Previously, we have obtained a recasted system for the pendulum in equation (52), constrained by the set $\Omega_2 = \{x_3^2 + x_4^2 - 1 = 0\}$. We define the set for the Young measures $\mathcal{D}$. In this case, the constrained sets Ω_1 and Ω_2 are redefined as

$$\Omega_1 = \{s \in \mathbb{R}|\ Q_1 = s(s-1) \geq 0,\ Q_2 = s(1-s) \geq 0\},$$

$$\Omega_2 = \{g_1 = x_3^2 + x_4^2 - 1 \geq 0, g_2 = -x_3^2 - x_4^2 + 1 \geq 0\},$$

therefore we obtain the constrained set

$$\mathcal{D} = \{(x, s) \in \mathbb{R}^5 \times \Omega_1 \times \Omega_2\},$$

The degree of the equivalent polynomial system is the highest degree of the field plus the running cost, i.e., $r = 2$. The number of variables is the number of states obtained in the recasting process plus s, i.e., $n = 5$. Hence, we obtain with this data the basis in a lexicographical order, i.e., $1, x_1, x_2, x_3, x_4, s, x_1^2, x_1x_2,\ x_1x_3, x_1x_4,\ x_1s, x_2^2, x_2x_3, x_2x_4, x_2s, x_3^2, x_3x_4, x_3s, x_4^2, x_4s, s^2$. Which leads to the moment and localizing matrices. We recall that moment and localizing matrices have the rows and columns indexed in the previous basis of polynomials. We obtain a Hamiltonian in moment sequences and co-states $H(m, \lambda)$. We obtain a relaxed Maximum Principle (RMP) conditions for the moment sequence $m = \{m_{\beta,k}\}$ to be optimal, and we obtain the positive semidefinite programs GLP_i. With sets defined above we obtain the moment matrix with $i \geq \max\deg = 2$, $M_2(m)$, and the localizing matrices, with $d_i = (q+1)/2 = 1$, $M_1(Q_1 m)$, $M_1(Q_2 m)$, $M_1(g_1 m)$, and $M_1(g_2 m)$. Using the set $\mathcal{D}$ and moment and localizing matrices we set the problem in moment variables obtaining the positive semidefinite programs (GLP_i). Solving the (GLP_i) programs in time, with a relaxation order of $i = 2$, we obtain a suboptimal value of $H^* = 2,7745$, the moment sequence, which allows us to calculate the switching signal, the suboptimal trajectories, and the co-states (see Figure 6). The system response shows that the trajectories reach an equilibrium point $(0, 0)$.

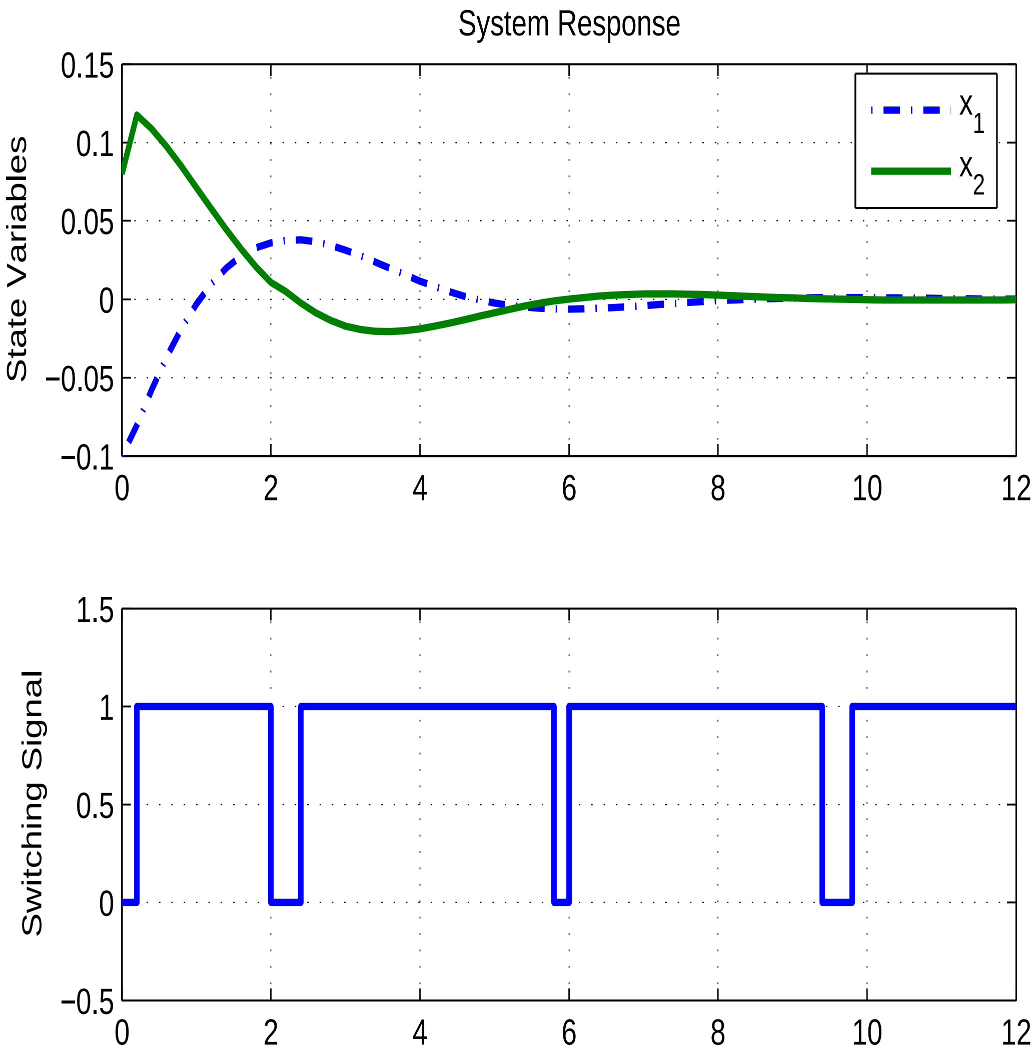

Figure 6: States and switching signal for the pendulum's example

CHAPTER IV

PIECEWISE-LINEAR APPROACH TO NONLINEAR CELLULAR GROWTH CONTROL

In this chapter, the peculiar features of mammalian cells growth in fed-batch operating condition are addressed. The task of the controller is to determine, in every instant, the best feed substrate, using the compilation of information on-line from the sensor. The determination of an optimal strategy of feed substrate using the nonlinear modeling either if the kinetics are known, is not a straightforward matter and is often further complicated by the presence of constraints imposed on the state variables [14]. This chapter is related with a modeling class of hybrid systems, piecewise-linear (PWL) systems. The PWL approximation, i.e., systems which are linear or affine on each of the components of a polyhedral partition of the state space [96], have shown advantages of implementation, performance analysis, and calculations [39], [84], [90], [89]. In this work, it is used a canonical piecewise linear approximation over simplicial partitions. It provides a partition of the state space into polytopic cells based on value at vertices [42], [89], [38]. This choice is motivated by several facts. First, this class of functions uniformly approximate any continuous nonlinear function defined over a compact domain $\mathbb{R}^n$ (see [42]). Moreover, the canonical expression introduced in [42] uses the minimum and exact number of parameters, and it is the first PWL expression able to represent PWL mappings in arbitrary dimensional domains. As a consequence of this, an efficient characterization is obtained from the viewpoint of memory storage and numerical evaluation [26]. Second, the approximation can be used in real implementations, the points from the nonlinear model may be replaced for points from sensors or data directly from the process, namely, it addresses the problem of finding a PWL approximation of system where a reasonable number of measure samples of the vector field is available (regression set) [98]. Third, this alternative approach deal with an approximation which is easier to handle than the nonlinear model. In fact, it can use many tools developed for hybrid systems, e.g., the MLD model based approach [16], since algorithms for translating MLD systems into PWL systems are available [15], [104]. Finally, this CPWL is used in a model based control termed, probing control in [68], being a first step to develop a hybrid probing control. This work refers to a probing control as it is presented in [7] for *E. coli*. Short pulses to the feed rate are added, and taking into account the system response, the pulse is increased or decreased according with the tuning rule. The probing control strategy avoids acetate accumulation while maintaining a high growth rate [7], [103]. The approximation model is tested by the probing control strategy. It is implemented in simulations for this mathematical model. The comparative analysis and error approximation between this new biological model and a nonlinear model developed first in [65], [64] are shown. This method is satisfactory for a implementation purpose of a hybrid probing control.

4.1 Process Description

The mammalian cells of Baby Hamster Kidney are used in the production of the vaccine against the foot-and-mouth disease. They are cultivated in a bioreactor located in the biotechnology laboratory *LIMOR de Colombia S.A*, in Bogotá, Colombia. It has a capacity of 2500 liters, operating in fed-batch mode, and has temperature and agitation speed controlled. Several experiments are carried out, allowing to obtain information of several cellular cultures. Then, the data are used for parameters adjustment of the mathematical nonlinear model [65]. In the nonlinear model approximation, it is used four state variables in order to analyze the basic cellular behavior, and the capability of the CPWL model to approximate better this behavior, compared with a simpler approximation of PWL based on a partition form of the state space.

4.1.1 Nonlinear Model

The biological system class of nonlinear dynamics may be described using the following model:

$$\begin{cases} \dot{x}(t) = f(x(t)) + B(x(t))u(t) \\ y(t) = Cx(t) \end{cases} \tag{58}$$

where $x(t) \in \mathbb{R}^n$ is a vector of state at time t, $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ are nonlinear vector-valued functions, B is a state-dependent $n \times m$ input matrix, $u : \mathbb{R} \rightarrow \mathbb{R}^m$ is an input signal, C is an $n \times k$ output matrix and $y : \mathbb{R} \rightarrow \mathbb{R}^k$ is the output signal. The control variable u is the dilution rate, it can be shown in equation (58) that the control variable has a direct relationship with state variables, therefore it has a highly nonlinear behavior.

It is presented a description with four state variables; cellular concentration, glucose concentration, dissolved oxygen concentration, and a waste product, i.e., acetate. The latter, it is the most relevant waste component produced in the cellular growth process. In addition, this model facilitates the reactor start-up and steady-state operation conditions, since the presence of less-desired steady states at the same inlet conditions, makes this bioreactor a challenging problem for control design. It can be shown in the simulation Section that the transient analysis evidence this behavior. As it has been shown in [64] the BHK cells have several growth phases:

i *Latency*, immediately after inoculation.

ii *Acceleration state*, when cells begin the growth process.

iii *Exponential state*, the cells reach a constant growth with the major growth rate.

iv *Deceleration state*, cells reduce the constant growth for absence of substrate and accumulation of toxic substrates.

v *Stationary state*, when the cellular growth reaches a constant value, since substrate is over, or it is in a low constant value.

For further details about the nonlinear model refer to [64]. The nonlinear model is as follows

$$\begin{cases} \dot{x}_1 = \mu_m \left(\frac{x_2}{\frac{x_2^2}{k_i} + x_2 + k_s} \right) \left(\frac{x_3}{x_3 + k_a} \right) x_1 - \left(\frac{k_{dm} k_{ds}}{(\mu_m - k_{dm} x_4)(k_{ds} + x_2)} \right) x_1 - x_1 u \\ \dot{x}_2 = -k_1 \mu_m \left(\frac{x_2}{\frac{x_2^2}{k_i} + x_2 + k_s} \right) \left(\frac{x_3}{x_3 + k_a} \right) x_1 + (S_{1i} - x_2) u \\ \dot{x}_3 = k_3 \mu_m \left(\frac{x_2}{\frac{x_2^2}{k_i} + x_2 + k_s} \right) \left(\frac{x_3}{x_3 + k_a} \right) x_1 - x_3 u \\ \dot{x}_4 = Kla(C^* - x_4) - k_1 \mu_m \left(\frac{x_2}{\frac{x_2^2}{k_i} + x_2 + k_s} \right) \left(\frac{x_3}{x_3 + k_a} \right) x_1 - x_4 u \end{cases} \tag{59}$$

where x_1 (g/L) is the cellular concentration, x_2 (g/L) is the glucose concentration, μ_m $(1/h)$ is the maximum cellular growth rate, k_s (g/L) is the glucose saturation parameter, x_3 (g/L) is the lactate concentration, x_4 (g/L) is the dissolved oxygen concentration, $k_a = 0.06$ (g/L) is the lactate saturation parameter, $k_{1,2,3}$ are yield coefficients, 1.46, 0.9, and 0.06 respectively. $S_{1i} = 1.5$ (g/L) is the initial glucose concentration in the feeding medium, $k_{dm} = 0.012$ $(1/h)$ is the maximum cellular death rate, and $k_{ds} = 0.08$ is the death saturation parameter. The volumetric oxygen transfer coefficient, Kla is a function of stirred speed but is also affected by the air flow rate and factors like viscosity, and foaming. C^* (g/L) is the dissolved oxygen concentration in equilibrium with the oxygen in gas bubble. All the nonlinear model parameters are taken from the real data through several experiments and are presented in [64].

Figure 1 shows the schematic of the bioreactor in a fed-batch operation mode with dissolved oxygen concentration as input variable and feed rate as control variable.

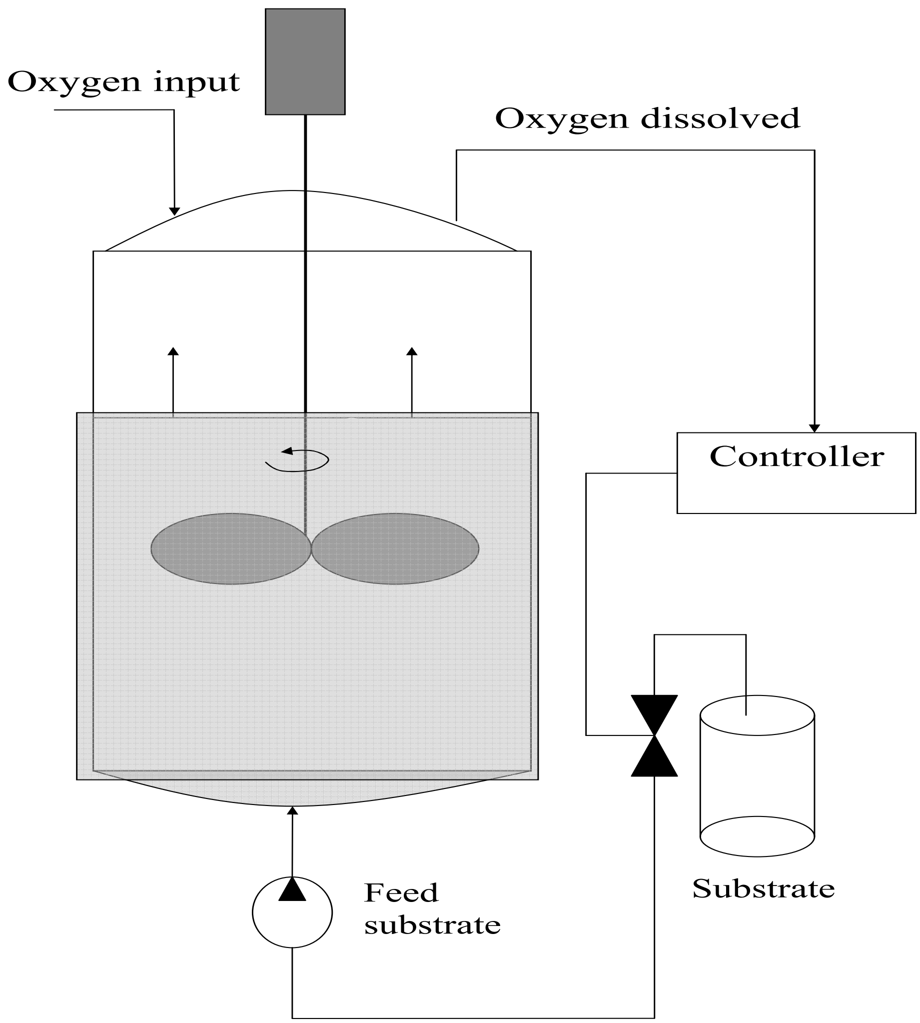

Figure 7: A cell bioreactor in a feed-batch operation mode

4.2 Biological CPWL model

The approximation deals with the PWL approximation technique developed in the past few years by [41]. To guarantee that the dynamical behavior of the PWL-approximate flow will be faithful to that of the original system for any values of some significant parameters, it is used the input variable u like a parameter, and simulated the PWL-flow. Figure 2 shows the schematic of the Bioreactor with CPWL model with the control loop and variables involved in this approach. First, some definitions about CPWL (Canonical Piecewise-linear) are given.

