UNIVERSITÉ CÔTE D'AZUR

ÉCOLE DOCTORALE
SCIENCES ET TECHNOLOGIES DE L'INFORMATION ET DE LA COMMUNICATION

THÈSE DE DOCTORAT

Allocation optimale de ressources dans la croissance bactérienne : étude théorique et applications à la production de métabolites

Agustín Gabriel Yabo

Institut National de Recherche en Sciences et Technologies du Numérique (Inria)
Équipe Projet Biocore / McTao

Présentée en vue de l'obtention du grade de docteur en Automatique, Traitement du Signal et des Images **d'**Université Côte d'Azur
Dirigée par: Jean-Luc Gouzé
Co-dirigée par: Jean-Baptiste Caillau

Soutenue le: 9 decembre 2021

Devant le jury, composé de:
Julio R. Banga, Professeur, Instituto de Investigaciones Marinas, CSIC
Jean-Baptiste Caillau, Professeur, Université Côte d'Azur
Madalena Chaves, Directrice de recherche, Inria Sophia Antipolis-Méditerranée
Jean-Luc Gouzé, Directeur de recherche, Inria Sophia Antipolis-Méditerranée
Hidde de Jong, Directeur de recherche, Inria Grenoble - Rhône-Alpes
Alain Rapaport, Directeur de recherche, INRAE Montpellier
Mohab Safey El Din, Professeur, Sorbonne Université

ANR AGENCE NATIONALE DE LA RECHERCHE

Inria

Allocation optimale de ressources dans la croissance bactérienne : étude théorique et applications à la production de métabolites

Optimal resource allocation in bacterial growth: theoretical study and applications to metabolite production

Jury:

Président du jury

Madalena Chaves, Directrice de recherche, Inria Sophia Antipolis-Méditerranée (France)

Rapporteurs

Julio R. Banga, Professeur, Instituto de Investigaciones Marinas, CSIC (Espagne)

Alain Rapaport, Directeur de recherche, INRAE Montpellier (France)

Examinateurs

Jean-Baptiste Caillau, Professeur, Université Côte d'Azur (France)

Jean-Luc Gouzé, Directeur de recherche, Inria Sophia Antipolis-Méditerranée (France)

Hidde de Jong, Directeur de recherche, Inria Grenoble - Rhône-Alpes (France)

Mohab Safey El Din, Professeur, Sorbonne Université (France)

Allocation optimale de ressources en croissance bactérienne : étude théorique et application à la production de métabolite

Agustín Gabriel Yabo

Résumé

Les micro-organismes évoluent sous la pression de la sélection naturelle, améliorant leur capacité à proliférer dans leur environnement en développant des réseaux métaboliques optimisés. Des études ont montré que les populations bactériennes peuvent atteindre un taux de croissance presque maximal dans certaines conditions, ce qui leur permet de supplanter les espèces concurrentes. Considérer l'auto-réplication microbienne comme un problème d'allocation de ressources est une nouvelle approche qui a répondu avec succès à certaines des questions sous-jacentes dans le domaine. Dans cette approche, les ressources cellulaires disponibles sont affectées dynamiquement à différentes fonctions comme le métabolisme ou la synthèse des protéines. Ce cadre a également motivé de nombreuses applications à la production artificielle de métabolites d'intérêt ; l'objectif principal est alors de détourner les ressources cellulaires des voies natives vers une voie hétérologue dans le but de synthétiser efficacement un composé spécifique (par exemple des agents antitumoraux, des antibiotiques, de l'insuline, des agents immunosuppresseurs, etc.) À cette fin, des techniques biotechnologiques récentes permettent de contrôler de manière externe la croissance bactérienne en interrompant l'expression de l'ARN polymérase. Cette thèse se concentre sur les aspects mathématiques d'une certaine classe de modèles d'auto-réplicateurs basés sur les principes d'allocation des ressources susmentionnés. Ces modèles, basés sur des hypothèses minimales, sont étonnamment efficaces pour rendre compte des lois de croissance empiriques bien étudiées des cultures microbiennes. Tout au long du manuscrit, nous revisitons certains des cadres industriels les plus pertinents pour la croissance bactérienne parmi lesquels la culture par lots et les bioréacteurs continus, ainsi que d'autres modèles simplifiés où la concentration en nutriments reste constante. L'idée est de comparer les stratégies d'allocation des ressources évoluant naturellement, où l'objectif est de maximiser la biomasse de la population bactérienne, avec les stratégies artificielles visant à maximiser la production d'un métabolite voulu. L'étude implique une analyse dynamique et une optimisation des modèles proposés, et nous avons recours à la théorie du contrôle optimal pour déterminer les stratégies d'allocation conformes à ces objectifs, tant d'un point de vue analytique que numérique. Ces stratégies optimales constituent des références de choix pour le développement de stratégies de rétroaction basées sur la mesure en temps réel des processus industriels.

Mots clés : biologie des systèmes, croissance bactérienne, contrôle optimal, biotechnologie, systèmes dynamiques non linéaires.

Optimal resource allocation in bacterial growth: theoretical study and applications to metabolite production

Agustín Gabriel Yabo

Abstract

Microorganisms have evolved under the pressure of natural selection, improving their capacity to proliferate in the environment by developing highly optimized metabolic networks. Studies showed that bacterial populations can achieve nearly maximal growth rate under certain conditions, allowing them to outgrow competing species. Considering microbial self replication as a resource allocation problem is a novel approach that successfully answered some of the underlying questions in the field. This perspective considers the problem of dynamically assigning the available cellular resources to different cellular functions, such as metabolism and protein synthesis. The framework has also motivated numerous application to the artificial production of metabolites of interest, where the main objective is to divert the cellular resources from the native pathways into a heterologous pathway in order to efficiently synthesize a specific compound (e.g. antitumour agents, antibiotics, insulin, immunosuppressive agents, etc.). To this end, recent biotechnological techniques allow to externally control bacterial growth by shutting off the expression of RNA polymerase. This thesis focuses on the mathematical aspects of a certain class of self-replicator models based on the aforementioned resource allocation principles. These models, based on minimal assumptions, are surprisingly effective in accounting for well-studied empirical growth laws of microbial cultures. Throughout the manuscript, we revisit some of the most relevant industrial frameworks for bacterial growth, such as batch cultivation and continuous bioreactors, and other simplifying approximations where the nutrient concentration remains constant. The objective is to compare the naturally-evolved resource allocation strategies, where the objective is to maximize the biomass of the bacterial population, with the artificial strategies aiming to maximize the production of the metabolite of interest. The study involves dynamical analysis and optimization of the proposed models, and we resort to Optimal Control theory to find the allocation strategies complying with these objectives, both from analytical and numerical points of view. Ultimately, such optimal strategies may provide guidance for developing feedback strategies based on real-time measuring of industrial processes.

Keywords: systems biology, bacterial growth, optimal control, biotechnology, nonlinear dynamical systems.

Acknowledgements

I would like to start by thanking both my supervisors Jean-Luc Gouzé and Jean-Baptiste Caillau, without whom this manuscript would not have been written. I am specially grateful for their trust, the freedom I was given to make my own path and to explore my own ideas, and for pushing me to learn "la langue de Molière".

I am also very grateful to the members of the jury in the PhD defense, in particular to Alain Rapaport and Julio R. Banga for the effort of thoroughly going through this PhD manuscript, and the rich feedbacks they gave me. I also appreciate the participation of Hidde de Jong and Madalena Chaves, who have been part of the project in many different levels (as members of both the Personal Monitoring Committee and the PhD jury), and with whom we shared very fruitful scientific discussions. Additionally, I am grateful to Hidde and the whole Microcosme team from Inria Grenoble - Rhône-Alpes (specially to Johannes Geiselmann, Eugenio Cinquemani and Antrea Pavlou) for welcoming me and helping me during the (brief but) exciting experimental phase of my PhD. Finally, I thank Mohab Safey El Din for taking the time to participate in the jury and for the ongoing joint work.

I am more than grateful to the teams Biocore and McTao for the excellent research environment we shared these years, and for the friends I have done during that period, with whom we shared lots of excellent moments around a coffee/mate/beer. I particularly enjoyed the scientific works we have done with my friend Nicolas Augier. I also acknowledge the inspiring time I spent at the Ctrl-A and Datamove teams in Grenoble, during my master's internship.

I would like to thank the Université Côte d'Azur for funding the project, as well as its Inria research center and the ANR project Maximic (ANR-17-CE40-0024-01).

Agradezco fuertemente el apoyo que me brindó Céline, qui m'a accompagné pendant la dernière année de ma thèse, et qui m'a offert la meilleure journée de soutenance de thèse qu'un humble doctorant puisse souhaiter.

No puedo dejar de agradecer a quien me viera dar mis primeros pasos en el mundo de la investigación, Damián E. Oliva, que supo orientarme cuando más lo necesité; y a los miembros del Departamento de Ciencia y Tecnología de la Universidad Nacional de Quilmes.

Finalemente, les debo un infinito agradecimiento a mis padres Andrea y Gabriel, y a mi hermana Anita, quienes siempre fueron mi modelo a seguir. Gracias por brindarme tantas herramientas y motivarme siempre a hacer lo que me apasiona.

Contents

List of Figures

Thesis-related publications

[AYT1] Yabo, A., J.B. Caillau, J.L. Gouzé, Hidde de Jong, and Francis Mairet. Dynamical analysis and optimization of a generalized resource allocation model of microbial growth (accepted for publication). *SIAM Journal on Applied Dynamical Systems*, 2021.

[AYT2] Agustín Gabriel Yabo, Jean-Baptiste Caillau, and Jean-Luc Gouzé. Optimal bacterial resource allocation: metabolite production in continuous bioreactors. *Mathematical Biosciences and Engineering*, 17(6):7074–7100, 2020.

[AYT3] Yabo, A., J.B. Caillau, and J.L. Gouzé. Hierarchical MPC applied to bacterial resource allocation and metabolite synthesis (accepted for publication). In *The 60th IEEE conference on Decision and Control*, 2021.

[AYT4] Agustín Gabriel Yabo and Jean-Luc Gouzé. Optimizing bacterial resource allocation: metabolite production in continuous bioreactors. *IFAC-PapersOnLine*, 53(2):16753–16758, 2020.

[AYT5] Agustín Gabriel Yabo, Jean-Baptiste Caillau, and Jean-Luc Gouzé. Singular regimes for the maximization of metabolite production. In *2019 IEEE 58th Conference on Decision and Control (CDC)*, pages 31–36. IEEE, 2019.

[AYT6] Agustín Yabo, Jean-Baptiste Caillau, and Jean-Luc Gouzé. Bacterial growth strategies as optimal control problems: maximizing metabolite production. In *FGS'2019-19th French-German-Swiss conference on Optimization*, 2019.

Other publications

[AYO1] N. Augier and Yabo, A. Time-optimal control of piecewise affine bistable gene-regulatory networks (accepted for publication). *International Journal of Robust and Nonlinear Control*, 2021.

[AYO2] N. Augier and Yabo, A. Time-optimal control of piecewise affine bistable gene-regulatory networks: preliminary results. In *7th IFAC Conference on Analysis and Design of Hybrid Systems*, 2021.

[AYO3] C. Djuikem, Yabo, A., F. Grognard, and S Touzeau. Mathematical modelling and optimal control of the seasonal coffee leaf rust propagation. In *7th IFAC Conference on Analysis and Design of Hybrid Systems*, 2021.

Online contributions

All the works presented in this manuscript contain a numerical section including simulations of the dynamical systems and trajectories of the optimal control problems. In order to guarantee the usability and reproducibility of the results, a series of online notebooks have been developed, which allow readers to execute the simulations and optimal control algorithms. The examples are part of the gallery (`https://ct.gitlabpages.inria.fr/gallery/`) of the ct (control toolbox) project, which contains several examples of optimal control problems solved numerically using a python wrapper of bocop3 (`https://www.bocop.org/`) that can be executed on a Binder live web environment (`https://mybinder.org/`). The notebooks are listed below:

- Optimal resource allocation in a generalized model of microbial growth
  `https://ct.gitlabpages.inria.fr/gallery/bacteria/bacteria.html`
- Bacterial biomass maximization with substrate depletion
  `https://ct.gitlabpages.inria.fr/gallery/substrate/depletion.html`
- Bistable gene-regulatory networks
  `https://ct.gitlabpages.inria.fr/gallery/bistable/bistable.html`
- Seasonal coffee leaf rust propagation
  `https://ct.gitlabpages.inria.fr/gallery/clr/clr.html`

Chapter 1

Introduction

Microbiological research, while recent in comparison with the study of other living creatures, has been the main task of many scientists across the globe, even when the available technological means were scarce. There is a vast number of compelling reasons that explain this fact. For instance, understanding bacteria is a first step towards understanding—and potentially controlling—the diseases caused by them. In addition, most of the antibiotics manufactured by the pharmaceutical industry are produced by bacteria as a *secondary metabolite*, so named for not being involved in the growth and reproduction of these organisms. Numerous compounds essential in medical practices are also synthesized by bacteria, such as antitumor agents, immunosuppresive agents and insulin. The latter is also the case in the food industry, for products such as yogurt, cheese and pickles [7].

Whereas investigating such microorganisms usually starts with an experimental phase in a laboratory, a fundamental part of the discipline involves the exploitation of the obtained data, and the challenge of drawing knowledge from it. Due to its complexity, this task brings together scientists from very different backgrounds, such as mathematicians, biologists, statisticians and computer scientists. While the theoretical and practical approaches used in each discipline are quite diverse, scientists agree on a very common, interdisciplinary tool: mathematical models, widely used in the domain—and in science in general—for their capacity to capture and predict key behaviours of these living systems.

1.1 Mathematical modelling

A powerful approach used to study bacterial growth is mathematical modelling, for its proved success in representing complex networks of biochemical reactions behind internal cellular processes such as metabolism and gene expression [8]. Due to the complexity of even the smallest living organisms, simple features such as the growth rate of a bacterial population is usually governed by a large number of internal reactions forming these complex networks, and thus it is not expected to be explained by simple deterministic laws. However, through continuous experimental analysis of growing microbial culture in controlled environments, it is possible to observe certain recurring phenomena, and to describe them through empirical mathematical expressions. For instance, during the exponential growth phase of a bacterial culture, the rate of growth is constant, and thus it is reasonable—and customary—to consider all relative concentrations in the cell as constants [9]. At this stage, the number of cells in the population can be modelled using a simple dynamical equation

$$\dot{N} = \mu\, N, \tag{1.1}$$

where N is the number of cells, and μ the constant population growth rate. Thus, the amount of cells at time t is given by

$$N(t) = N_0\, e^{\mu t}.$$

In the same line, a major milestone in microbiology is the work of Monod [10], for successfully being able to describe the relation between microbial growth and its medium. The main contribution was the concept of limiting nutrient: an empirical relationship between the exponential steady-state growth rate μ of a growing culture of *Escherichia coli* and the concentration s of available substrate in the medium. The latter can be represented by an hyperbolic function

$$\mu(s) = \mu_{\max}\, \frac{s}{K_s + s},$$

where $\mu_{\max}$ is the maximal possible bacterial growth rate, and K_s a constant. Then, equation (1.1) can be written as

$$\dot{N} = \mu_{\max} \frac{s}{K_s + s} N$$

The latter can be related to classical Michaelis-Menten dynamics [11], an equation proposed more than a hundred years ago predicting a proportionality relation between the rate of an enzyme-catalyzed reaction and the concentration of the substrate consumed in the reaction.