4.2.1 Orthonormal Canonical Piecewise Linear Functions

The orthonormal definition of the PWL functions given before in [42] is used in this work to represent the nonlinear static mapping of cellular growth process in bioreator (59). A brief description of this representation and its most important characteristics are given as following. For further details refer to [41], [42]. For this approximation, given a nonlinear system as in (3), the following steps are necessary to obtain a CPWL approximation:

Step 1. Order all vertices of the grid partition.

Step 2. Group the vertices into simplicial cells.

Step 3. Find CPWL basis functions.

Step 4. Find the CPWL model approximation.

Figure 2 shows a schematic of this piecewise approximation. In order to perform these steps, the following definitions and methods are presented:

Definition 31 . *Let* $\mathbf{x}^0, \mathbf{x}^1, ..., \mathbf{x}^n$ *be* $n+1$ *points in the* $n-$*dimensional space. A simplex (or polytope)* $\triangle(\mathbf{x}^0, \mathbf{x}^1, ..., \mathbf{x}^n)$ *is defined by*

$$\triangle(\mathbf{x}^0, \mathbf{x}^1, ..., \mathbf{x}^n) = \left\{\mathbf{x} : \mathbf{x} = \sum\nolimits_{i=0}^{n} \mu_i \mathbf{x}^n \right\}$$

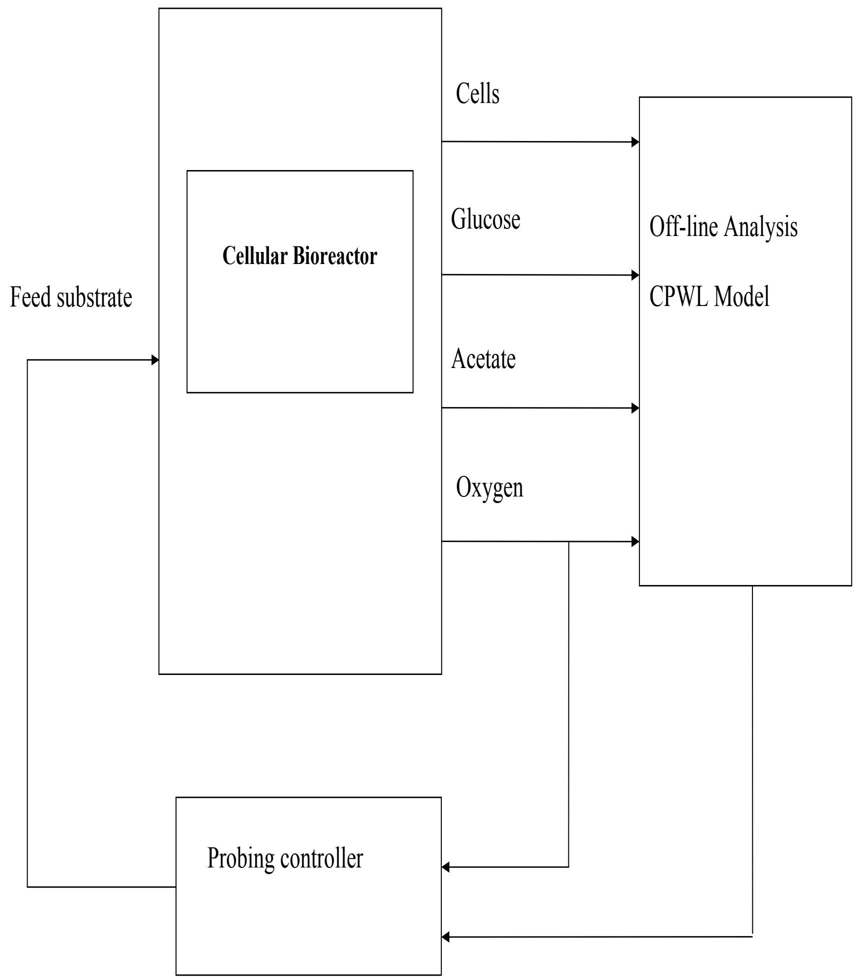

Figure 8: Schematic of the Bioreactor with CPWL model

where $0 \leq \mu_i \leq 1$, $i \in \{0, 1, ..., n\}$ *and* $\sum_{i=0}^{n} \mu_i$=1. *A simplex is said to be proper if and only if it cannot be contained in a* $(n-1)$ *dimensional hyperplane.*

The representation proposed in [41] requires the definition of a rectangular compact domain of the form

$$S = \{\mathbf{x} \in \mathbb{R}^m : 0 \leq x_i \leq n_i \delta_k, i = 1, 2, ..., m\} \tag{60}$$

where δ is the grid size and $n_i \in \mathbb{Z}_+$ (the set of positive integers). Then, this domain is subdivided using a simplicial boundary configuration H, a simplicial boundary configuration is characterized by the property that it produces a division of the domain into simplices. If S is defined as in (60), the space of all continuous PWL mappings defined over the domain S partitioned with a simplicial boundary configuration H is denoted by PWL[S_H]. At this point it is convenient to introduce the set V_S of all simplex vertices contained in S. Depending on the number of coordinates different from zero, these vertices are organized in classes; for instance, all the vertices in S that have k ($k \leq m$) coordinates different from zero are called *class* k vertices. The set V_S plays an important role due to the fact that the function values at the vertices of V_S constitute all the information that is needed to fully characterize any PWL function $f_p : S \longmapsto \mathbb{R}^1$. Figure 3 shows the basic concepts of simplicial partition and the grid size for an illustrative two dimensional system.

There are many possible choices for the PWL basis functions, each of which is made up of N (linearly independent) functions belonging to PWL $N-$dimensional linear space. Thus, any basis can be expressed as a linear combination of the elements of the β basis, which is defined by recursively applying the following generating function $\gamma(u, v)$ [41], [42]:

$$\gamma(x_1, x_2) \;=\; \{||x_1| + x_2| - ||x_2| - x_1| + |x_1| + |x_2| - |x_2 - x_1|\}/4 \tag{61}$$

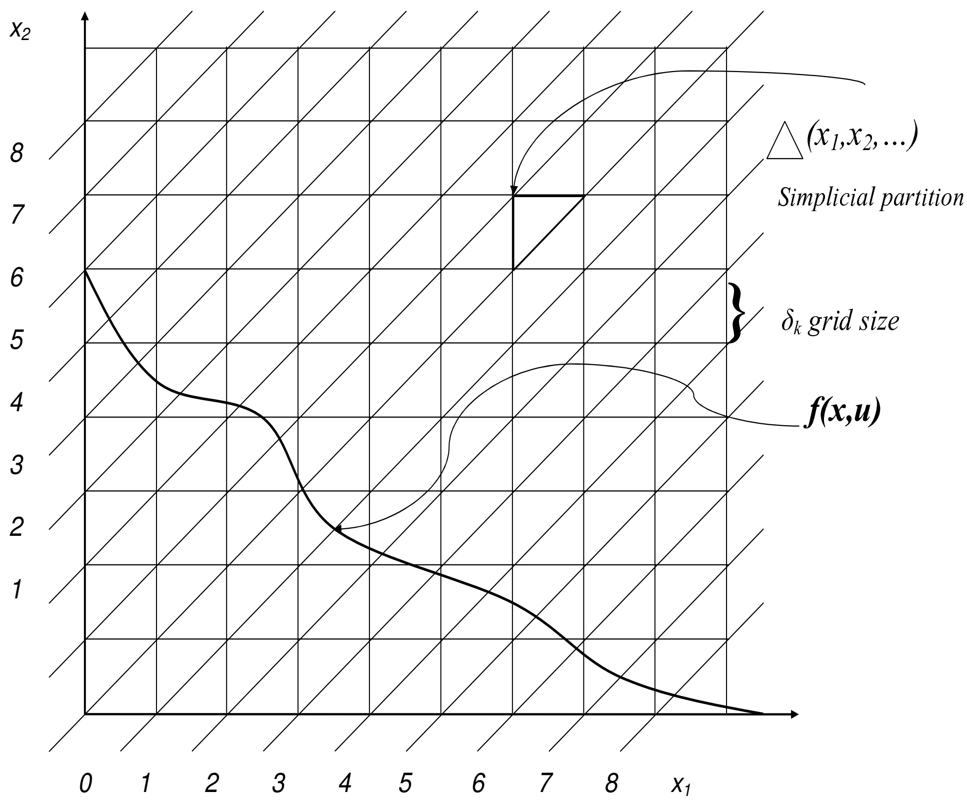

Figure 9: Simplicial partition and the grid size

Using (3), the functions becomes $\gamma^0(x_1) = x_1$, $\gamma^1(x_1) = \gamma(x_1, x_1)$, $\gamma^2(x_1, x_2) = \gamma(x_1, x_2)$, and in general

$$\gamma^k(x_1, ..., x_k) = \gamma(x_1, \gamma^{k-1}(x_2, ..., x_k)) \tag{62}$$

are defined. A distinctive property of (62) is that it possesses k nesting of absolute value functions, and accordingly, it is said to have *nesting level* ($n.l$) equal to k.

The first element of the basis is a constant term $\gamma^0(1)$. The remaining elements are formed by the composition of the function $\gamma^k(\cdot, ..., \cdot)$, $k = 1, 2, ..., m$, with the linear functions

$$\pi_{k,j_k}(\mathbf{x}) = x_k - j_k\delta_k$$

$k = 1, 2, ..., m$, $j_k = 0, 1, ..., n_k - 1$. As a result, the basis can be expressed in vector form, ordered according to its nesting level as

$$\Lambda = \left[\Lambda^{0^T}, \Lambda^{1^T}, ..., \Lambda^{m^T}\right] \tag{63}$$

where Λ^i is the vector containing the generating functions defined in [41] with i nesting levels. Accordingly, any $f_p \in \mathrm{PWL}[S_H]$ can be written as

$$f_p(\mathbf{x}) = \mathbf{c}^T\Lambda(\mathbf{x}) \tag{64}$$

where $\mathbf{c} = \left[\mathbf{c}^{0^T}, \mathbf{c}^{1^T}, ..., \mathbf{c}^{m^T}\right]^T$, and every vector $\mathbf{c}^i$ is a parameter vector associated with the vector function Λ^i.

In order to obtain an orthonormal basis, it is necessary to define an inner product on $\mathrm{PWL}[S_H]$. The new basis elements are linear combination of (64), that is $\Upsilon(\mathbf{x}) = T\Lambda(\mathbf{x})$, and the matrix T may be obtained by two different methodologies. In this work we use the one built using the Gran-schmidt procedure as given in [42]. To find the required approximation, we use a routine of [42] that finds a vector of parameters $\mathbf{c}$ that is the solution of the least square problem $\min_x |Ax - b|_2$, where $A = \Upsilon^T(\mathbf{X})$, $\mathbf{X}$ the input matrix and b the output to be approximated in sparse format. In accordance with [42], the CPWL approximation of the nonlinear function g is defined as the function $f_p \in$PWL$[S_H]$ satisfying

$$f_{\mathbf{CPWL}} = A\mathbf{c} \tag{65}$$

Equation (65) presents the CPWL approximation for the nonlinear system.

4.2.2 Analysis of CPWL Approximation: Error Estimation

In this section, the error estimation between the nonlinear system and the CPWL approximation (65) is presented. It can be noticed that it is only necessary to know the values of f at the vertices because, as stated above, this is all the information needed to uniquely define a function belonging to PWL$[S_H]$.

If the function f is continuous, the following results quantify the precision of the approximation.

Lemma 32 *If f is continuous in S, which is the union of nonoverlapping simplices, then*

$$\|f_{CPWL} - f(x)\| \leq \epsilon, \forall x \in S,$$

where

$$\epsilon = max_{\Delta \in S}\left\{max_{x_0,x_1 \in \Delta}\left(\|f_{CPWL} - f(x)\|\right)\right\}.$$

In addition, if the function is assumed to be Lipschitz continuous in S, a useful relationship between the approximation error and the grid size δ can be obtained.

Lemma 33 *Let $f \colon S \mapsto R^1$ be a function satisfying*

$$\|f_{CPWL} - f(x)\| \leq L\left\|x_1 - x_0\right\|,$$

$\forall x_1, x_0 \in S$ and $L \geq 0$ is the Lipschitz constant. Then, f_{CPWL} satisfies

$$\|f_{CPWL} - f(x)\| \leq \delta L \left\| \sum_{i=0}^{n} e_i \right\|, \forall x \in S,$$

where e_i is the error of each simplicial partition. It should be noticed that a choice of δ such that $\delta L \left\| \sum_{i=0}^{n} e_i \right\| = e$, or equivalently,

$$\delta = \frac{e}{L \left\| \sum_{i=0}^{n} e_i \right\|}$$

guarantees that the approximation error is bounded as follows: $\|f_{CPWL} - f(x)\| \leq e$

4.2.3 Cellular Growth CPWL Model

The cellular growth approximation scheme is based on the orthonormal canonical piecewise linear functions. It means the approximation of the right-hand side nonlinear function to a CPWL function for each differential equation. The PWL approximation based on a nonlinear model is obtained over the domain

$$S = \{\mathbf{y} \in \mathbb{R}^3 : a_1 \leq x_1 \leq b_1, a_2 \leq x_2 \leq b_2,$$
$$a_3 \leq x_3 \leq b_3, a_4 \leq x_4 \leq b_4, a_5 \leq u \leq b_5\}$$

with

$$a_1 = a_2 = a_3 = a_4 = a_5 = 0.0,\ b_1 = 4, b_2 = 4.5, b_3 = 2.6, b_4 = 1.8, b_5 = 0.3$$

The domain S is partitioned by performing $m_1,\ m_2, m_3, m_4$ subdivisions along the state components $x_1,\ x_2,\ x_3,\ x_4$ respectively, and m_5, subdivisions along the parameter component u, as the control variable. The coefficients are derived from a set of samples of f corresponding to a regular grid of $n_1 \times n_2 \times n_3 \times n_4 \times n_5$ points over the domain S. In particular, it is considered the following PWL approximation $f_{\mathbf{CPWL}}$ of f : $m_1 = m_2 = m_3 = m_4 = 6$, $m_5 = 3, n_1 = n_2 = 15$, $n_3 = n_4 = n_5 = 10$.