Almost ten years after Monod's work, experimental studies found that the concentration of RNA (ribonucleic acid) in microbial cells was directly linked to the growth rate [12]. More precisely, the results showed an empirical linear relation between the RNA/total protein ratio and the growth rate at steady state in bacteria [1], as illustrated in Figure 1.1. This study was one of the first relating growth rate and the cell macromolecular composition, and motivated further investigation on these empirical laws, currently known in the literature as *growth laws*. These

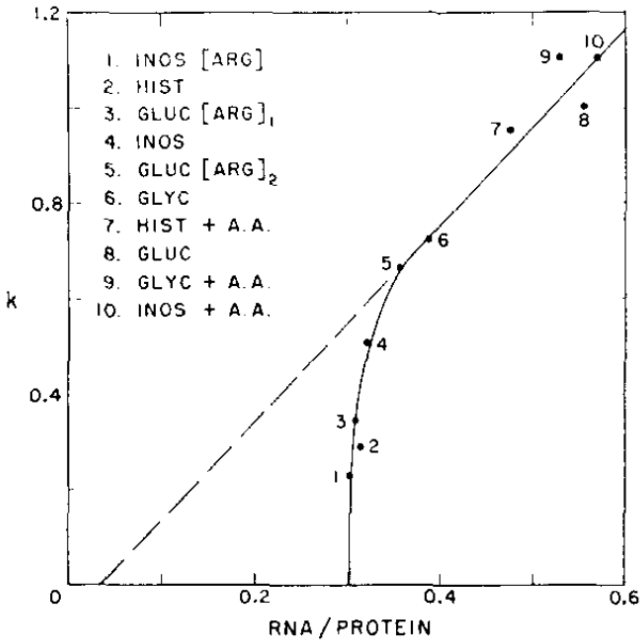


Figure 1.1: The relation between steady-state growth rate and RNA content of *A. aerogenes*, strain 5-P14 [1].

results have been instrumental in the way we understand microorganisms, and are

still valid nowadays [2], as seen in Figure 1.2.

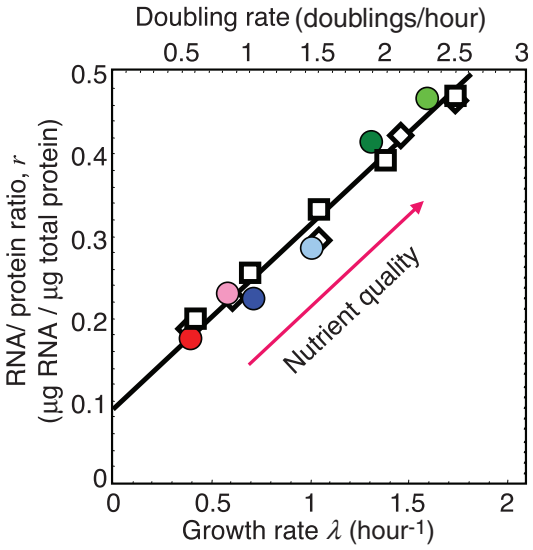


Figure 1.2: Correlation of the RNA/protein ratio r with growth rate λ for various strains of E. coli. The growth rate is modulated by changing the quality of nutrients [2].

An essential machinery of bacteria is the metabolism, defined as the ensemble of internal chemical reactions responsible for the production of energy from nutrients, growth and replication. There has been numerous works focusing on the prediction of metabolic networks during exponential growth (i.e. steady-state conditions) aiming to better understand the internal resources distribution [13, 14]. These studies aim to understand and predict how bacteria adapt to different mediums, such as acetate, glucose, and glycerol. The obtained results support an assumption that has been adopted in many previous works: under certain conditions, bacteria adapt the way they distribute their internal resources so as to maximize their growth rate.

1.2 Optimality in nature

A common hypothesis when studying bacterial growth is that bacteria seek to maximize their growth rate [15, 2, 16, 17]. From an evolutionary perspective, the latter can be viewed as a result of the pressure of natural selection: in environments with scarce nutrient, a higher growth rate represents an evolutionary advantage allowing

species to outgrow competitors. This concept is known as fitness, and is usually represented in mathematical models as optimality with respect to a specific criteria (i.e. growth rate, growth yield, ...) [18]. Different modelling techniques have been employed to predict the optimality principles behind bacterial growth. A classical example is FBA (Flux Balance Analysis), an approach that aims to predict metabolic flux distributions in a quantitative way based on certain optimization criteria, while complying with a number of physical and chemical constraints such as mass balance and energy balance [19]. As a result, this technique is able to predict (sometimes multiple) steady-state solutions of internal fluxes inside a cell based solely on the stoichiometry of the metabolic networks. However, in certain cases, this technique can require imposing very strong simplifying assumptions in order to predict maximal growth rate, as the approach tends to exclude some important biomass constituents from the model [8]. Other similar approaches such as RBA (Resource Balance Analysis) consider the optimization problem together with the production of proteins and enzyme-related cost functions [20].

More recent studies focus on the problem of adapting to changing environments from a resource allocation perspective. The approach is based on the idea that the mechanisms behind the allocation of internal resources to different cellular functions have been optimized through evolution to maximize the population growth rate [21]. The advantage of this framework is the capacity to yield models encompassing mechanistic trade-offs related to physical and biochemical limitations. An example can be seen in [16], where the model is constructed taking into account constraints such as the limited size of the pools of energy and ribosomes in a cell. Intuitively, producing proteins to tackle a specific task implies diverting resources from other tasks, which creates numerous inherent trade-offs between cellular functions. There are well-studied examples in the literature, such as the trade-off between nutrient uptake and protein synthesis [2], or growth and stress response [22]. Through this methodology, it has been possible to link the optimal regulation of ribosome synthesis in the cell to a well-known natural regulation mechanism based on the presence of guanosine tetraphosphate (also known as the ppGpp alarmone) [23]. Recent experimental results were able to confirm such theoretical predictions by showing how the artificial increase or reduction of the concentration of the ppGpp molecule in *E. coli* leads to non-optimal resource allocation strategies [24].

A natural way of representing bacterial growth is through self-replicator models,

which are becoming increasingly popular for their simplicity and their potential to predict the empirical growth laws above mentioned. In this context, the impact of simple natural trade-offs in bacterial growth can be studied through the question: what is the resource allocation strategy that maximizes the rate of replication of a self-replicating system?.

1.3 Self-replicator models

The study of self-replicator systems is a domain that has drawn the attention of scientists for more than 70 years. The questions of whether a system is able to self replicate, and how, have been of interest even before the discovery of the DNA and the current available knowledge in biological research. Its interest does not only come from living beings, but also from its applications to space exploration, nanotechnology, and computer science in general [25]. However, cellular replication has been a major driver in the study of self replication, for its potential to understand the biological mechanisms behind it, to transfer this knowledge to artificial systems, and to optimize the existing biological processes.

Bacterial growth is not strictly a replication process, as each new individual in the culture is subject to a strong genetic variability, producing mutations from one generation to another. However, representing microbial growth as a process of self replication has proven surprisingly helpful in elucidating multiple natural mechanisms of unicellular organisms [26]. A very simple example of a real self-replicating biological system consists of an RNA enzyme that synthesises itself from a precursor [27]. As briefly explained in previous sections, if the substrate in the medium is constant, the population is expected to undergo exponential growth. Self-replicator models have also been useful in studying fundamental biological mechanisms as homeostasis and circadian rhythms [28]. Other examples were able to reproduce a metabolic strategy adopted by unicellular organisms known as overflow metabolism, which consists of an energetic inefficiency in bacteria growing at a high growth rate [26].

Most of these models consider the process of self replication in a steady-state setting, which is rarely a condition attained by bacteria in nature. While these works can be useful to study phenomena that occur in reduced timescales, or in controlled environments such as a laboratory, some questions regarding the adaptation of living

systems to changing environments require an explicit formulation of such dynamism, which entail a different mathematical approach to the problem.

1.4 A dynamical perspective

The changing nature of the environments in which bacteria inhabit motivates a dynamical perspective of the resource allocation problem [29]. A recent paper has focused in the dynamical allocation of cellular resources facing nutrient upshifts [30], where the bacterial cell is represented through a coarse-grained self-replicator dynamical model described by a system of ODEs. The approach considers a simplified proteome (i.e. the entire set of proteins in a cell) divided into metabolic enzymes (responsible for nutrient uptake) and ribosomal proteins (responsible for the synthesis of macromolecules) and tackles the question of how to allocate resources between these two cellular functions so as to optimally transition between two different environments. The main hypothesis behind the study is that the resource allocation strategy has been optimized by natural selection to maximize the bacterial growth rate. In a system of ODEs, the mathematical problem of finding the time-varying function that maximizes a certain criterion is called an OCP (Optimal Control Problem), and it can be solved through PMP (Pontrjagin's Maximum Principle) [31]. Results show that the optimal allocation strategy consists of a transient period followed by a steady-state phase matching previous static results [26]. The transient is composed of oscillations between the two limit allocations: maximal enzyme synthesis and maximal ribosomal production. Surprisingly, the natural regulatory mechanism behind the ppGpp molecule produces a similar control structure to that of the optimal allocation found through the optimal control approach.

The approach of Giordano *et al.* is a prime example of how simple natural trade-offs can be studied through abstract mathematical models. A similar work investigates the trade-off between storage and enzymatic compounds in cells [32]. By hypothesis, the natural objective is to cumulate biomass, which is described by both compounds, but the metabolic capacity of the bacterial cell is supposed to be affected only by enzymes. The paper also tackles the dynamical optimization problem using a PMP approach. A quantitative model of dynamical resource allocation describing transitions between different nutrients has been proposed in [33], which is independent of the kinetic parameters behind each reaction, thus providing robust-

ness of the results to parametric uncertainty. A resource allocation model was used to describe bacterial growth during nutritional upshifts [34], suggesting that, in unstable environments, ribosomal sub-saturation can represent a selective advantage for allowing faster growth rate increases after nutrient upshifts. A more recent work proposed a minimalistic whole-cell coarse-grained dynamical model representing the proteome composition, growth rate and size of the cell, that takes into account a class of proteins dedicated to cell division [35]. While the model considers that the growth rate is affected by environmental conditions such as changes in the nutrient quality or over-expression of growth rate-independent proteins, it is not based on the maximal growth rate principle, in contrast with most of the studies in the literature.

Dynamical approaches based on minimal models have been key in understanding the regulation mechanisms behind natural trade-offs, but it have also proved fruitful in analyzing trade-offs between natural and artificially engineered resource pathways as, for instance, those targeting the synthesis of value-added compounds.

1.5 Metabolite production

A classic strategy for producing value-added metabolites and other chemical compounds is to artificially engineer the production of an heterologous enzyme into a cell host, which will thus synthesize the targeted metabolite. For instance, *E. coli* can be modified to produce glucuronic and glucaric acid (used in the production of detergents) from glucose by reenginering a synthetic pathway into the microbial cell [36]. However, the limitation of cellular resources yields a competition between the native pathways of the cell (dedicated to growth) and the synthetic circuits (dedicated to metabolite synthesis). Numerous works have studied this trade-off, aiming to improve the current industrial bioprocesses. Bacterial growth can be inhibited through growth inhibitors, thus increasing the yield of metabolites, as shown in recent experimental results [37]. While this is an effective technique, it can also be expensive, and can produce toxicity to the host cell [38]. A less costly, innovative tool is the optogenetic regulation of gene expression, which has been successfully implemented in a real-time feedback loop in continuous liquid cultures [39].

A promising alternative is the engineering of synthetic gene circuits. An example can be seen in [40], where a synthetic control circuit is able to compensate for the demand of resources from both the native and engineered pathways by regulating

the expression of enzymes. A method for engineering metabolic control circuits is proposed in [41], which tackles the burden of deviating resources from growth into the heterologous pathway through a multiobjective optimization problem, towards a framework for automated design of genetic circuits. Other approaches, based on coarse-grained models of the proteome, focuses on the loss of fitness of individuals in the population produced by the burden of the heterologous pathway, and on how to overcome such effect [42].

The capacity to externally affect gene expression through external real-time controllers, combined with the increasing understanding of natural resource allocation strategies, has motivated approaches based on self-replicator models [43]. The latter focuses on how to regulate the expression of RNA polymerase through a synthetic growth switch [44] in order to maximize the production of a metabolite of interest. The model represents the trade-off biomass/compound synthesis through a resource allocation problem similar to the one studied in Giordano *et al.*, and the results emphasizes the differences between the control strategies depending on the particular objective.

Based on these works, we propose a series of coarse-grained self-replicator dynamical models representing the proteome of a bacterial cell, and we study the trade-offs arising from inherent natural allocation compromises, as well as those related to the artificial production of compounds of interest. The mathematical tools employed for this task range from dynamic systems analysis of the given models, to optimization and optimal control theory for studying the associated static and dynamical maximization objectives.

1.6 Studied problems

In this section, a brief description of the nature of the models used in this manuscript is given, all devised from a resource allocation perspective, that consider some of the most relevant industrial production frameworks. Each chapter is written in a self-contained manner, as they are based on (already published or in production) scientific papers, and thus there are substantial differences in the biological hypotheses and the overall notation. However, some common definitions can be provided to describe the organization of the manuscript.

The time-varying quantities represented in the models are divided into the extra-

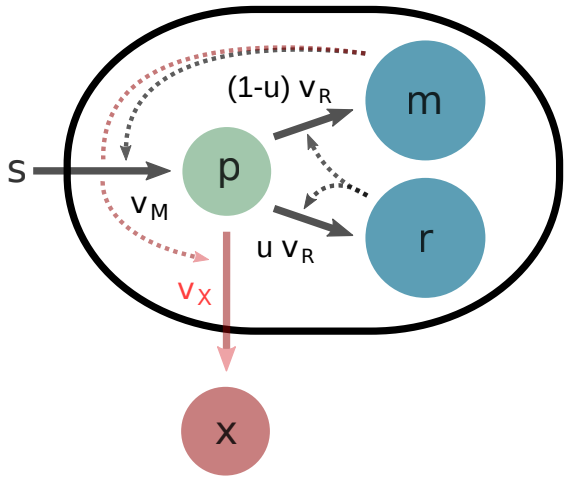


Figure 1.3: Schema of the system.

cellular concentrations (i.e. with respect to a bioreactor of volume $\mathcal{V}_{\text{ext}}$) and intracellular concentrations (i.e. with respect to the bacterial volume $\mathcal{V}$). The extracellular variables are s and x, corresponding to the substrate and metabolite concentrations with respect to the external volume, respectively. The intracellular variables are p, r and m, which correspond to the concentrations of precursor metabolites, proteins of the gene expression machinery, and proteins of the metabolic machinery, respectively. The main reaction rates are v_M, v_R and v_X, and are functions of the concentrations. The precursor synthesis rate $v_M(s, m)$ is increasing with respect to the concentration of substrate in the medium s and catalyzed by m. The macromolecule synthesis rate $v_R(p, r)$ is increasing with respect to the concentration of precursors p and catalyzed by r. Finally, $v_X(p, m)$ is the metabolite synthesis rate, also increasing with respect to the concentration of precursors p, and catalyzed by m. A schema of the system is depicted in Figure 1.3. While the derivations of each particular model are detailed in each chapter throughout the manuscript, and thus they are omitted in this introduction, we describe here the obtained differential equations governing the time evolution of the defined quantities. The dynamics of the pool of precursors is described by the differential equation

$$\dot{p} = v_M(s, m) - v_X(p, m) - v_R(p, r)(p + 1).$$

The dynamics of the concentrations of macromolecules in the cell, described by

ribosomal and enzymatic proteins, are given by

$$\dot{r} = (u - r)v_R(p, r),$$
$$\dot{m} = (1 - u - m)v_R(p, r),$$

where $u \in [0, 1]$ is the time-varying allocation strategy, which can represent either the natural allocation or an artificial human-driven allocation resulting from the effect of an external control, depending on the objective to be studied. The dynamics of the bacterial volume and the metabolite concentration are

$$\dot{\mathcal{V}} = v_R(p, r)\mathcal{V},$$
$$\dot{x} = v_X(p, m)\frac{\mathcal{V}}{\mathcal{V}_{ext}}.$$

The simplified case, where the dynamics of s are omitted—representing the case where the substrate concentration is constant for the system—was first studied in [3], and is further investigated in Chapter 2. In this work, s is assumed to be constant, and so the dynamics of p becomes

$$\dot{p} = \boldsymbol{v_M(m)} - v_X(p, m) - v_R(p, r)(p + 1),$$

where v_M only depends on m. Intuitively, this case is expected to undergo exponential bacterial growth after a certain transient. An extension of this work is presented in Chapter 5, where a hierarchical MPC (Model-Predictive Control) loop is proposed, based on the studied open-loop optimal control.