Using several values of subdivisions, it is found that the last values perform the system well, from both qualitative and quantitative points of view. It is noticed that the algorithm using orthonormal basis allows a larger number of inputs due to the ability of handling sparse matrices. Thus, it is possible to detect the simplices that contribute to the approximation, and it gives a measure of the nonlinearity of the function. It is possible to take smaller values of m in order to reduce the number of parameters, but the quality of the approximation become worse.

4.3 Simulation Results

4.3.1 Model Simulation Results - Nonlinear vs CPWL

In order to analyze the system behavior with different input values, corresponding to the dilution rate u, the original nonlinear system with the CPWL approximation for $u = 0$ is simulated. It is the batch operation mode, i.e., the natural system behavior. The simulation provides a plausible explanation of the different steady states. First, as in the previous section, it is simulated the CPWL model without input variable, dilution rate, in order to obtain a batch mode behavior. All the simulations are made with initial value $x_0 = (0.12, 4.2, 2.6, 0.0)'$. Figure 4 shows a comparison between the nonlinear system and the CPWL approximation obtained by numerically integrating the dynamical system having on the right-hand side of $f_{\mathbf{CPWL}}$ for a dilution rate $u = 0$.

In this case, it can be shown that the behavior of the four state variables using the CPWL model is qualitatively similar to the original system. For simulation purpose, the dilution rate is increased to a best experimental value, it has been obtained form the data of many experiments [64]. The CPWL presents very good agreement in this case. The dilution rate was then increased to reach a high dilution rate, where normally the system is not working, and the cellular concentration is too low. It is of particular interest to show that the obtained model is able to indicate the multiple

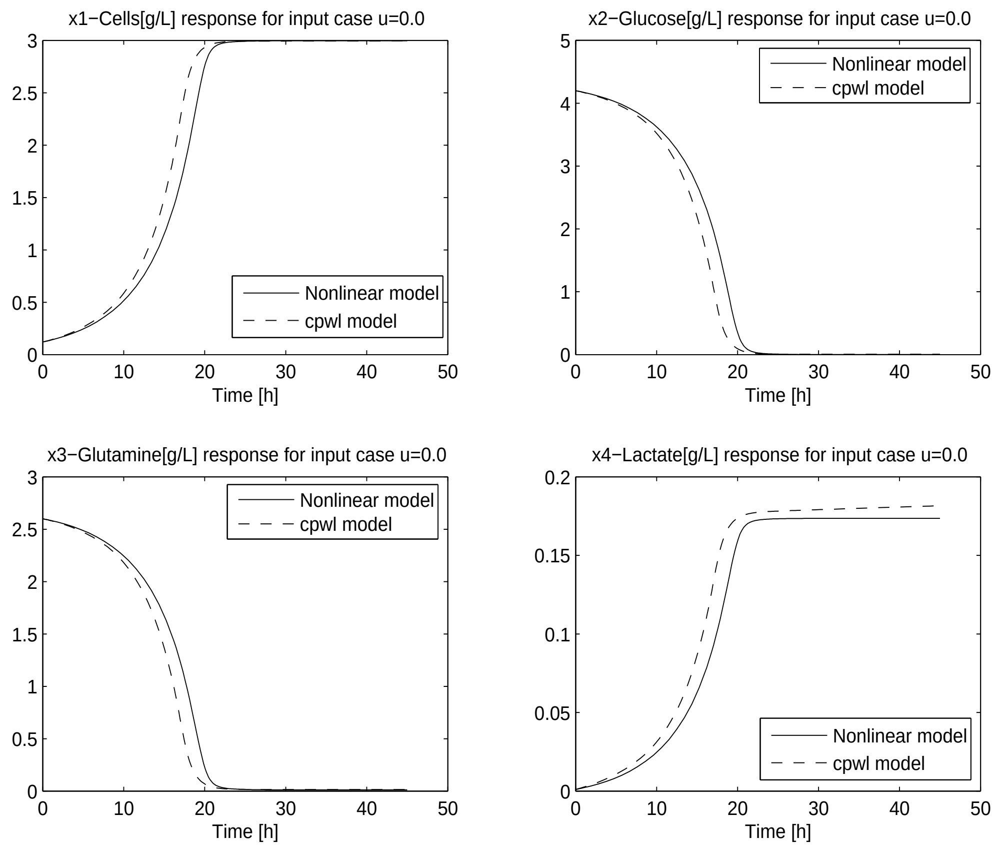

Figure 10: Comparison of system response computed with the nonlinear and cpwl models to a dilution rate u = 0.0

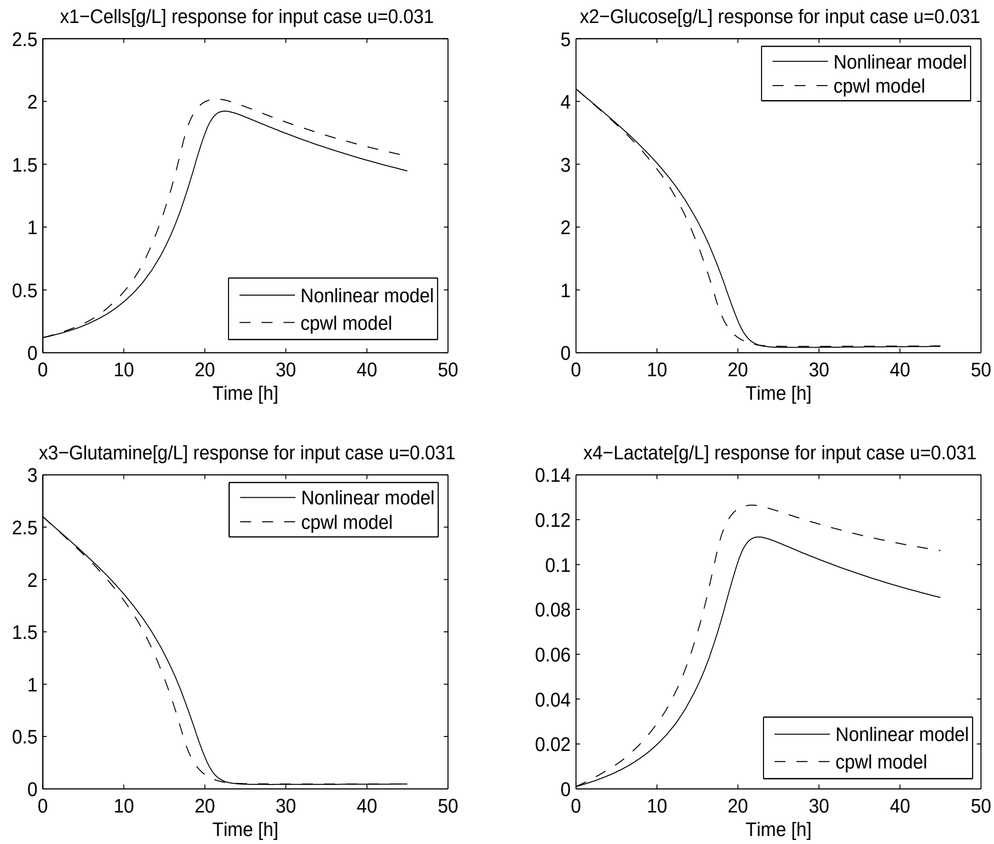

Figure 11: Comparison of system response computed with nonlinear and cpwl models to the dilution u = 0.031

steady states. Figure 5 shows the comparison between the nonlinear system and the CPWL approximation for these cases, the dilution rate is $u = 0.031/h$.

4.3.2 Transient Analysis of Cells Concentration

In this section, the transient analysis for different initial conditions is considered at the same dilution rate, and several steady states are found. In order to further investigation of these steady states, the dilution rate is established to, $u = 0.036/h$, and several initial conditions, with cellular and lactate initial conditions maintained unmodified, varying glucose and glutamine initial concentration, see Table 1.

It can be observed clearly three steady states (see Figure 6). Steady state 1 results

Table 1: Start-up Conditions for Transient Analysis

	S_1(Glucose g/L)	S_2(Glutamine g/L)
1	1.2	2.6
2	4.2	4.2
3	7.2	2.6
4	4.2	0.6
5	7.2	4.2

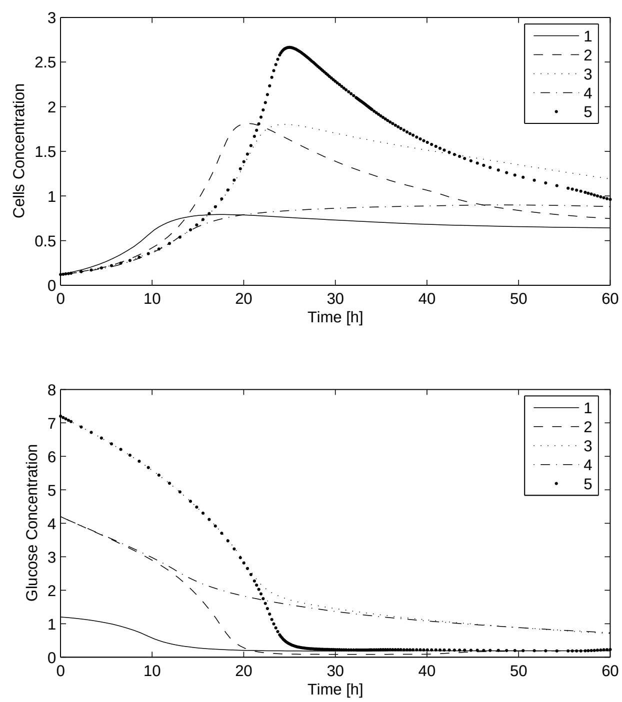

Figure 12: Transient response of Cells and Glucose Concentration for u=0.036 for different start-up conditions

from low levels of S_1 as well as S_2 is maintained in a middle value. In a low value, the curves 1 and 4 are shown. Steady state 2 results from middle value of S_1 and high value for S_2, as also S_2 middle and S_1 high, see curves 2 and 3. Finally, the steady state 3 results from high levels of S_1 and S_2, see curve 5. It can be noticed the capability of the CPWL model to show the three steady states. High concentration of glucose and glutamine in the initial conditions produces high concentration of lactate generating the cellular death, see Figure 6.

4.4 Probing Feed Controller

This section briefly describes the control technique developed in [7] for *E. coli* cultivations. The key idea is to exploit the characteristic saturation in the oxygen supply system that occurs at acetate formation, when the glucose uptake rate exceeds the glucose uptake rate critic. By superimposing a short pulse on the feed rate and evaluating the response in the dissolved oxygen signal. If a pulse response is visible, the feed rate is increased at the end of a pulse. When overfeeding is detected (no pulse response) the feed rate is decreased. This control has some optimal characteristics in the sense that it gives the maximum growth rate without overflow metabolism, thus minimizing the cultivation time. A simple flow diagram of the control algorithm is shown in Figure 7.

4.4.1 Feedback algorithm

A simple algorithm that can be interpreted as a proportional incremental controller is used to adjust the feed rate F. At each cycle of the algorithm a pulse is given and depending on the response, the feed rate is adjusted according to the flow. A reaction to a pulse is said to occur if the amplitude of the response exceeds a critic oxygen reactive during the pulse [7]. In this specific case, a proportional probing feed controller is used, which changes depending if it is in a pulse up or a pulse down. For the first case the increase in the feed F is decided by

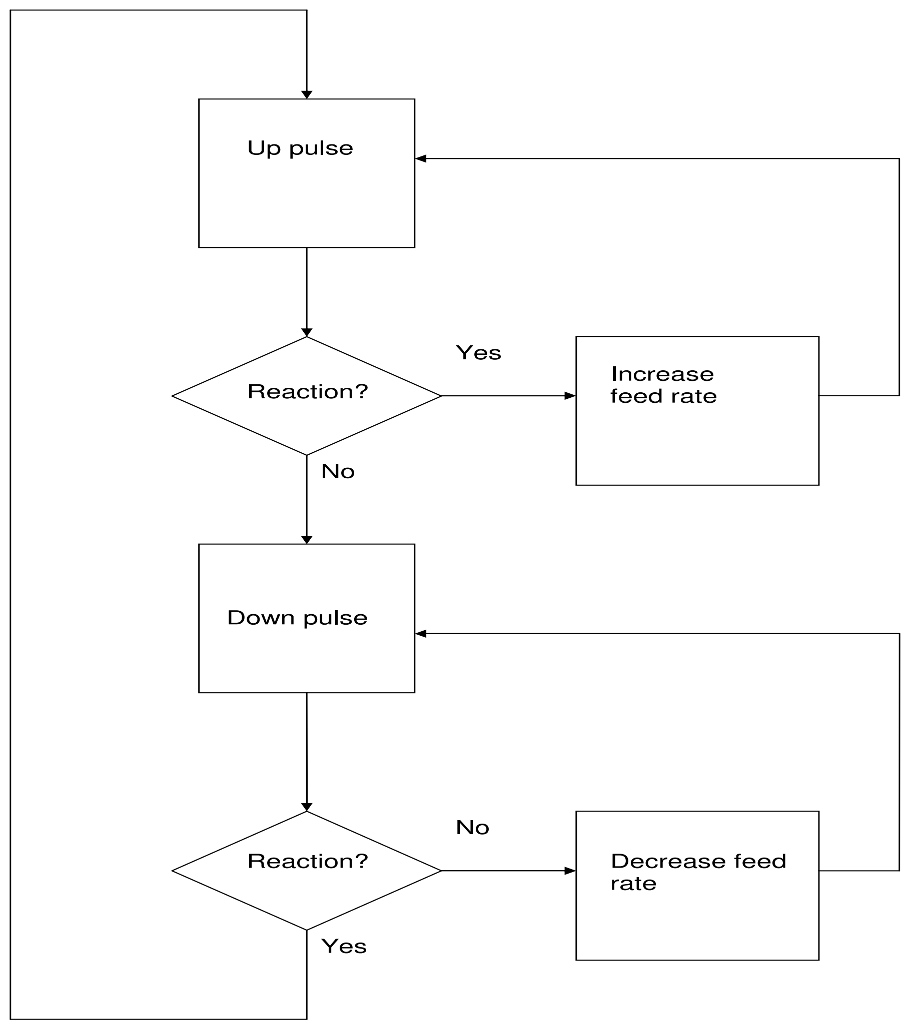

Figure 13: Flow diagram of the probing control algorithm ([7])

$$\frac{dF(t)}{dt} = k\frac{C_{pulse}(t) - C_{pref}}{C^* - x_4}F$$

where $C_{pulse}(t)$ is the pulse response, C_{pref} is the desired pulse response and k the controller gain. In the case for down pulses the feed is given by

$$\frac{dF(t)}{dt} = \left(k\frac{C_{pulse}(t) - C_{pref}}{C^* - x_4} - \gamma_p\right)F$$

where γ_p is the pulse amplitude. The pulse height, F_{pulse}, must give an oxygen response that exceeds the reaction levels, but it must also be ensured that the dissolved oxygen level does not reach zero during the pulse.

$$\begin{aligned} F_{pulse} &= \gamma_p F \\ \gamma_p &\approx 4C_{reac}/(C^* - C_{sp}) \end{aligned}$$

where C_{reac} is the reaction amplitude, it should be chosen large enough for the algorithm to be unaffected by the background variability in C_{pulse}, a value in $3 - 5\%$, is a reasonable default choice.