The case where an initial amount of substrate s_0 in the bioreactor is gradually consumed represents a batch process, and is studied in Chapter 3. The latter is modelled through the dynamics of s given by the equation

$$\dot{s} = -v_M(s, m)\frac{\mathcal{V}}{\mathcal{V}_{ext}},$$

and is expected to eventually reach a steady state when all the available substrate in the bioreactor is depleted. A scheme of metabolite production in continuous bioreactors is studied in Chapter 4. In this case, there is a continuous constant inflow of substrate of concentration s_{in} to the bioreactor, as well as an outflow (of the same volumetric flow rate) of substrate, biomass and metabolites. This flow is

characterized by the constant D denoted as the dilution rate. Thus, the dynamics are modified with additional terms

$$\begin{aligned}
\dot{s} &= \boldsymbol{D s_{in}} - v_M(s,m)\frac{\mathcal{V}}{\mathcal{V}_{ext}} - \boldsymbol{D s}, \\
\dot{\mathcal{V}} &= v_R(p,r)\mathcal{V} - \boldsymbol{D\mathcal{V}}, \\
\dot{x} &= v_X(p,m)\frac{\mathcal{V}}{\mathcal{V}_{ext}} - \boldsymbol{D x}.
\end{aligned}$$

The latter system is also expected to reach a steady state, depending on the dilution rate D and on the allocation variable u. An experimental stage of this project, consisting of *E. coli* cultures in continuous bioreactors, was conducted in collaboration with Ibis team (Inria Grenoble - Rhône-Alpes). However, it was not possible to perform the required amount of experiments due to the pandemic, and so these results were not included in the manuscript.

Chapter 6 describes a self-replicator model that considers growth rate-independent proteins in the proteome. The latter aims to model a wild-type bacteria with no metabolite production, and so the synthesis rate of metabolites $v_X(p,m) = 0$. Growth rate-independent proteins are divided into ribosomal proteins and housekeeping proteins. The fact that there are proteins of the gene expression machinery that do not contribute to growth produces a minimal ribosomal concentration $r_{\min}$ required to have bacterial growth. Thus, the synthesis rate of macromolecules becomes $v_R(p, r - r_{\min})$. Additionally, since a part of the proteome is dedicated to housekeeping proteins q, there is a maximal ribosomal concentration $r_{\max}$. The equations for the intracellular concentrations in the proteome are then given by

$$\begin{aligned}
\dot{r} &= (\boldsymbol{r_{\max}} u - r) v_R(p, r - \boldsymbol{r_{\min}}), \\
\dot{m} &= (\boldsymbol{r_{\max}}(1-u) - m) v_R(p, r - \boldsymbol{r_{\min}}), \\
\dot{q} &= ((1 - \boldsymbol{r_{\max}}) - q) v_R(p, r - \boldsymbol{r_{\min}}).
\end{aligned}$$

Finally, an additional section in Appendix A describes supplementary work carried out on gene-regulatory networks, produced as part of a collaborative work during my PhD. The results fall within the realm of Systems Biology, but they have been excluded from the main body of the manuscript for being out of the specific scope of the thesis.

Chapter 2

Constant substrate inflow

This chapter reproduces [AYT5], published in the 58th Conference on Decision and Control (IEEE CDC 2019), and some unpublished results on the stability of the system of interest and the uniqueness of the solution of the static optimization problem.

2.1 Introduction

In nature, microorganisms are continuously facing nutrient availability changes in the environment, and thus they have evolved to dynamically adapt their physiology to cope with this phenomenon. This is achieved through reorganization of the gene expression machinery, by dynamically allocating resources to different cellular functions. Among all possible allocation strategies, only few will guarantee survival when competing for nutrients, leading to complex and highly optimized organisms. In the specific case of *Escherichia coli*, studies have shown that, under certain conditions, bacterial populations achieve nearly maximal growth-rate, suggesting that this feature is indeed a design objective resulting from evolutionary processes [14]. These experimental results have triggered a large number of studies where the growth-rate maximization strategy is a central assumption when approaching resource allocation problems [21]. However, most of these works consider the resource allocation problem in steady-state conditions, which do not represent the natural environment of bacterial populations, thus motivating a dynamical approach to the problem. Such dynamical behaviours can be modeled through the so-called self-replicator models, widely used in bacterial growth representations for its simplicity and its capacity to

reproduce observed experimental behaviours [45].

The starting point for this line of research is [30], where the authors addressed the problem of dynamical allocation of cellular resources, showing that maximizing the steady-state growth is a sub-optimal strategy under changing environments. The dynamical growth-rate maximization can be interpreted as a biomass maximization problem during a fixed time period. Thus, it is possible to reformulate the question as an OCP (Optimal Control Problem), to be solved by means of the PMP (Pontryagin's Maximum Principle). This theoretical approach can provide gold standard allocation strategies, that can be then compared to feasible growth control implementations in bacterial cells. These results have provided a baseline understanding upon which it is possible to re-engineer the naturally-evolved behaviors of the cell in order to improve certain productivity measures. In particular, we consider the problem of producing a certain metabolite of interest, as considered in [43], for its relevance in biotechnological processes. In this regard, Optimal Control theory can not only shed light on the natural bacterial intracellular behaviours, but also help enhance industrial processes, as well as provide guidance in biotechnological research.

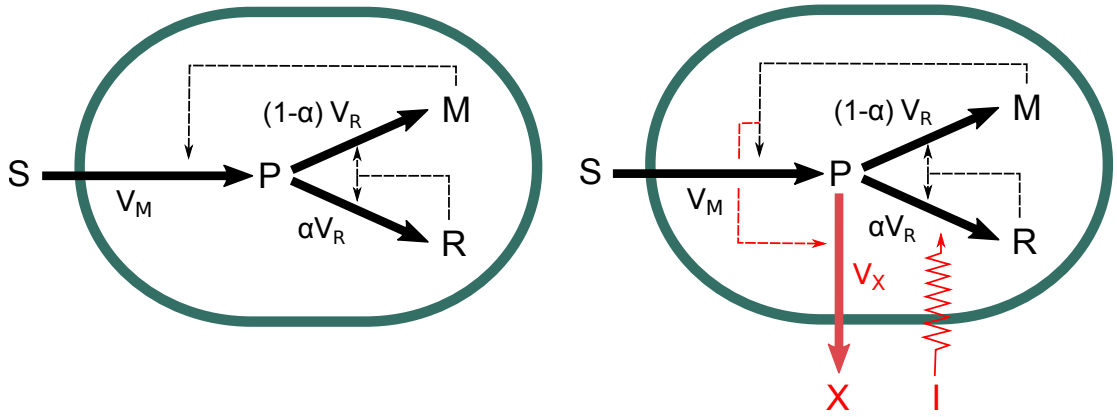
In this work, we present a general coarse-grained model for a self-replicating system extended with the metabolite production pathway based on [30, 43]. We start, in Section 2.2, by considering the case of the CSTR (Continuous Stirred-Tank Reactor) Bioreactor scheme as our baseline, and we show that is it possible to derive the models previously analyzed in the field (fed-batch, constant substrate, no production of metabolites) as an initial step towards a full analysis of the new system. Then, in Section 2.4, we focus on two dynamical problems: 1) biomass maximization when there is no production of metabolites, a feature assumed to be achieved by living organisms through evolution; and 2) product maximization under constant environmental conditions (corresponding to the fed-batch bioreactor), an artificial objective stated purely for biotechnological purposes. Ultimately, the comparison between this two problems should help to understand how to dynamically disrupt the natural allocation process in order to prioritize the metabolite production pathway instead of the population's growth rate, which is the natural behavior of bacteria. From the biological point of view, our results show that in order to optimally produce the artificial compound, the dynamical allocation of resources should be progressively altered to allocate more resources to the metabolic machinery of

the cell population. Both problems are tackled through Optimal Control theory, and then solved using PMP. The solutions of both OCPs turn out to be *singular controls*, characterized by the existence of a singular arc along the solution. Consequently, we proceed in Section 2.5 to characterize the singular trajectories in both cases, detailing the computations required to obtain the second order singular arc, and providing a numerical check of the suitable Legendre-Clebsch condition.

2.2 Self-replicator models

2.2.1 CSTR Bioreactor model with metabolite production

As previously stated, the problem of resource allocation in bacteria can be studied through the so-called self-replicator models. We consider a self-replicating system in a CSTR Bioreactor of volume $\mathcal{V}_{ext}$. The cell is composed of the gene expression machinery (R) and the metabolic machinery (M), as seen in Figure 2.1a. Based on the extension introduced in [43], a metabolic pathway for the production of a certain metabolite of interest X is included (Figure 2.1b).



(a) Original scheme [30]. (b) Extended scheme [43].

Figure 2.1: Self-replicator models of bacterial growth.

The system is described by three chemical macroreactions

$$\mathrm{S} \xrightarrow{\mathrm{V_M}} \mathrm{P},$$
$$\mathrm{P} \xrightarrow{\mathrm{V_R}} \alpha\mathrm{R} + (1\text{-}\alpha)\,\mathrm{M},$$
$$\mathrm{P} \xrightarrow{\mathrm{V_X}} \mathrm{X}.$$

The first reaction transforms an external substrate (S) into precursor metabolites (P) and is catalyzed by M. The second one converts precursors into macromolecules, and is catalyzed by R. Finally, a third reaction transforms precursors into the product X, and is also catalyzed by M. The parameter $\alpha \in [0, 1]$ represents the resource allocation choice, and determines for each time instant the proportion of precursor allocated to the gene expression machinery, while $1 - \alpha$ indicates the proportion allocated to the metabolic machinery. The rates at which these reactions occur are V_M, V_R and V_X [g h^{-1}]. The system is subject to a constant volumetric flow rate F [L h^{-1}] which generates both an inflow of fresh medium rich in substrate, and an outflow of biomass and metabolites [46]. Moreover, the scheme is extended with the growth switch described in [44] that allows to shut off the production of ribosomes and other components of the gene expression machinery. Then, the new resource allocation variable becomes

$$u(t) = I(t)\,\alpha(t), \quad u \in [0, 1], \tag{2.1}$$

where I is the external control and α the natural allocation mechanism used in [30]. While these two control functions are supposed to act independently, we are interested in obtaining the optimal combination of them. Thus, in this work, we restrict the analysis to calculate optimal control u, without decoupling the individual controls. Then, the time evolution of the mass of each component can be written as

$$\begin{cases} \dot{S} = V_{S_{in}} - V_M - V_{S_{out}}, \\ \dot{P} = V_M - V_R - V_X - V_{P_{out}}, \\ \dot{M} = (1 - u)\,V_R - V_{M_{out}}, \\ \dot{R} = uV_R - V_{R_{out}}, \\ \dot{X} = V_X - V_{X_{out}}, \end{cases} \tag{2.2}$$

where the inflow/outflow rates are defined as $V_{S_{out}} = DS$, $V_{P_{out}} = DP$, $V_{M_{out}} = DM$, $V_{R_{out}} = DR$, $V_{X_{out}} = DX$, $V_{S_{in}} = Fs_{in}$, s_{in} [g L^{-1}] being the concentration of the nutrient input, and D [h^{-1}] the dilution rate given by the relation $F/\mathcal{V}_{ext}$. Under the assumption that the cytoplasmic density of the cells is constant throughout the

population, we define the volume of the cell population $\mathcal{V}$ [L] as

$$\mathcal{V} \doteq \beta(M + R), \tag{2.3}$$

where β [L g^{-1}] corresponds to the inverse of the cytoplasmic density. The definition (2.3) is based on the experimental fact that macromolecules explain most of the biomass in microbial cells [6]. Then, for the sake of convenience, quantities of the system are expressed as concentrations,

$$p \doteq \frac{P}{\mathcal{V}}, \quad r \doteq \frac{R}{\mathcal{V}}, \quad m \doteq \frac{M}{\mathcal{V}}, \quad s \doteq \frac{S}{\mathcal{V}_{ext}}, \quad x = \frac{X}{\mathcal{V}}$$

where p, r and m [g L^{-1}] are intracellular concentrations of precursor metobolites, ribosomes (and other components of the gene expression machinery) and metabolic enzymes respectively (with respect to the cell population volume), s [g L^{-1}] is the extracellular concentration of substrate with respect to a constant external volume $\mathcal{V}_{ext}$ [L]; and x [g L^{-1}] a concentration without a precise biological interpretation (since the metabolite is excreted from the cell population). As a result, it is possible to exclude the dynamics of m from the analysis since, by construction, $r + m = 1/\beta$. Replacing with concentrations leads to the following system,

$$\begin{cases} \dot{s} = D(s_{in} - s) - v_M(s, m)\dfrac{\mathcal{V}}{\mathcal{V}_{ext}}, \\ \dot{p} = v_M(s, m) - v_R(p, r) - v_X(p, m) - \mu(t)p, \\ \dot{r} = uv_R(p, r) - \mu(t)r, \\ \dot{m} = (1 - u)\, v_R(p, r) - \mu(t)m, \\ \dot{x} = v_X(p, m) - \mu(t)x \\ \dot{\mathcal{V}} = (\mu(t) - D)\, \mathcal{V}, \end{cases} \tag{2.4}$$

where v_M, v_R and v_X [g L^{-1} h^{-1}] are the mass fluxes per unit volume obtained from dividing the rates V_M, V_R and V_X by $\mathcal{V}$; and $\mu(t)$ [h^{-1}] is the growth rate of the self-replicator system that, using (2.2), is defined as

$$\frac{\dot{\mathcal{V}}}{\mathcal{V}} = \frac{\dot{M} + \dot{R}}{M + R} = \beta v_R(p, r) - D \quad \rightarrow \quad \mu(t) \doteq \left.\frac{\dot{\mathcal{V}}}{\mathcal{V}}\right|_{F=0} = \beta v_R(p, r). \tag{2.5}$$

The latter basically means that the growth rate is defined as the relative variation of cell volume ($\dot{\mathcal{V}}/\mathcal{V}$) when there is no volumetric flow rate. The synthesis rates are modeled as Michaelis-Menten kinetics

$$\begin{aligned} v_M(s,m) &\doteq k_M\, m\, \frac{s}{K_M+s}, \\ v_R(p,r) &\doteq k_R\, r\, \frac{p}{K_R+p}, \\ v_X(p,m) &\doteq k_X\, m\, \frac{p}{K_X+p}, \end{aligned} \tag{2.6}$$

with rate constants k_M, k_R, k_X [h^{-1}] and half-saturation constants K_M, K_R, K_X [g L^{-1}]. Moreover, quantities and time-scale are nondimensionalized in order to simplify the analysis, by defining appropriate new variables and constants

$$\begin{gathered} \hat{t} \doteq k_R\, t, \quad \hat{p} \doteq \beta\, p, \quad \hat{r} \doteq \beta\, r, \quad \hat{x} \doteq \beta\, x \quad \hat{X} \doteq \beta\, X, \\ K \doteq \beta\, K_R, \quad K_1 \doteq \beta K_X, \quad K_2 \doteq \beta K_M \\ k_1 \doteq \tfrac{k_X}{k_R}, \quad k_2 \doteq \tfrac{k_M}{k_R}. \end{gathered}$$

By replacing all variables and dropping all hats, system (2.4) becomes

$$\begin{cases} \dot{s} = D(s_{in}-s) - k_2 \dfrac{(1-r)s}{(K_2+s)}\dfrac{\mathcal{V}}{\mathcal{V}_{ext}}, \\ \dot{p} = k_2 \dfrac{s\,(1-r)}{K_2+s} - k_1 \dfrac{p(1-r)}{K_1+p} - (p+1)\,\dfrac{pr}{K+p}, \\ \dot{r} = (u-r)\,\dfrac{pr}{K+p}, \\ \dot{x} = k_1 \dfrac{p(1-r)}{K_1+p} - \dfrac{pr}{K+p}x, \\ \dot{\mathcal{V}} = \left(\dfrac{pr}{K+p} - D\right)\mathcal{V}. \end{cases} \tag{2.7}$$

where the dynamical expression of m has been removed since $m = 1-r$. The analysis of model (2.7) is part of a more recent work [AYT2], which is described in this thesis in Chapter 4. In the next section, we present the particular case where $D = 0$, meaning that there is no substrate inflow, and no biomass output.