Figure 8 shows a simulation result with a proportional controller according to the description above. It is shown the oxygen concentration, the variable from the process, and the cell concentration. It shows the exponential cell growth by simulation.

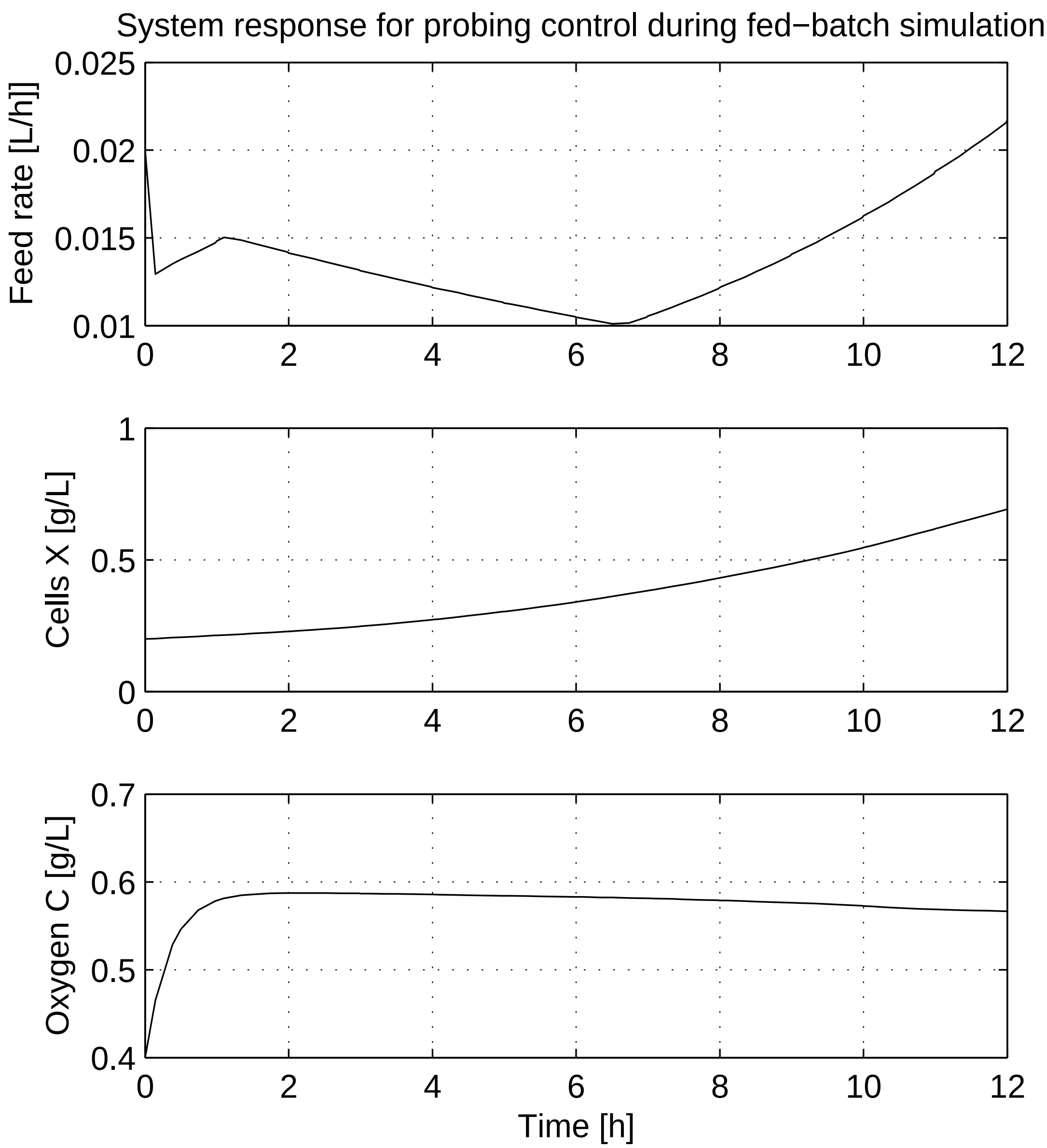

Figure 14: Systems response for probing control in a part of fed-batch cultivation

CHAPTER V

CONCLUSIONS AND FUTURE WORK

In this chapter, we overview the main contributions of this dissertations and discuss briefly future research directions.

5.1 Summary of Contributions

- In Chapter 2, we developed a new method for stability analysis of switched systems based on a polynomial approach. First, we transform the original problem into a polynomial system, which is able to mimic the switching behavior but with a continuous differential-algebraic nonlinear representation. From a theoretical point of view, we show that the representation of the original switched problem into a continuous polynomial system allows us to use the dissipation inequality for polynomial systems. With this method and from a theoretical point of view, we provide an alternative way to search for a common Lyapunov function for switched systems.

- In Chapter 3, we considered a new method for solving the optimal control problem of nonlinear switched systems based on a polynomial approach. First, we transform the original problem into a polynomial system, which is able to mimic the switching behavior with a continuous polynomial representation. After that, we transform the polynomial problem into a relaxed convex problem, using the method of moments. From a theoretical point of view, we provide necessary and sufficient conditions for the existence of minimizer by using particular features of the relaxed, convex formulation. Even in the absence of classical minimizers of the switched system, the solution of its relaxed formulation provides minimizers. However, in some cases some functions of the system are not in a polynomial form, to solve this issue, we apply the recasting process to obtain a complete polynomial system extended the results to a more general nonlinear switched systems.

- In Chapter 4, simulation results on a piecewise linear approximation based on orthonormal CPWL functions of nonlinear cellular growth have been presented. It has also been proved that this structure allows us to approximate the dynamical behavior for different initial conditions in the transient analysis. Also, the nonlinear characteristic presented over cellular growth about three steady states is shown for the CPWL model capturing the partially substitutable and partially complementary nature of the two substrates, glucose and glutamine. This model is useful to nonlinear control, in special to hybrid control. It is possible to take a nonlinear system, to obtain a PWL approximation from it, and then the hybrid control is applied.

  One interesting point over this CPWL model is the fact that some points of the nonlinear system model are used, and in a real implementation these points instead of to be taken from nonlinear systems, they could be taken from the sensors or data from the process.

5.2 Future Research Directions

In this section we outline several future research directions that are related to our work. Further directions for the optimal control can be focused on the development of a computational tool to solve the convex relaxed problem in general cases, i.e., nonlinear vector fields, and to prove the computational efficiency of the proposed method. Besides, an extension of this approach using SDP relaxation should be done for subsystems modeled by differential-algebraic equations [44], [91]. On the other hand, we have several tools for switched systems with this polynomial representation, different of the sum of squares decomposition. Some of them opens several possibilities for the system analysis, such as controllability and observability.

In the area of model and control of bioreactors using hybrid systems, future work is focused on to develop an optimal control of this class of hybrid system to implement it on cellular cultures based on the relaxed approach developed in Chapter 3. This model, although identified for the BHK experiments seems to be generally applicable to mammalian systems other than BHK, such as Hybridoma and Chinese Hamster Ovary (CHO) cells that display similar behavior [74]. It has been tested the CPWL model with the probing proportional controller based on the dissolved oxygen signal, its behavior exhibited some optimal characteristics, we can extend this result to a more general biological process.

APPENDIX A

MATHEMATICAL BACKGROUND

In this appendix, we present some of the mathematical ideas, concepts, and definitions that are employed in this thesis based mainly on the work of [18], [20], [28], [46], [47], [61], [85], [78], [73]. This appendix is a brief compilation of notions of measure theory and integration, probability theory, convex optimization, and optimization over polynomials using the theory of moments.

A.1 Brief Introduction to Measure Theory and Integration

Measure theory provides a way to extend our notions of length, area, volume, etc. to a much larger class of sets (but not all of its applications have to do with physical sizes). Informally, given some base set, a "measure" is any consistent assignment of "sizes" to the subsets of the base set. Depending on the application, the "size" of a subset may be interpreted as its physical size, the amount of something that lies within the subset, or the probability that some random process will yield a result within the subset. The main use of measures is to define general concepts of integration over domains with more complex structure than intervals of the real line. Such integrals are used extensively in probability theory, and in much of mathematical analysis. In order to be able to make use of measures and integrals, we need to know that the class of measurable sets is closed under certain types of operations.

A.1.1 Systems of Sets

Our universe is denoted by Ω, i.e., all the sets we shall consider are subsets of Ω. Recall some standard notation. 2^Ω everywhere denotes the set of all subsets of a given set Ω. The complement (in Ω) of a set A is denoted by A^c. By $A \triangle B$ the symmetric difference of A and B is denoted, i.e. $A \triangle B = (A \backslash B) \cap (B \backslash A)$.

Definition 34 *A ring of sets is a non-empty subset in* 2^Ω *which is closed with respect to the operations* $\cap$ *and* $\backslash$.

Definition 35 *A semi-ring is a collection of sets* $\mathcal{A} \subset 2^\Omega$ *with the following properties:*

1. *If* $A, B \in \mathcal{A}$ *then* $A \cap B \in \mathcal{A}$*;*
2. *For every* $A, B \in \mathcal{A}$ *there exists a finite disjoint collection* (C_j) $j = 1, 2, ..., n$ *of sets (i.e.*$C_i \cap C_j = \emptyset$ *if* $i \neq j$*) such that* $A \backslash B = \bigcup_{j=1}^{n} C_j$.

Example 36 *Let* $\Omega = \mathbb{R}$, *then the set of all semi-segments,* $[a, b)$, *forms a semi-ring.*

Definition 37 *An algebra (of sets) is a ring of sets containing* $\Omega \in 2^\Omega$.

Example 38 *Let* $\Omega = [a, b)$ *be a fixed interval on* $\mathbb{R}$. *Then the system of finite unions of subintervals* $[\alpha, \beta) \subset [a, b)$ *forms an algebra.*

Definition 39 *A* σ*-algebra is an algebra of sets which is closed with respect to all countable unions.*

In other words, a family $\mathcal{B}$ of subsets of Ω is a σ-algebra if:

(a) $\mathcal{B}$ contains Ω (or, $\mathcal{B}$ contains the empty set)

(b) $\mathcal{B}$ is closed under complements

(c) $\mathcal{B}$ is closed under countable unions

Definition 40 *A* σ*-algebra of sets,* $\mathcal{B}(\mathcal{U})$ *generated by a collection of sets* $\mathcal{U} \subset 2^\Omega$ *is the minimal* σ*-algebra of sets containing* $\mathcal{U}$.

In other words, it is the intersection of all σ-algebras of sets containing $\mathcal{U}$.

A.1.2 Measures

Formally, a measure is a function defined on a σ-algebra $\mathcal{B}$ over a set Ω such that the following properties are satisfied:

Definition 41 *Let Ω be a set, $\mathcal{B}$ an algebra on Ω. A function μ: $\Omega \to \mathbb{R}_+ \cup \{\infty\}$ is called a measure if*

1. *The empty set has measure zero. $\mu(A) \geq 0$ for any $A \in \mathcal{B}$ and $\mu(\emptyset) = 0$;*

2. *Countable additivity or σ-additivity. If $(A_i)_{i\geq 1}$ is a disjoint family of sets in $\mathcal{B}$ ($A_i \cap A_j = \emptyset$ for any $i \neq j$) such that $\bigcup_{i=1}^{\infty} A_i \in \mathcal{B}$, then*

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i).$$

A.1.2.1 Properties

Some elementary properties of a measure can be derived from the definition of a countably additive measure.

(a) If $A, B \in \mathcal{B}$ and $B \subset A$ then $\mu(B) \leq \mu(A)$.

(b) If $(A, B \in \mathcal{B}$ and $B \subset A$ and $\mu(B) < \infty$ then $\mu(A \backslash B) = \mu(A) - \mu(B)$.

(c) If $A, B \in \mathcal{B}$ and $\mu(A \cap B) < \infty$ then $\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$.

(d) If $(A_i)_{i\geq 1} \subset \mathcal{B}$ not necessarily disjoint, such that $\bigcup_{i=1}^{\infty} A_i \in \mathcal{B}$, then

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mu(A_i).$$

Theorem 42 *Let $\mathcal{B}$ be an algebra, $(A_i)_{i\geq 1} \subset \mathcal{B}$ a monotonically increasing sequence of sets ($A_i \subset A_{i+1}$) such that $\bigcup_{i\geq 1} \in \mathcal{B}$. Then*

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{n\to\infty} \mu(A_n).$$

Some important measures. Mainly we are interested in three important measures.

(a) The Lebesgue measure on $\mathbb{R}$ is the measure on a σ-algebra containing the intervals in $\mathbb{R}$ such that $\mu([0,1]) = 1$.

(b) The probability measure. $\mu(\Omega) = 1$, i.e. takes the value 1 on the whole space.

(c) The Dirac measure μ_A. The measure of a set is 1 if it contains the point A and 0 otherwise.

A.1.2.2 The Lebesgue Measure

Bounded Sets of $\mathbb{R}$. Let $\mathcal{B}$ be the algebra of all finite unions of semi-segments (semi-intervals) on $\mathbb{R}^1$, i.e. all sets of the form

$$A = \bigcup_{j=1}^{k} [a_j, b_j).$$

the Lebesgue measure μ is a mapping $\mu : \mathcal{B} \to \mathbb{R}$ defined by:

$$\mu(A) = \sum_{j=1}^{k} (b_j - a_j).$$

Definition 43 *The **Borel algebra** of sets, $\mathcal{B}$ on the real line is a σ-algebra generated by all open sets on $\mathbb{R}$. Any element of $\mathcal{B}$ is called a **Borel set**.*

In other words, The `Borel algebra` on $\mathbb{R}$ is the smallest σ-algebra containing all open sets in $\mathbb{R}$. It means that is the σ-algebra

$$\mathcal{B} = \cap \{\mathcal{C} : \mathcal{B}\text{is a } \sigma\text{-algebra, } U \text{ open} \Rightarrow U \in \mathcal{C}\}.$$

i.e. it is the intersection of the class of all σ-algebra that contain all the open sets in $\mathbb{R}$.