2.2.2 Substrate depletion

In this case, the substrate is not replenished from the outside, a situation that can describe batch cultivation. The dynamical equation for s becomes

$$\dot{s} = -k_2 \frac{s(1-r)}{K_2 + s} \frac{\mathcal{V}}{\mathcal{V}_{ext}}. \tag{2.8}$$

The optimal product maximization problem was partly analyzed for this particular case in [43] where, due to the complexity of the computations in the PMP, most of the analysis was performed through a numerical approach. In Chapter 3, a more detailed study of this model and its associated optimal control problems is performed. To allow an analytical study, we simplify the system by assuming there is substrate in excess and the depletion occurs slowly enough (if, for example, $\mathcal{V} \ll \mathcal{V}_{ext}$). Thus, it is possible to exclude the dynamics of s from the analysis, which yields the model of interest in this paper.

2.2.3 Constant environmental conditions

The environmental conditions can be modeled as constant over time as a result of s being constant due some external regulation of the variable, but it can also describe an environment with abundant substrate, where $s \gg K_M$ in (2.6). A constant value

$$E_M \doteq k_2 \frac{s}{K_2 + s}$$

is defined, such that model (2.8) becomes

$$\begin{cases} \dot{p} = E_M(1-r) - k_1 \dfrac{p(1-r)}{K_1 + p} - (p+1)\dfrac{pr}{K+p}, \\ \dot{r} = (u-r)\dfrac{pr}{K+p}, \\ \dot{x} = k_1 \dfrac{p(1-r)}{K_1 + p} - \dfrac{pr}{K+p}x, \\ \dot{\mathcal{V}} = \dfrac{pr}{K+p}\mathcal{V}. \end{cases} \tag{S}$$

This assumption can also represent fed-batch cultivation where the nutrient concentration is maintained high enough in order to achieve exponential growth rate.

For this particular case study, we state the problem of product maximization as an OCP, and characterize the singular arcs of the solution.

2.2.4 Allocation problem with no metabolite production

By overriding the production of the compound X, the model becomes the coarse-grained self-replicator depicted in Figure 2.1a. This is a particular case of the model (S), when $k_1 = 0$, and the external control $I(t)$ introduced in (2.1) is overridden (so that $u(t) = \alpha(t)$),

$$\begin{cases} \dot{p} = E_M(1-r) - (p+1)\dfrac{pr}{K+p}, \\ \dot{r} = (\alpha - r)\dfrac{pr}{K+p}. \end{cases} \tag{2.9}$$

The problem of biomass maximization by natural mechanisms has been extensively analyzed for this model in [30], so in this paper we merely recall its dynamics and the OCP associated, in order to compare the solution with that of the metabolite production problem.

2.3 Dynamical analysis of model (S)

In this section, we perform a local and global analysis of model (S). Given the nullcline p that satisfies the equation

$$0 = E_M(1-r) - w_X(p)(1-r) - w_R(p)(p+1)r,$$

with

$$w_R(p) = \frac{p}{K+p}, \quad w_X(p) = \frac{k_1 p}{K_1 + p},$$

we can define the function

$$r_n(p) \doteq \frac{E_M - w_X(p)}{E_M - w_X(p) + w_R(p)(p+1)} \in (0,1], \tag{2.10}$$

which is monotonically decreasing w.r.t. p, such that the equation for p can be rewritten as

$$\frac{dp}{dt} = (r_n(p) - r)[E_M - w_X(p) + w_R(p)(p+1)],$$

leading to the system

$$\begin{cases} \dfrac{dp}{dt} = (r_n(p) - r)[E_M - w_X(p) + w_R(p)(p+1)], \\ \dfrac{dr}{dt} = (u - r)w_R(p)r, \\ \dfrac{dx}{dt} = w_X(p)(1 - r) - w_R(p)rx, \\ \dfrac{d\mathcal{V}}{dt} = w_R(p)r\mathcal{V}. \end{cases} \tag{S_1}$$

2.3.1 Global behavior

Given that the volume $\mathcal{V}$ admits no steady state except the trivial $\mathcal{V} = 0$, we will study the dynamical behavior of the system without considering this variable. Additionally, since the dynamics of variables (p, r) do not depend on x, we will resort to well-known decomposition techniques presented in [47] to study the stability properties of S_1 through its subsystems. More precisely, we will separate the system into two subsystems arranged in a hierarchical form

$$\begin{aligned} z_1 &= \begin{bmatrix} p \\ r \end{bmatrix}, \quad \frac{dz_1}{dt} = f_1(z_1), \\ z_2 &= \begin{bmatrix} x \end{bmatrix}, \quad \frac{dz_2}{dt} = f_2(z_1, z_2), \end{aligned}$$

where f_1 depends only on z_1. Therefore, in the present section, we will study the stability of the subsystem given by variables (p, r)

$$\begin{cases} \dfrac{dp}{dt} = (r_n(p) - r)[E_M - w_X(p) + w_R(p)(p+1)] \\ \dfrac{dr}{dt} = (u - r)w_R(p)r \end{cases} \tag{S_1'}$$

to conclude, towards the end of the section, on the stability of the global system (S_1).

Lemma 2.3.1. *The set,*

$$\Gamma \doteq \left\{(p,r) \in \mathbb{R}^2 \,:\, p \geq 0,\, 1 \geq r \geq 0\right\},$$

is positively invariant in (S_1') *for the initial value problem.*

Proof. Let us analyze the boundaries of Γ,

$$\begin{aligned}
&\dot{p}|_{p=0} = E_M(1-r) \geq 0 \quad &\Rightarrow \quad &p = 0 \text{ is repulsive or invariant.}\\
&\dot{r}|_{r=0} = 0 \quad &\Rightarrow \quad &r = 0 \text{ is invariant.}\\
&\dot{r}|_{r=1} = (u-1)w_R(p) \leq 0 \quad &\Rightarrow \quad &r = 1 \text{ is repulsive or invariant.}
\end{aligned}$$

□

In order to study the steady-state behavior of the system, we fix $u(t) = \bar{u} \in (0,1)$.