Theorem 44 *If $\mathcal{B}$ is the **Borel σ-algebra** on $\mathbb{R}$, there is a unique measure μ (called Lebesgue measure) defined on $\mathcal{B}$ such that $\mu((a,b)) = b - a$ provided that $b > a$. For every* `Borel set` *B,*

$$\mu(B) = \sup\{\mu(K) : K\,compact, K \subseteq B\} = \inf\{\mu(U) : U\,open, U \supseteq B\}$$

A.1.2.3 The Probability Measure

The probability measure is a function from $\mathcal{B}$ σ-algebra to the real numbers that assigns to each event a probability between 0 and 1. It must satisfy $\mu(\Omega) = 1$. Because the probability measure is a function on $\mathcal{B}$ and not on Ω, the set of events is not required to be the complete 2^{Ω} of the sample space; that is, not every set of outcomes is necessarily an event.

A.1.2.4 The Dirac Measure

A Dirac measure is a measure δ_x on a set Ω (with any σ-algebra of subsets of Ω) that gives the singleton set $\{x\}$ the measure 1, for a chosen element $x \in \Omega : \delta_x(\{x\}) = 1$.

In general, the measure is defined by

$$\delta_x(A) = \begin{cases} 0, & x \notin A \\ 1, & x \in A \end{cases}$$

for any measurable set $A \in \Omega$.

The Dirac measure is a probability measure, and in terms of probability it represents the almost sure outcome x in the sample space Ω. The Dirac measures are the extreme points (in a convex set S, is a point in S which does not lie in any open line segment joining two points of S, intuitively, is a "corner" of S) of the convex set of probability measure on Ω.

A.1.3 Integration

Let Ω be a set, $\mathcal{B}$ is a σ-algebra of subsets of Ω and μ is a measure on it.

Definition 45 *A triple $(\Omega, \mathcal{B}, \mu)$ is called a measure space.*

The most important example of a measure space is the Lebesgue measure space.

A.1.3.1 Lebesgue Integration

The integral of a function f between limits a and b can be interpreted as the area under the graph of f. This is easy to understand for familiar functions such as polynomials, but what does it mean for more exotic functions? In general, what is the class of functions for which "area under the curve" makes sense? The answer to this question has great theoretical and practical importance. As part of a general movement toward rigor in mathematics in the nineteenth century, attempts were made to put the integral calculus on a firm foundation. The Riemann integral, proposed by Bernhard Riemann (1826-1866), is a broadly successful attempt to provide such a foundation for the integral. Riemann's definition starts with the construction of a sequence of easily-calculated integrals which converge to the integral of a given function. This definition is successful in the sense that it gives the expected answer for many already-solved problems, and gives useful results for many other problems. However, Riemann integration does not interact well with taking limits of sequences of functions, making such limiting processes difficult to analyze. This is of prime importance, for instance, in the study of Fourier series, Fourier transforms and other topics. The Lebesgue integral is better able to describe how and when it is possible to take limits under the integral sign. The Lebesgue definition considers a different class of easily-calculated integrals than the Riemann definition, which is the main reason the Lebesgue integral is better behaved. The Lebesgue definition also makes it possible to calculate integrals for a broader class of functions.

In Lebesgue's theory, integrals are limited to a class of functions called measurable functions. A function f is measurable if the pre-image of every closed interval is in X:

$$f^{-1}([a, b]) \in X \text{ for all } a < b.$$

It can be shown that this is equivalent to requiring that the pre-image of any Borel subset of $\mathbb{R}$ be in X. We will make this assumption from now on. The set of measurable functions is closed under algebraic operations, but more importantly the class is closed under various kinds of pointwise sequential limits.

We build up an integral

$$\int_{\Omega} f d\mu = \int_{\Omega} f(x)\mu(dx)$$

for measurable real-valued functions f defined on Ω in stages.

The Lebesgue integral has the following properties:

Linearity: If f and g are Lebesgue integrable functions and a and b are real numbers, then $af + bg$ is Lebesgue integrable and

$$\int (af + bg)d\mu = a \int f d\mu + b \int g d\mu$$

Monotonicity: If $f \leq g$, then

$$\int f d\mu \leq \int g d\mu$$

Monotone convergence theorem: Suppose (f_n) is a sequence of real, non-negative measurable functions with limit f such that

$$f_n(x) \leq f_{n+1}(x) \; \forall k \in \mathbb{N}, \forall x \in \Omega$$

Then

$$\int_A f d\mu = \lim_{n \to \infty} \int_A f_n d\mu, A \in \Omega$$

Note: The value of any of the integrals is allowed to be infinite.

Fatou's lemma: If (f_n is a sequence of non-negative measurable functions defined a.e. and

$$f(x) = \lim_{n\to\infty} \inf f_n(x)$$

then

$$\int_A f d\mu \leq \lim_{n\to\infty} \inf \int_A f_n d\mu, A \in \Omega$$

Note: Almost everywhere (or in probability theory almost surely, abbreviated $a.s$), abbreviated $a.e.$, means that if the set of elements for which the property does not is a null set, i.e. is a set with measure zero, or in cases where the measure is not complete, contained within a set of measure zero.

Theorem 46 *Let $A \in \mathcal{B}$, (f_n) be a sequence of non-negative measurable functions and*

$$f(x) = \sum_{n=1}^{\infty} f_n(x), x \in A$$

Then

$$\int_A f d\mu = \sum_{n=1}^{\infty} \int_A f_n d\mu$$

Theorem 47 ***Lebesgue's dominated convergence theorem:*** *Let $A \in \mathcal{B}$, (f_n) be a sequence of measurable functions such that $f_n(x) \to f(x)(x \in A)$ Suppose there exists a function $g \in L^1(\mu)$, i.e., Lebesgue integrable function, on A such that*

$$|f_n(x)| \leq g(x)$$

Then

$$\lim_n \int_A f_n d\mu = \int_A f d\mu$$

Theorem 48 *Suppose that* $f : [a, b] \to \mathbb{R}$ *is Riemann integrable (in particular, this is the case if* f *is continuous). Then* f *is Lebesgue integrable and*

$$\int_{[a,b]} f_n d\mu = \int_a^b f(t)dt$$

In other words, the Lebesgue integral of f *is equals the Riemann integral of* f*.*

A.2 Some Results on Probability Theory

The definition of the probability space is the foundation of probability theory. It was introduced by Kolmogorov in the 1930s.

Definition 49 *A measure space* $(\Omega, \mathcal{B}, \mu)$ *is a probability space if* $\mu(\Omega)$ *is a probability measure on* $\mathcal{B}$*, i.e.*

(a) $\mu(\Omega) = 1$

(b) $\mu(A) \geq 0, A \in \mathcal{B}$

(c) μ *is countably additive:* $A_i \in \Omega, \forall i \geq 1, A_i \cap A_j = \emptyset, \forall i \neq j \Rightarrow \mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$

The sample space Ω, is a nonempty set whose elements are known as outcomes or states of nature and are often given the symbol ω. The set of all possible outcomes of an experiment is known as the sample space of the experiment.

$\mathcal{B}$ is a σ-algebra of subsets of Ω. Its elements are called events, which are sets of outcomes for which one can ask a probability. The complement of any event is an event, and the union of any (finite or countable infinite) sequence of events is an event. usually, the events are the Lebesgue-measurable or Borel-measurable sets of real numbers.

The probability measure μ is a function from $\mathcal{B}$ to the real numbers that assigns to each event a probability between 0 and 1. Because μ is a function defined on $\mathcal{B}$ and not on Ω, the set of events is not required to be the complete power set of the sample space; that is, not every set of outcomes is necessarily an event. Measure theory give us a way to deal simultaneously with continuous, discrete, and mixed distributions in a unified way.

A.2.0.2 Expectation

Definition 50 *If $f : \Omega \to \mathbb{R}$ is a random variable on $(\Omega, \mathcal{B}, \mu)$ then the mean or expectation of f is defined by*

$$E(f) = \int_{\Omega} f d\mu$$

It has all usual properties of the integrals. Recall, f is a measurable function.

The analogy with discrete values and the classical probability theory, the expectation is

$$E(X) = \sum_{j=1}^{k} x_j p_j$$

where X is a discrete variable that can take values $x_1, x_2, ..., x_k$ with probabilities $p_1, p_2, ..., p_k$.

A.2.1 Some Facts About Young Measures

This section is a summary of a few basic, general, important facts to be used to providing results for the Chapter 3. Our fundamental reference for this material is [78] and [73]. The fist one is a basic existence theorem for Young measures, which are parameterized measures that are associated with certain subsequences of a given bounded sequence of measurable functions. The second one relates to the fact that lack of oscillations is reflected on triviality of the Young measure.

A basic definition about normed spaces or L^p spaces, which are in fact vector spaces, since they are stable under the vector space operations.

Definition 51 *The space of sequences* l^p*,* $p \geq 1$*, consists of sequences* (x_n) *satisfying*

$$\sum_{n=1}^{\infty} |x_n|^p < \infty$$

the norm in l^p *is* $\|x_n\|_p = (\sum_{n=1}^{\infty} |x_n|^p)^{1/p}$*.*

Theorem 52 *Let* $\Omega \subset \mathbb{R}^n$ *be a measurable set and let* $u_j : \Omega \mapsto \mathbb{R}^m$ *be measurable functions such that*

$$\sup_j \int_\Omega \psi(|u_j|)dx < \infty,$$

where $\psi : [0, \infty) \mapsto [0, \infty]$ *is a continuous, nondecreasing function such that* $\lim_{t \mapsto \infty} \psi(t) = \infty$*. There exists a subsequence, not relabeled, and a family of probability measures,* $\nu = \{\nu_x\}_{x \in \Omega}$ *(the associated parametrized measure), depending measurably on* x*, with the property that whenever the sequence* $\{\psi(x, u_j(x))\}$ *is weakly convergent in* $L^1(\Omega)$ *for any Carathéodory function* $\psi(x, \lambda) : \Omega \times \mathbb{R}^m \mapsto \mathbb{R}^*$*, which implies that this function is measurable in* x*, continuous in* t*, and has a bounded Lebesgue integral, the weak limit is the (measurable) function*

$$\overline{\psi}(x) = \int_{\mathbb{R}^m} \psi(x, \lambda) d\nu_x(\lambda).$$

Notice that there is always a Young measure associated to a bounded sequence $\{u_j\}$ in $L^p(\Omega)$ for $p > 1$.

Proposition 53 *Let* $z_j = (x_j, u_j) : \Omega \mapsto \mathbb{R}^n \times \mathbb{R}^m$ *be a bounded sequence in* L^p *such that* $\{x_j\}$ *converges strongly to* x *in* L^p*. If* $\mu = \{\mu_t\}_{t \in \Omega}$ *is the parametrized measure associated to* $\{z_j\}$ *then* $\mu_t = \delta_{x(t)} \otimes \nu_t$ *a.e.* $t \in \Omega$*, where* $\{\nu_t\}_{t \in \Omega}$ *is the parametrized measure corresponding to* $\{u_j\}$*.*

The following fact is a remarkable, convenient property.

Lemma 54 *If* $\nu = \{\nu_x\}_{x \in \Omega}$ *is a family of probability measures supported in* $\mathbb{R}^m$ *such that*

$$\int_\Omega \int_{\mathbb{R}^m} |\lambda|^p d\nu_x(\lambda) dx < \infty, \quad p > 1,$$

then there exists a bounded sequence in $L^p(\Omega)$, $\{u_j\}$, whose corresponding Young measure is ν and such that $\{|u_j|^p\}$ is equiintegrable. If moreover $\text{SUPP}(\nu_t) \subset K$ *for a.e. $x \in \Omega$, where $K \subset \mathbb{R}^m$ is a convex set, then each u_j can be chosen taking values in K.*

one more lemma.

Lemma 55 *Let $g(x,u) : \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}$ be continuous in x and measurable in u such that*

$$\lim_{|u| \mapsto \infty} \frac{g(x,u)}{|u|^p} = 0, \quad p > 1,$$

for every fixed $x \in \mathbb{R}^n$. If $\{x_j\}$ and $\{u_j\}$ are bounded in $L^\infty(\Omega)$ and $L^p(\Omega)$, respectively, then $\{g(x_j,u_j)\}$ is equiintegrable in $L^1(\Omega)$.

A.2.2 The Problem of Moments

> Nous appellerons *problème des moments* le problème suivant: Trouver une distribution de masse positive sur une droite $(0,\infty)$, les moments d'ordre $k(k = 0,1,...)$ étant donnés.
>
> T.J. Stieltjes, 1894-memoir

Given a sequence $m_0, m_1, ...$ of real numbers. Find necessary and sufficient conditions for the existence of a measure μ on $[0,\infty)$ so that

$$m_n = \int_0^\infty x^n d\mu(x) \text{ for } n = 0,1,...$$

The number m_n is called the n'th moment of μ, and the sequence (m_n) is called the moment sequence of μ. Then the definition of the integral of a continuous function with respect to an increasing function is defined - the Stieltjes integral.

Results from the original Stieltjes moment problem have been presented, leading to three classical moment problems: the Hamburger moment problem in which the support of μ is allowed to be the whole real line; the Stieltjes moment problem, for $[0, \infty)$; and the Hausdorff moment problem for a bounded interval.

Existence It was realized that the problem of moments is closely connected to Hilbert spaces and spectral theory. In more concrete term, there is a condition on a positive measure μ, namely that

$$\int |P(x)|^2 \, d\mu(x) > 0$$

for every complex-valued polynomial $P(x)$, unless P vanishes on the support of μ. This gives rise to matrix conditions, necessary on any sequence of moments, namely that certain Hankel matrices are positive semi definite.

A general description of the problem of moments is as follows.

Given a set of functions $h_1, ..., h_k$ defined in $\Omega \subset \mathbb{R}^n$ and a sequence of values (m_k), the problem of moments consist of determining a positive measure μ such that

$$m_i = \int_\Omega h_i(x) d\mu(x) \; \forall i = 1, ..., k,$$

whenever it is possible. Thus, the problem of moments also includes the search for requirements in order to characterize the sequence (m_k) as a set of moments. Depending on the function basis $h_1, ..., h_k$ and the set Ω, the problem of moments can take different forms. Usually, we refer to the function basis as an algebraic systems, i.e. $h_i = x^i$.

A.3 Basics of Convex Optimization

In convex optimization we find a fusion of three different domains: Convex analysis, optimization, and numerical computation.

A.3.1 Convex Sets

We call a point of the form $\theta_1 x_1 + \cdots + \theta_k x_k$, where $\theta_1 + \cdots + \theta_k = 1$ and $\theta_i \geq 0$, $i = 1, \cdots, k$, a convex combinations of the points $x_1 + \cdots + x_k$. It can be shown that a set is convex if and only if it contains every convex combination of its points.

Definition 56 *The convex hull of a set C, denoted $\overline{co}C$, is the set of all convex combinations of points in C:*

$$\overline{co}C = \{\theta_1 + \cdots + \theta_k | x_i \in C, \theta_i \geq 0, i = 1, \cdots, k, \sum_k \theta_i = 1\}$$

The convex hull $\overline{co}C$ is always convex. It is the smallest convex set that contains C.

The idea of a convex combination can be generalized to include infinite sums, integrals, and, in the most general form, probability distributions.

Suppose $\theta_1 + \theta_k + \cdots$ satisfy

$$\theta_i \geq 0, i = 1, \cdots, k, \sum_{i=1}^{\infty} \theta_i = 1$$

and $x_1, x_2, ... \in C$, where $C \in \mathbb{R}^n$ is convex. Then $\sum_{i=1}^{\infty} \theta_i x_i \in C$, if the series converges.

More generally, suppose $p : \mathbb{R}^n \to \mathbb{R}$ satisfy $p(x) \geq 0$ for all $x \in C$ and $\int_C p(x)dx = 1$, where $C \subset \mathbb{R}^n$ is convex. Then

$$\int_C x p(x) dx \in C,$$

if the integral exists. In the most general form, suppose $C \subset \mathbb{R}^n$ is convex and x is a random vector with $x \in C$ with probability one. Then the expectation $E(x) \in C$. Indeed, this form includes all the others as special cases. For example, suppose the random variable x only takes on the two values x_1 and x_2, with $\textbf{prob}(x = x_1) = \theta$

and $\mathbf{prob}(x = x_1) = 1 - \theta$, where $0 \leq \theta \leq 1$. Then $E(x) = \theta x_1 + (1 - \theta)x_2$, and we are back to a simplex convex combination of two points.

A.3.2 Convex Functions

Definition 57 *A function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is convex if its domain $\mathbf{dom} f$ is convex and for all $x, y \in \mathbf{dom} f, \theta \in [0, 1]$*

$$f(\theta x + (1 - \theta)y) \leq \theta f(x) + (1 - \theta)f(y)$$

Definition 58 *The epigraph of a function f is*

$$\mathbf{epi} f = \{(x, a) \,| x \in \mathbf{dom} f, f(x) \leq a\}$$

From the basic definition of convexity, it follows that f is a convex function if and only if its epigraph $\mathbf{epi} f$ is a convex set.

The convexity of a differentiable function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ can also be characterized by conditions on its gradient ∇f and Hessian $\nabla^2 f$. Recall that, in general, the gradient yields a first order Taylor approximation at x_0.

The first-order condition: f is convex if and only if for all $x, x_0 \in \mathbf{dom} f$,

$$f(x) \geq f(x_0) + \nabla f(x_0)^T (x - x_0),$$

i.e., the first order approximation of f is a global underestimator.

Recall that the Hessian of f, $\nabla^2 f$, yields a second order Taylor series expansion around x_0. We have the necessary and sufficient second-order condition: a twice differentiable function f is convex if and only if for all $x \in \mathbf{dom} f, \nabla^2 f(x) \succeq 0$, i.e., its Hessian is positive semidefinite on its domain.

Elementary properties helpful in verifying convexity:

(a) nonnegative sums of convex functions are convex

(b) nonnegative infinite sums, integrals:

$$p(y) \geq 0, g(x,y) \text{ convex in } x \Rightarrow \int p(y)g(x,y)dy \text{ convex},$$

(c) expected value: $f(x,u)$ convex in $x \Rightarrow g(x) = E_u f(x,u)$ convex

Jensen's inequality: If $f : \mathbb{R}^n \to \mathbb{R}$ is convex

(a) $\sum_i \theta_i = 1, \theta_i \geq 0 \Rightarrow f(\sum_i \theta_i x_i) \leq \sum_i \theta_i f(x_i)$

(b) continuous version:$\int p(x)dx = 1, p(x) \geq 0 \Rightarrow f(\int xp(x)dx) \leq \int f(x)p(x)dx)$

(c) more generally, $f(E(x)) \leq E(f(x))$.

Another convexity preserving operation is that of minimizing over some variables. Specially, if $h(x,y)$ is convex in x and y, then

$$f(x) = \inf_y h(x,y)$$

is convex in x. This is because the operation above correspond to projection of the epigraph, $(x,y,a) \to (x,t)$.

Definition 59 *The convex hull or envelope of a function* $f : \mathbb{R}^n \to \mathbb{R}$ *is defined as*

$$h(x) = \overline{co} f(x) = \inf\{a \,|(x,a) \in \overline{co}\,\mathbf{epi} f(x)\}$$

Geometrically, the epigraph of $h(x)$ is the convex hull of the epigraph of f. $h(x)$ is the largest convex underestimator of f.

Definition 60 *(**Convex Relaxation**) Let* $f : S \Rightarrow \mathbb{R}$ *where* $S \subset \mathbb{R}^n$ *is a nonempty convex set. Then, a convex function* $h : S \Rightarrow \mathbb{R}$ *is a convex relaxation of* f *if* $h(x) \leq f(x) \forall x \in S$.

Definition 61 *(**Convex Envelope**) Let $f : S \Rightarrow \mathbb{R}$ where $S \subset \mathbb{R}^n$ is a nonempty convex set. The convex envelope of f over S (denoted f_S) is a convex relaxation such that for any other convex relaxation h of f on S, we have*

$$f_S \geq h(x) \forall x \in S.$$

The convex envelope is the tightest possible convex relaxation of a nonconvex function. The convex envelopes of many functions are known. However, in general, finding the convex envelope of an arbitrary function is as hard as finding the global minimum. On the other hand, a number of polynomial algorithms exist for constructing convex relaxations of quite general classes of functions.

A.3.3 Convex Optimization

Without more ado. The most important theorem of convex optimization.

Theorem 62 *Let S be a nonempty convex set in $\mathbb{R}^n$ and let $f : S \Rightarrow \mathbb{R}$ be convex on S. Consider the problem to minimize $f(x)$ subject to $x \in S$. Suppose $x^* \in S$ is a local minimum. Then x^* is a global minimum.*

In other words, for convex optimization (minimization of a convex function on a convex set):

$$\text{local minimum} \Longrightarrow \text{global minimum.}$$

In this way, we know that convex optimization problems have three crucial properties that makes them fundamentally more tractable than generic nonconvex optimization problems:

(a) no local minima: any local optimum is necessarily a global optimum;

(b) exact infeasibility detection: using duality theory, hence algorithms are easy to initialize

(c) efficient numerical solution methods.

There are several canonical optimization problem formulations, for which extremely efficient solution codes are available. Thus, if a real problem can be cast into one of these forms, then it can be considered as essentially solved.

The most important category of these canonical problems is known as Conic programming, it is called conic because the inequalities are specified in terms of affine functions and generalized inequalities. Geometrically, the inequalities are feasible if the range of the affine mapping intersects the cone of the inequality.

Linear Program

A general linear program (LP) has the form

$$\begin{aligned} \text{minimize} \quad & c^T x + d \\ \text{subject to} \quad & Gx \preceq h \\ & Ax = b, \end{aligned}$$

where $G \in \mathbb{R}^{m \times n}$ and $A \in \mathbb{R}^{p \times n}$

Second-order Cone Program

A problem that subsumes both linear and quadratic programming is the second-order cone program (SOCP):

$$\begin{aligned} \text{minimize} \quad & f^T x \\ \text{subject to} \quad & \|A_i x + b_i\|_2 \preceq c_i^T x + d_i, i = 1, ..., m \\ & Fx = g, \end{aligned}$$

where $x \in \mathbb{R}^n$ is the optimization variable, $A_i \in \mathbb{R}^{n_i \times n}$, and $F \in \mathbb{R}^{p \times n}$.

A problem which subsumes linear, quadratic and second-order cone programming is called semidefinite program (SDP), and has the form

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & x_1 F_1 + \cdots + x_n F_n + G \preceq 0 \\ & Ax = b, \end{array}$$

where $G, F_i \in \mathbb{S}^k$, and $A \in \mathbb{R}^{p \times n}$. The inequality here is a linear matrix inequality. As shown earlier, since SOCP constraints can be written as LMI's, SDP's subsume SOCP's, and hence LP's as well. (If there are multiple LMI's they can be stacked into one large block diagonal LMI, where the blocks are the individual LMIs).

A.4 Optimization over Polynomials Using The Method of Moments

The method of moments is a general method for treating non convex optimization problems. It takes a proper formulation in probability measures of a non convex optimization problem. In this way, when the problem can be stated in terms of polynomial expressions, we can transform the measures into algebraic moments to obtain a new convex program defined in a new set of variables that present the moments of every measure.

In global optimization the main objective is to find the global minima of a function f defined on a subset Ω of the Euclidean space $\mathbb{R}^n$. In other words, we are interested in solving a mathematical programs given in the general form

$$\min_{x \in \Omega} f(x) \tag{66}$$

where the objective function $f(x)$ is a linear combination of simple functions.

One approach to this problem comes from convex analysis, since we can use the convex envelope of the function f in order to locate its global minima. As we have shown, every convex combination of points in Ω can be described as a discrete probability distribution μ supported in Ω such that every integral

$$\int_{\Omega} f(x)d\mu(x)$$

represent one point over the convex envelope of the function f. For this reason, we study the relaxed problem

$$\min_{\mu} \int_{\Omega} f(x)d\mu(x) \tag{67}$$

in order to find the global minima of the objective function f in Ω. The relaxed problem (67) contains information about all the global minima of the function f in Ω. However, it cannot be solved easily in practice: consider for instance, the difficulty of describing all possible convex combinations of points in Ω. It can be shown how transform problem (67) in order to make it more treatable. Since $f(x)$ is a polynomial of degree, suppose r, the criterion $\int f d\mu$ involves only the moments of μ up to order m and, in addition, is linear in the moment variables. Therefore, one next replaces μ with the finite sequence $m = \{m_\alpha\}$ of all its moments, up to order r, that is

$$\min_{x\in\mathbb{R}} f(x) \rightarrow \min_{\mu} \int_{\Omega} f(x)d\mu(x)$$

$$\begin{aligned}
\min_{\mu} \int_{\Omega} f(x)d\mu(x) &= \min_{\mu} \int_{\Omega} \left(\sum_{i=0}^{2n} c_i x^i \right) d\mu(x) \\
&= \min_{\mu} \sum_{i=0}^{2n} c_i \underbrace{\int_{\Omega} x^i d\mu(x)}_{\text{moments}} \\
&= \min_{m_i} \sum_{i=0}^{2n} c_i m_i
\end{aligned} \tag{68}$$

and one works with the finite sequence m of the moments of μ, up to order r, instead of μ itself. Of course, not every sequence m has a representing measure μ; that is, given an arbitrary finite sequence m, there might not be any probability measure μ, all of whose moments up to order r coincide with the m_α scalars. Consider the function polynomial f and hence it can be expressed as a linear combination of simple functions, in this case the simple functions are the algebraic system of integer exponents. Let $\mathbb{R}^n[x] = \mathbb{R}[x_1, ..., x_n]$ denote the space of real polynomials in n real variables, and let $\mathbb{R}^n_r[x]$ denote the polynomials of degree at most r, then $d(r) = \dim \mathbb{R}^n_r[x] = \binom{n+r}{r}$ is its dimension. For $x = (x_1, ..., x_n) \in \mathbb{R}^n$, a monomial can then be associated to a string $\alpha = (\alpha_1, ..., \alpha n)$ of integers $\alpha_i \in \{1, ..., n\}$. Then, an r-degree polynomial $p(x) : \mathbb{R}^n \mapsto \mathbb{R}^n$ can be expressed as

$$p(x) = \sum_{|\alpha| \leq r} p_\alpha x^\alpha$$

where

$$x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$$

is a monomial of degree r and p_α is a coefficient. A polynomial p of degree r can thus be identified with the sequence of its coefficients $(p_\alpha)_{|\alpha| \leq r}$ in the canonical basis of monomial $\{x_\alpha : \ |\alpha| \leq r\}$, and the space of polynomial of degree r can be viewed as a vector space, which we denote $\mathcal{P}_r$. If necessary, a polynomial of degree r can be viewed as a polynomial of higher degree r' by setting the coefficients of monomials of degree higher than r to zero. Given two index string $\alpha = (\alpha_1, ..., \alpha_s)$ and $\beta = (\beta_1, ..., \beta_t)$, we define their concatenation as $\alpha \circ \beta = (\alpha_1, ..., \alpha_s, \beta_1, ..., \beta_t)$. It thus follows that $x_\alpha x_\beta = x_{\alpha \circ \beta}$.

In order to study moment problems on a semialgebraic set of $\mathbb{R}^n$, we consider the following notions of the truncated moment problem. Given a real sequence $m = \{m_\alpha\}_{\alpha \in \mathbb{Z}^n, |\alpha| \leq N}$, the truncated moment problem for m concern conditions for the existence of positive Borel measure μ on $\mathbb{R}^n$ satisfying

$$m_\alpha = \int x^\alpha d\mu(x) (\equiv \int x_1^{\alpha_1} \cdots x_n^{\alpha_n} d\mu(x_1, ..., x_n)) \ (|\alpha| \leq N). \tag{69}$$

A measure μ satisfying (69) is a representing measure for m; if, in addition, $\mathbf{K} \subseteq \mathbb{R}^n$ is closed and $\text{supp}\mu \subseteq \mathbf{K}$, then μ is a $\mathbf{K}$-representing measure for m. we next introduce the definitions of moment matrix and localizing matrix. Let $N = 2r$, in this case m corresponds to a real moment matrix $M_r = M_r^n(m)$ defined as follows.

Definition 63 ***Moment matrix:*** *Let $\mathcal{P}_r$ denote the basis of monomials in $\mathbb{R}^n_r[x]$, ordered lexicographically, e.g., for $n = 3$, $r = 2$, this ordering is 1, x_1, x_2, x_3, x_1^2, x_1x_2, x_1x_3, x_2^2, x_2x_3, x_3^2. The size of M_r is $\dim \mathbb{R}^n_r[x] = \binom{r+n}{n}$, with rows and columns indexed as $\{T^i\}_{i\in\mathbb{Z}^n,|i|\le r}$, following the same lexicographic order as above. The entry of M_r in row T^i, column T^j is $m_{i+j}, i, j \in \mathbb{Z}^n, |i| + |j| \le 2r$. Note that for $n = 1$, $M^n_r(m)$ is the Hankel matrix (m_{i+j}) associated with the classical Hamburger moment problem ($\mathbf{K} = \mathbb{R}$). Another way of constructing $M_r(m)$ is as follows. For a given real sequence $m = \{m_\alpha\}_{\alpha\in\mathbb{N}^n\times\mathbb{N}^q}$ of real numbers, the moment matrix $M_r(m)$ of order r associated with m, has its rows and columns indexed in the canonical basis $\{x^\alpha\}$, and is defined by*

$$M_r(m)(\alpha,\gamma) = m_{\alpha+\gamma},$$

$$\gamma, \alpha \in \mathbb{N}^n,\ |\gamma|, |\alpha| \le r, \tag{70}$$

where $|\alpha| := \sum_j \alpha_j$.

For illustration, consider the two-dimensional case. The moment matrix $M_r(m)$ is the block matrix $\{M_{i,j}(m)\}_{0\le i,j\le 2r}$ defined by

$$M_{i,j}(m) = \begin{bmatrix} m_{i+j,0} & m_{i+j-1,1} & \cdots & m_{i,j} \\ m_{i+j-1,1} & m_{i+j-2,2} & \cdots & m_{i-1,j+1} \\ \cdots & \cdots & \cdots & \cdots \\ m_{j,i} & m_{i+j-1,1} & \cdots & m_{0,i+j} \end{bmatrix},$$

where $m_{i,j}$ represents the $(i+j)$-order moment $\int x^i y^j d\mu(x,y)$ for some probability measure μ. For instance, consider the particular case when $n = 2$ and $r = 2$, one obtains

$$M_2(m) = \begin{bmatrix} 1 & | & m_{1,0} & m_{0,1} & | & m_{2,0} & m_{1,1} & m_{0,2} \\ & - & - & - & - & - & - & - \\ m_{1,0} & | & m_{2,0} & m_{1,1} & | & m_{3,0} & m_{2,1} & m_{1,2} \\ m_{0,1} & | & m_{1,1} & m_{0,2} & | & m_{2,1} & m_{1,2} & m_{0,3} \\ & - & - & - & - & - & - & - \\ m_{2,0} & | & m_{3,0} & m_{2,1} & | & m_{4,0} & m_{3,1} & m_{2,2} \\ m_{1,1} & | & m_{2,1} & m_{1,2} & | & m_{3,1} & m_{2,2} & m_{1,3} \\ m_{0,2} & | & m_{1,2} & m_{0,3} & | & m_{2,2} & m_{1,3} & m_{0,4} \end{bmatrix},$$

We next define the localizing matrix $M_r(\theta m)$ whose positivity is directly related to the existence of a representing measure for m with support in $\mathbb{K} = \{x \in \mathbb{R}[x] : \ \theta(x) \geq 0\}$.

Definition 64 ***Localizing matrix:*** *For a given polynomial $\theta \in \mathbb{R}[x]$, written as*

$$\theta(x) = \sum_{\beta} \theta_\beta x^\beta,$$

We define the localizing matrix $M_r(\theta m)$ associated with m, θ, and with rows and columns also indexed in the canonical basis of $\mathbb{R}[x]$, by

$$M_r(\theta m)(\gamma) = \sum_{\beta} \theta_\beta m_{(\theta)+\gamma},$$

$$\gamma \in \mathbb{N}^n, \ \ |\gamma| \leq r. \tag{71}$$

Therefore, $M_r(\theta m) \succeq 0$ whenever μ_m has its support contained in the set $\mathbb{K}$.

The $\mathbb{K}$-moment problem identifies those sequences m that are moments-sequences of a measure with support contained in the semialgebraic set $\mathbb{K}$.

To illustrate how construct a localizing matrix, consider a moment matrix

$$M_1(m) = \begin{bmatrix} 1 & m_{1,0} & m_{0,1} \\ m_{1,0} & m_{2,0} & m_{1,1} \\ m_{0,1} & m_{1,1} & m_{0,2} \end{bmatrix},$$

and a polynomial $\theta(x) = a - x_1^2 - x_2^2$, then we obtain a localizing matrix as follows,

$$M_1(m) = \begin{bmatrix} a - m_{2,0} - m_{0,2} & am_{1,0} - m_{3,0} - m_{1,2} & am_{0,1} - m_{2,1} - m_{0,3} \\ am_{1,0} - m_{3,0} - m_{1,2} & am_{2,0} - m_{4,0} - m_{2,2} & am_{1,1} - m_{3,1} - m_{1,3} \\ am_{0,1} - m_{2,1} - m_{0,3} & am_{1,1} - m_{3,1} - m_{1,3} & am_{0,2} - m_{2,2} - m_{0,4} \end{bmatrix},$$

We briefly outline the idea developed for the optimization in polynomials with Ω as an arbitrary subset of $\mathbb{R}^n$, one first reduces the optimization polynomial problem to the equivalent convex optimization (67) on the space of probability measures μ with support contained in Ω. We have the following proposition summarizing the main result for constrained optimization in polynomials.

Proposition 65 *The problems (66) and (67) are equivalent, that is,*

(a) $\inf(66) = \inf(67)$.

(b) If x^* is a global minimizer of (66), then $\mu^* := \delta_{x^*}$ is a global minimizer of (67).

(c) Assuming (66) has a global minimizer, then, for every optimal solution μ^* of (67), $f(x) = \min(66)$, μ^*-almost everywhere ($\mu^* - a.e$).

(d) If x^* is the unique global minimizer of (66), then $\mu^* := \delta_{x^*}$ is the unique global minimizer of (67).

A.4.1 Convergent Semi-definite Relaxations

Consider the constrained optimization problem in (66) where the $g_i(x)$ are all real-valued polynomials, $i = 1, ..., m$. Let

$$\Omega = \{x \in \mathbb{R}^n \mid g_i(x) \geq 0,\ i = 1, ..., m\}$$

be the feasible set. Depending on its parity, let $w_k = 2v_k$ or $w_k = 2v_k - 1$ be the degree of the polynomial $g_k(x)$, $k = 1, ..., m$. When needed below, for $i \geq \max_k w_k$, the vector $g_k \in \mathbb{R}^{d(w_k)}$ are extended to vectors of $\mathbb{R}^{d(w_k)}$ by completing them with zeros.

For $i \geq \max[\deg(f), \max_i \deg(g_i)]$, consider the positive semidefinite program

$$SDP_i \begin{cases} f^i = \min_m & \sum_\alpha f_\alpha m_\alpha \\ s.t & M_i(m) \succeq 0, \\ & M_{i-d_i}(g_k m) \succeq 0, \ \forall k = 1, ..., m \end{cases}$$

where $d_i = \lceil \deg(g_k)/2 \rceil$. Note that the optimum f^i is a lower-bound on the global optimum f^* of the original problem (66), since any feasible solution x of (66) yields a feasible solution m of SDP_i through Eq. (68). Moreover, $f^i \leq f^{i'}$ when $i \geq i'$. We refer to problem SDP_i as the semidefinite program relaxation of order i of (66). If any feasible point of the relaxation of order i is bounded, then $f^i \to f^*$ as $i \to \infty$. The LMI constrains of SDP_i state necessary conditions for m to be the vector of moments up to order $2i$, of some probability measure μ_m with support contained in Ω. This clearly implies that $\inf SDP_i \leq f^*$, as the vector of moments of the Dirac measure at a feasible point of (66) is feasible for SDP_i. It has been established that there are no more than $2^n - 1$ variables. We now can state an important results in the following theorem.

Theorem 66 *Consider the problem defined in (66) and let $v = \max_{k=1,\ldots,m} v_k$. Then for every $i \geq n + v$*

(a) SDP_i is solvable with $f^* = \min SDP_i$, and to every optimal solution x^* of (66) corresponds the optimal solution

$$m^* = (x_1^*, ..., x_n^*, ..., (x_1^*)^{2i}, ..., (x_n^*)^{2i}) \tag{72}$$

of SDP_i;

(b) every optimal solution m^* of SDP_i is the (finite) vector of moments of a probability measure finitely supported on t optimal solutions of (66), with $t = rankM_i(m) = rankM_n(m)$.

In many cases, low order relaxations (that is with $i \ll n$) will provide the optimal value f^*. Therefore, one would like to have a test to detect whether some relaxation SDP_i achieves the optimal value f^*. One way is to determine by inspection whether an optimal solution m of SDP_i is a moment vector. This will be the case if, for instance, $rank M_r(m) = 1$. However, in the case in which (66) has multiple optimal solutions, it can happen that m is a convex combination of moments of Dirac measures supported on the optimal solutions, which in general is not easy to detect. We next provide a criterion to test whether the SDP relaxation SDP_i indeed achieves the optimal value f^*.

Theorem 67 *Consider the problem defined in (66) and let* $v = \max_{k=1,\ldots,m} v_k$. *Let* m^* *be an optimal solution of* SDP_i *with* $i < n + v$. *If*

$$rank M_{i-v+1}(m^*) = rank M_{i-v}(m^*),$$

then $\min SDP_i = f^*$ *and* m^* *is the vector of moments of a probability measure supported on* $t = rank M_i(m^*) = rank M_{i-v}(m^*)$ *optimal solutions of (66).*

REFERENCES

[1] "Special issue on hybrid systems and applications," *Nonlinear analysis*, vol. 62, no. 8, 2005.

[2] "Special issue on hybrid systems and applications (2)," *Nonlinear analysis*, vol. 63, no. 3, 2005.

[3] "Special issue on hybrid systems and applications (3)," *Nonlinear analysis*, vol. 64, no. 2, 2005.

[4] "Special issue on hybrid systems and applications (6)," *Nonlinear analysis*, vol. 65, no. 9, 2006.

[5] "Special issue on hybrid systems and applications (7)," *Nonlinear analysis*, vol. 65, no. 11, 2006.

[6] "Special issue on hybrid systems and applications (8)," *Nonlinear analysis*, vol. 65, no. 12, 2006.

[7] Åkesson, M., Hagander, P., and Axelsson, J., "Probing control of fed-batch cultures: Analysis and tuning," *Control Engineering Practice*, vol. 9, no. 7, pp. 709–723, 2001.

[8] Alamir, M. and Attia, S., "An efficient algorithm to solve optimal control problems for nonlinear switched hybrid systems," *The 6th IFAC symposium, NOLCOS*, 2004.

[9] Antsaklis, P., "Special issue on hybrid systems: theory and applications a brief introduction to the theory and applications of hybrid systems," *Proceedings of the IEEE*, vol. 88, no. 7, pp. 879–887, 2000.

[10] Åström, K., Aracil, J., and Gordillo, F., "A family of smooth controllers for swinging up a pendulum," *Automatica*, vol. 44, no. 7, pp. 1841–1848, 2008.

[11] Attia, S., Alamir, M., and de Wit, C., "Sub optimal control of switched nonlinear systems under location and switching constraints," *16th IFAC World Congress*, 2005.

[12] Axelsson, H., Boccadoro, M., Wardi, Y., and Egerstedt, M., "Optimal mode-switching for hybrid systems with unknown initial state,"

[13] Azuma, S., Imura, J., and Sugie, T., "Lebesgue piecewise affine approximation of nonlinear systems and its application to hybrid system modeling of biosystems," *Proc. of the 45th IEEE Conference on Decision and Control*, pp. 2128–2133, 2006.

[14] Bastin, G. and Van Impe, J., "Nonlinear and adaptive control in biotechnology: a tutorial," *European Journal of Control*, vol. 1, no. 1, pp. 37–53, 1995.

[15] Bemporad, A., "Efficient conversion of mixed logical dynamical systems into an equivalent piecewise affine form," *IEEE Trans. on Automatic Control*, vol. 49, no. 5, pp. 832–838, 2004.

[16] Bemporad, A. and Morari, M., "Control of systems integrating logic, dynamics, and constraints," *Automatica*, vol. 35, no. 3, pp. 407–427, 1999.

[17] Bemporad, A., Morari, M., Dua, V., and Pistikopoulos, E., "The explicit solution of model predictive control viamultiparametric quadratic programming," *American Control Conference, 2000. Proceedings of the 2000*, vol. 2, 2000.

[18] Ben-Tal, A. and Nemirovsky, A., *Lectures on Modern Convex Optimization*. SIAM, 2001.

[19] Bengea, S. C. and DeCarlo, R. A., "Optimal control of switching systems," *Automatica*, vol. 41, pp. 11–27, 2005.

[20] Berg, C., "Moment problems and polynomial approximation," *Ann. Fac. Sci. Toulouse, Stieltjes special*, pp. 9–32, 1996.

[21] Blondel, V., Gevers, M., and Lindquist, A., "Survey on the state of systems and control," *European Journal of Control*.

[22] Boccadoro, M., Egerstedt, M., Valigi, P., and Wardi, Y., "Beyond the construction of optimal switching surfaces for autonomous hybrid systems,"

[23] Branicky, M., Borkar, V., and Mitter, S., "A unified framework for hybrid control: model and optimal control theory," *Automatic Control, IEEE Transactions on*, vol. 43, no. 1, pp. 31–45, 1998.

[24] Branicky, M. and Mitter, S., "Algorithms for optimal hybrid control," *Decision and Control, 1995., Proceedings of the 34th IEEE Conference on*, vol. 3, 1995.

[25] Burden, R. and Faires, J. D., *Numerical Analysis*. Boston: PWS, 1985.

[26] Castro, L., Agamennoni, O., and D'Attellis, C., "Wiener-like modelling: a different approach," *Proc. of the 9th Mediterranean Conference on Control and Automation*, june 2001.

[27] Colaneri, P., Geromel, J., and Astolfi, A., "Stabilization of continuous-time switched systems," *Systems & Control Letters*, vol. 57, pp. 95–103, 2008.

[28] Curto, R. and Fialkow, L., "The truncated complex K-moment problem," *Trans. American Mathematical Society*, vol. 352, no. 6, pp. 2825–2856, 2000.

[29] Das, T. and Mukherjee, R., "Optimally switched linear systems," *Automatica*, vol. 44, no. 5, pp. 1437–1441, 2008.

[30] Dayawansa, W. and Martin, C., "A converse Lyapunov theorem for a class of dynamical systems which undergo switching," *IEEE Transactions on Automatic Control*, vol. 44, no. 4, pp. 751–760, 1999.

[31] Ebenbauer, C. and Allgöwer, F., "Analysis and design of polynomial control systems using dissipation inequalities and sum of squares," *Computers and Chemical Engineering*, vol. 30, pp. 1590–1602, 2006.

[32] Ebenbauer, C. and Allgöwer, F., "Stability analysis of constrained control systems: An alternative approach," *Systems & Control Letters*, vol. 56, pp. 93–98, 2007.

[33] Egerstedt, M., Wardi, Y., and Delmotte, F., "Optimal control of switching times in switched dynamical systems," *Decision and Control, 2003. Proceedings. 42nd IEEE Conference on*, vol. 3, 2003.

[34] Fotiou, I., Parrilo, P., and Morari, M., "Nonlinear parametric optimization using cylindrical algebraic decomposition," *Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC'05. 44th IEEE Conference on*, pp. 3735–3740, 2005.

[35] Fotiou, I., Rostalski, P., Parrilo, P., and Morari, M., "Parametric optimization and optimal control using algebraic geometry methods," *International Journal of Control*, vol. 79, no. 11, pp. 1340–1358, 2006.

[36] Galbraith, G. and Vinter, R., "Optimal control of hybrid systems with an infinite set of discrete states," *Journal of Dynamical and Control Systems*, vol. 9, no. 4, pp. 563–584, 2003.

[37] Garavello, M. and Piccoli, B., "Hybrid necessary principle," *44th IEEE Conference on Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC'05*, pp. 723–728, 2005.

[38] Girard, A., "Approximative solutions of odes using piecewise linear vector fields," *Proc. of the 5th International Workshop on Computer Algebra in Scientific Computing*, pp. 107–120, 2002.

[39] Hassibi, A. and Boyd, S., "Quadratic stabilization and control of piecewise-linear systems," *Proc. of the American Control Conference*, pp. 3659–3664, 1998.

[40] Henrion, D., Lasserre, J., and Savorgnan, C., "Nonlinear optimal control synthesis via occupation measures," *Proceedings of the 47th IEEE Conference on Decision and Control*, pp. 4749–4754, 2008.

[41] Julián, P., Desages, A., and Agamennoni, O., "High-level canonical piecewise linear representation using a simplicial partition," *IEEE Trans. on Circuits Syst. I*, vol. 46, pp. 463–480, 1999.

[42] Julián, P., Desages, A., and D'amico, B., "Orthonormal high level canonical pwl functions with applications to model reduction," *IEEE Trans. on Circuits Syst. I*, vol. 47, pp. 702–712, 2000.

[43] Kolker, Y., "A piecewise-linear growth model: comparison with competing forms in batch culture," *Journal of Mathematical Biology*, vol. 25, no. 5, pp. 543–551, 1987.

[44] Kunkel, P. and Mehrmann, V., "Stability properties of differential-algebraic equations and spin-stabilized descretizations," *Electronic Transactions on Numerical Analysis*, vol. 26, pp. 385–420, 2007.

[45] Lakshmikantham, V., "Special issue on hybrid systems and applications (4)," *Nonlinear analysis*, vol. 65, no. 5, 2005.

[46] Lasserre, J., "Global optimization with polynomials and the problem of moments," *SIAM J. Optimization*, vol. 11, pp. 796–817, 2001.

[47] Lasserre, J. B., "An explicit equivalent positive semidefinite program for nonlinear 0-1 programs," *SIAM Journal on Optimization*, vol. 12, pp. 756–769, 2002.

[48] Lasserre, J., "Convexity in semi-algebraic geometry and polynomial optimization," *SIAM J. Optim.*, vol. 19.

[49] Lasserre, J., "A semidefinite programming approach to the generalized problem of moments," *Mathematical Programming*, vol. 112, no. 1, pp. 65–92, 2008.

[50] Lasserre, J., Henrion, D., Prieur, C., and Trélat, E., "Nonlinear optimal control via occupation measures and LMI-relaxations," *SIAM Journal on Control and Optimization*, vol. 47, no. 4, pp. 1643–1666, 2008.

[51] Lasserre, J., Prieur, C., and Henrion, D., "Nonlinear optimal control: Numerical approximations via moments and LMI-relaxations," in *IEEE CONFERENCE ON DECISION AND CONTROL*, vol. 44, pp. 1648–1653, 2005.

[52] Li, R., Teo, K., Wong, K., and Duan, G., "Control parameterization enhancing transform for optimal control of switched systems," *Mathematical and Computer Modelling*, vol. 43, no. 11-12, pp. 1393–1403, 2006.

[53] Li, Z., Qiao, Y., Qi, H., and Cheng, D., "Stability of Switched Polynomial Systems," *Journal of Systems Science and Complexity*, vol. 21, no. 3, pp. 362–377, 2008.

[54] Liberzon, D., *Switching in systems and Control*. Boston: Birkhauser, 2003.

[55] Lin, H. and Antsaklis, P. J., "Stability and Stabilizability of Switched Linear Systems: A Short Survey of Recent Results," *Proceedings of the 2005 IEEE International Symposium on, Mediterrean Conference on Control and Automation Intelligent Control, 2005*, pp. 24–29, 2005.

[56] Lin, H. and Antsaklis, P., "Stability and Stabilizability of Switched Linear Systems: A Survey of Recent Results," *IEEE Transactions on Automatic Control*, vol. 54, no. 2, pp. 308–322, 2009.

[57] Lincoln, B. and Rantzer, A., "Relaxing dynamic programming," *IEEE Transactions on Automatic Control*, vol. 51, no. 8, pp. 1249–1260, 2006.

[58] Mancilla-Aguilar, J. and Garcia, R., "A converse Lyapunov theorem for nonlinear switched systems," *Systems & control letters*, vol. 41, no. 1, pp. 67–71, 2000.

[59] Margaliot, M., "Stability analysis of switched systems using variational principles: An introduction," *Automatica*, vol. 42, no. 12, pp. 2059–2077, 2006.

[60] Margaliot, M. and Liberzon, D., "Lie-algebraic stability conditions for nonlinear switched systems and differential inclusions," *Systems & Control Letters*, vol. 55, no. 1, pp. 8–16, 2006.

[61] Meziat, R., "The method of moments in global optimization," *Journal of mathematics science*, vol. 116, pp. 3303–3324, 2003.

[62] Meziat, R., Patiño, D., and Pedregal, P., "An alternative approach for non-linear optimal control problems based on the method of moments," *Computational Optimization and Applications*, vol. 38, pp. 147–171, 2007.

[63] Meziat, R., "Analysis of Nonconvex Polynomial Programs by the Method of Moments," *NONCONVEX OPTIMIZATION AND ITS APPLICATIONS*, vol. 74, pp. 353–372, 2003.

[64] Mojica, E., *Identificación y Control de un Proceso de Crecimiento Celular en Bioreactor.* http://biblioteca.uniandes.edu.co/Tesis_2005_segundo_semestre/00005127: Master thesis, University of Los Andes, 2005.

[65] Mojica, E. and Gauthier, A., "Desarrollo y simulación de un modelo para un cultivo celular en bioreactor," *Proc. of the 2nd Congreso colombiano de Bioingeniería e Ingeniería Biomédica*, 2005.

[66] Mojica, E., Gauthier, A., and Rakoto-Ravalontsalama, N., "Canonical piecewise-linear approximation of nonlinear cellular growth," *Proc. of the 46th IEEE Conference on Decision and Control*, pp. 1640–1645, 2007.

[67] Mojica, E., Gauthier, A., and Rakoto-Ravalontsalama, N., "Piecewise-linear approximation of nonlinear cellular growth," *Proc. of the 3rd IFAC Symposium on System, Structure and Control*, pp. 266–271, 2007.

[68] Mojica, E., Gauthier, A., and Rakoto-Ravalontsalama, N., "Probing control for pwl approximation of nonlinear cellular growth," *Proc. of the 13th IEEE Conference on Control Applications*, pp. 140–145, 2007.

[69] Mojica-Nava, E., Meziat, R., Quijano, N., Gauthier, A., and Rakoto-Ravalontsalama, N., "Optimal control of switched systems: A polynomial approach," *Proceedings of the 17th IFAC World Congress*, pp. 7808–7813, 2008.

[70] Mojica-Nava, E., Quijano, N., Gauthier, A., and Rakoto-Ravalontsalama, N., "Stability analysis of switched polynomial systems using dissipation inequalities," *Proceedings of the 47th IEEE Conference on Decision and Control*, 2008.

[71] Morari, M. and Baric, M., "Recent developments in the control of constrained hybrid systems," *Computers & Chemical Engineering*, vol. 30, p. 16191631, sep 2006.

[72] Morse, A., Pantelides, C., Sastry, S., and Schumacher, J., "Special issue on hybrid control systems," *Automatica J. IFAC*, vol. 35, 1999.

[73] Muñoz, J. and Pedregal, P., "A refinement on existence results in nonconvex optimal control," *Nonlinear Analysis*, vol. 46, no. 3, pp. 381–398, 2001.

[74] Namjoshi, A., Hu, W., and Ramkrishna, D., "Unveiling steady-state multiplicity in hybridoma cultures: The cybernetic approach," *Biotechnol. Bioeng*, vol. 81, pp. 80–91, 2003.

[75] Papachristodoulou, A. and Prajna, S., *Positive Polynomials in Control*, ch. Analysis of non-polynomial systems using the sum of squares decomposition, pp. 23–43. Springer, 2005.

[76] Papachristodoulou, A., Prajna, S., and Parrilo, P., "Introducing SOSTOOLS: A general purpose sum of squares programming solver," *Proceedings of the IEEE Conference on Decision and Control*, pp. 741–746, 2002.

[77] Parrilo, P., *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization.* California Institute of Technology: Ph.D. Thesis, 2000.

[78] Pedregal, P., *Parametrized Measures and Variational Principles.* Birkhäuser, 1997.

[79] Pedregal, P., *Introduction to optimization.* Springer, 2004.

[80] Pedregal, P. and Tiago, J., "Existence Results for Optimal Control Problems with Some Special Nonlinear Dependence on State and Control," *SIAM Journal on Control and Optimization*, vol. 48, pp. 415–437, 2009.

[81] Prajna, S. and Papachristodoulou, A., "On the Construction of Lyapunov Functions using the Sum of Squares Decomposition," *Proceedings of the 41st IEEE Conference on Decision and Control*, pp. 3482–3487, 2002.

[82] Prajna, S. and Papachristodoulou, A., "Analysis of switched and hybrid systems - beyond piecewise quadratic methods," *Proceedings of the American Control Conference*, pp. 2779–2784, 2003.

[83] Prieur, C., "A robust globally asymptotically stabilizing feedback: the example of the Artstein's circle," *Nonlinear Control in the Year 2000*, vol. 258, pp. 279–300, 2000.

[84] Rantzer, A. and Johansson, M., "Piecewise linear quadratic optimal control," *Automatic Control, IEEE Transactions on*, vol. 45, no. 4, pp. 629–637, 2000.

[85] Rao, M., *Measure theory and integration.* Marcel Dekker Inc, 2004.

[86] Riedinger, P., Daafouz, J., and Iung, C., "Suboptimal switched controls in context of singular arcs," *Proceedings of the 42nd IEEE Conference on Decision and Control*, pp. 6254– 6259, 2003.

[87] Riedinger, P., Iung, C., and Kratz, F., "An optimal control approach for hybrid systems," *European Journal of Control*, vol. 9, no. 5, pp. 449–458, 2003.

[88] Riedinger, P., Kratz, F., Iung, C., and Zanne, C., "Linear quadratic optimization for hybrid systems," *Decision and Control, 1999. Proceedings of the 38th IEEE Conference on*, vol. 3, 1999.

[89] Rodrigues, L. and How, J., "Automated control design for a piecewise-affine approximation of a class of nonlinear systems," *Proc. of the American Control Conference*, pp. 3189–3194, 2001.

[90] Rodrigues, L. and How, J., "Observer-based control of piecewise-affine systems," *Proc. of the 40th IEEE Conference on Decision and Control*, pp. 1366–1371, 2001.

[91] Roubi'cek, T. and Vala'sek, M., "Optimal control of causal differential-algebraic systems," *Journal of Mathematical Analysis and Applications*, vol. 269, no. 2, pp. 616–641, 2002.

[92] Savageau, M. and Voit, E., "Recasting nonlinear differential equations as S-systems: a canonical nonlinear form," *Mathematical biosciences*, vol. 87, no. 1, pp. 83–115, 1987.

[93] Seatzu, C., Corona, D., Giua, A., and Bemporad, A., "Optimal control of continuous-time switched affine systems," *Automatic Control, IEEE Transactions on*, vol. 51, no. 5, pp. 726–741, 2006.

[94] Seidman, T., "OPTIMAL CONTROL FOR SWITCHING SYSTEMS 1," *21st Annual Conference on Information Science and Control*, pp. 485–489, 1987.

[95] Shaikh, M. and Caines, P., "On the hybrid optimal control problem: Theory and algorithms," *IEEE Transactions on Automatic Control*, vol. 52, no. 9, pp. 1587–1603, 2007.

[96] Sontag, E., "Nonlinear regulation: The piecewise linear approach," *IEEE Trans. on Automatic Control*, vol. 26, pp. 346–358, 1981.

[97] Spinelli, W., Bolzern, P., and Colaneri, P., "A note on optimal control of autonomous switched systems on a finite time interval," *Proceedings of the 2006 American Control Conference*, 2006.

[98] Storace, M. and Feo, O. D., "Piecewise-linear approximation of nonlinear dynamical systems," *IEEE Trans. on Circuits Syst. I*, vol. 22, pp. 830–842, 2004.

[99] Sturn, J. F., "Using sedumi 1.02, a Matlab toolbox for optimization over symmetric cones," *Optimization Methods and Software*, vol. 11-12, pp. 625–653, 1999.

[100] Sun, Z. and Ge, S. S., "Analysis and synthesis of switched linear control systems," *Automatica*, vol. 41, pp. 181–195, 2005.

[101] Sussmann, H., "A maximum principle for hybrid optimal control problems," *Decision and Control, 1999. Proceedings of the 38th IEEE Conference on*, vol. 1, 1999.

[102] Toh, K. C., Todd, M. J., and Tutuncu, R., "SDPT3 — a Matlab software package for semidefinite programming," *Optimization Methods and Software*, vol. 11, pp. 545–581, 1999.

[103] Velut, S., de Maré, L., and Hagander, P., "Bioreactor control using a probing feeding strategy and mid-ranging control," *Control Engineering Practice*, vol. 15, no. 2, pp. 135–147, 2007.

[104] Villa, J., Duque, M., Gauthier, A., and Rakoto-Ravalontsalama, N., "A new algorithm for translating MLD systems into PWA systems," *Proc. of the American Control Conference*, pp. 1208–1213, 2004.

[105] Vu, L. and Liberzon, D., "Common Lyapunov functions for families of commuting nonlinear systems," *Systems & Control Letters*, vol. 54, pp. 405–416, 2005.

[106] Wei, S., Uthaichana, K., Žefran, M., DeCarlo, R., and Bengea, S., "Applications of numerical optimal control to nonlinear hybrid systems," *Nonlinear Analysis: Hybrid Systems*, vol. 1, no. 2, pp. 264–279, 2007.

[107] Willems, J., "Dissipative dynamical systems part I: General theory," *Arch. Rational Mech. Anal.*, vol. 45, pp. 321–351, 1972.

[108] Witsenhausen, H., "A class of hybrid-state continuous-time dynamic systems," *IEEE Transactions on Automatic Control*, vol. 11, no. 2, pp. 161–167, 1966.

[109] Xu, X. and Antsaklis, P., "Optimal control of switched systems via nonlinear optimization based on direct differentiations of value functions," *International Journal of Control*, vol. 75, no. 16-17, pp. 1406–1426, 2002.

[110] Xu, X. and Antsaklis, P., "Quadratic optimal control problems for hybrid linear autonomous systems with state jumps," *American Control Conference, 2003. Proceedings of the 2003*, vol. 4, 2003.

[111] Xu, X. and Antsaklis, P., "Optimal control of switched systems based on parameterization of the switching instants," *IEEE Trans. Automatic Control*, vol. 49, pp. 2–16, 2004.

[112] Young, L., *Lectures on the Calculus of Variations and Optimal Control Theory*. Saunders, 1969.

[113] Zaytoon, J., "Special issue on hybrid systems and applications (5)," *Nonlinear analysis*, vol. 65, no. 6, 2006.

[114] Zefran, M., Bullo, F., and Stein, M., "A notion of passivity for hybrid systems," *Proceedings of the 40th IEEE Conference on Decision and Control*, pp. 768–773, 2001.

[115] Zhao, J. and Hill, D., "Passivity and stability of switched systems: A multiple storage function method," *Systems & Control Letters*, vol. 57, pp. 158–164, 2008.

[116] Zhao, J. and Hill, D., "Hybrid h_∞ control based on multiple lyapunov functions," *Proc. of the 6th IFAC Symposium on Nonlinear Control Systems*, pp. 567–572, 2004.