Lemma 2.3.2. *The system admits two equilibria:*

- *The interior equilibrium:* $E_i = (p_i, \bar{u})$
- *The no-growth equilibrium:* $E_w = (0,1)$

with p_i *solution of the equation*

$$E_M(1-\bar{u}) - (p+1)w_R(p)\bar{u} - w_X(p)(1-\bar{u}) = 0 \tag{2.11}$$

Proof. To prove uniqueness of E_i, it is sufficient to prove that the value of p_i is unique in (2.11). We define the constant value,

$$c_1 \doteq E_M(1-\bar{u}) \geq 0,$$

replace them in (2.11) and define the function f as

$$f(p) \doteq w_X(p)(1-\bar{u}) + (p+1)\, w_R(p)\bar{u},$$

with the following properties

$$f(0) = 0, \quad \lim_{p\to\infty} f(p) = \infty, \quad \frac{\partial f(p)}{\partial p} > 0.$$

Therefore, $f(p) = c_1$ has a unique solution p_i, which also leads to a unique value x_i and thus the equilibrium E_i is unique. □

2.3.2 Local stability of equilibria

We first write the Jacobian matrix

$$J = \begin{bmatrix} -(p+1)w'_R(p)r - w_R(p)r - w'_X(p)(1-r) & -E_M - w_R(p)(p+1) + w_X(p) \\ w'_R(p)r(\bar{u}-r) & -w_R(p)r + (\bar{u}-r)w_R(p) \end{bmatrix}, \tag{2.12}$$

where

$$w'_X(p) \doteq \frac{\partial w_X(p)}{\partial p}, \quad w'_R(p) \doteq \frac{\partial w_R(p)}{\partial p}.$$

Equilibrium E_i

Replacing the values of E_i in (2.12)

$$J_i = \begin{bmatrix} -(p+1)w'_R(p)\bar{u} - w_R(p)\bar{u} - w'_X(p)(1-\bar{u}) & -E_M - w_R(p)(p+1) - w_X(p) \\ 0 & -w_R(p)\bar{u} \end{bmatrix},$$

with eigenvalues

$$\lambda = (-(p+1)w'_R(p)\bar{u} - w_R(p)\bar{u} - w'_X(p)(1-\bar{u}), -w_R(p)\bar{u}),$$

and so the equilibrium is locally stable.

Equilibrium E_w

In E_w, one has $w_R(p) = w_X(p) = 0$ and $w'_R(p)r = \frac{1}{K}$, and so the Jacobian matrix becomes

$$J_w = \begin{bmatrix} -\frac{1}{K} & -E_M \\ -\frac{1}{K}(1-\bar{u}) & 0 \end{bmatrix},$$

with characteristic polynomial

$$P(\lambda) = \left(\lambda + \frac{1}{K}\right)\lambda - E_M\frac{1}{K}(1-\bar{u}) = \lambda^2 + \lambda\frac{1}{K} - E_M\frac{1}{K}(1-\bar{u}).$$

Then, the eigenvalues are given by

$$\lambda_1 = \frac{-\frac{1}{K} - c_1}{2} < 0, \quad \lambda_2 = \frac{-\frac{1}{K} + c_1}{2} > 0, \quad \textit{with } c_1 = \sqrt{\frac{1}{K}^2 + 4E_M\frac{1}{K}(1-\bar{u})}$$

showing that the equilibrium is a saddle point. Now, let us find the eigenvector $v_1 = [p_1, r_1]^T$ associated to the stable eigenvalue $\lambda_1 < 0$,

$$\begin{bmatrix} \lambda_1 + \frac{1}{K} & E_M \\ \frac{1}{K}(1-\bar{u}) & \lambda_1 \end{bmatrix}\begin{bmatrix} p_1 \\ r_1 \end{bmatrix} = 0, \quad \Rightarrow \quad \begin{cases} (\lambda_1 + w'_R(p)r)p_1 + E_M r_1 = 0 \\ w'_R(p)r(1-\bar{u})p_1 + \lambda_1 r_1 = 0 \end{cases}$$

Solving the system yields the relations

$$p_1 = -r_1\frac{E_M}{\lambda_1 + w'_R(p)r}, \quad p_1 = -r_1\frac{\lambda_1}{w'_R(p)r(1-\bar{u})},$$

where it can be seen that both equations are equivalent

$$\frac{E_M\frac{1}{K}(1-\bar{u})}{(\lambda_1 + \frac{1}{K})\lambda_1} = \frac{4E_M\frac{1}{K}(1-\bar{u})}{-(\frac{1}{K} - c_1)(\frac{1}{K} + c_1)} = \frac{4E_M\frac{1}{K}(1-\bar{u})}{\left(\frac{1}{K}\right)^2 + 4E_M\frac{1}{K}(1-\bar{u}) - \left(\frac{1}{K}\right)^2} = 1,$$

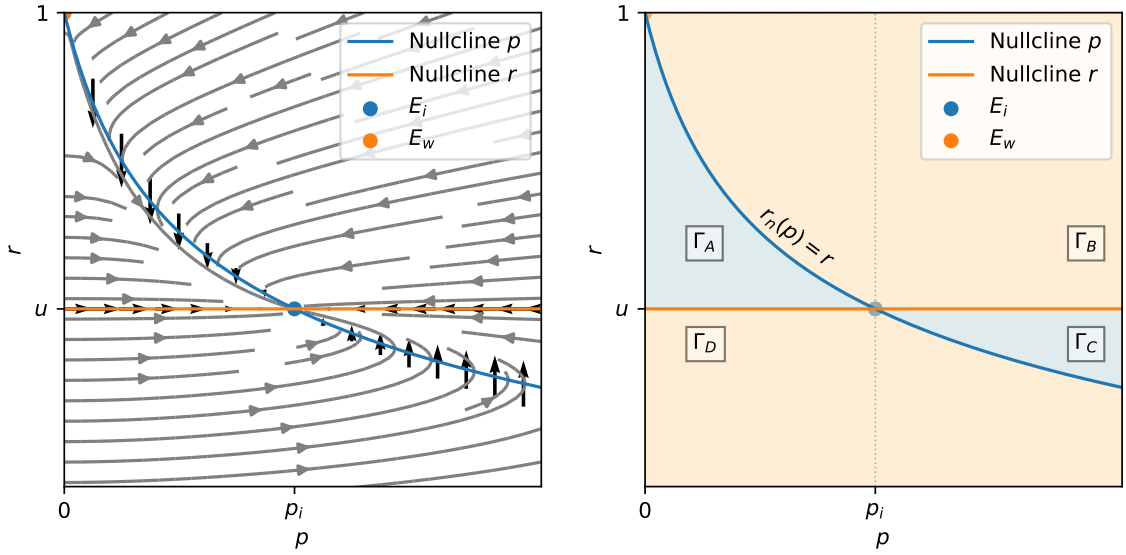
and so, since $\lambda_1 < 0$, p_1 and r_1 have the same sign, which means that the eigenvector associated to the stable eigenvalue points to the first and third quadrants, which is outside the invariant set Γ.

2.3.3 Global analysis of S_1'

The analysis of the global behavior consists on dividing Γ in four regions delimited by the system nullclines (as seen in Figure A.2) and studying how the system behaves in each region separately. We start by stating what has been showed in Section 2.3.2:

Lemma 2.3.3. *A trajectory originating in* $\Gamma \setminus \{E_w\}$ *can't converge to the unstable equilibria* E_w.

Proof. As shown in Section 2.3.2, E_w is a saddle point, and the eigenvector associated to the stable eigenvalue points to the first and third quadrant w.r.t. $E_w = (0, 1)$. As this is the only possible direction through which the saddle point can be attained, the equilibrium is not an attractor in Γ, except for trajectories starting in E_w. □



(a) Trajectories of S_1'.

(b) Division of Γ in 4 regions.

Figure 2.2: Phase plane of system S_1'.

Lemma 2.3.4. *Every solution of the initial value problem starting in the sets,*

$$\Gamma_A \doteq \left\{(p, r) \in \mathbb{R}^2 \,:\, p_i \geq p > 0,\, r_n(p) \geq r \geq \bar{u}\right\},$$
$$\Gamma_C \doteq \left\{(p, r) \in \mathbb{R}^2 \,:\, p \geq p_i,\, \bar{u} \geq r \geq r_n(p)\right\},$$

converges asymptotically to the equilibrium E_i.

Proof. First, let us show that Γ_A is positively invariant to the system

$$\begin{aligned}
&\dot{p}|_{p=0} = E_M(1-r) > 0,\ \forall r \in (\bar{u}, 1) && \Rightarrow\ p = 0 \textit{ is repulsive.}\\
&\dot{r}|_{r=r_n(p)} = (\bar{u} - r_n(p))w_R(p)r_n(p) < 0,\ \forall p \in (0, p_i) && \Rightarrow\ r = r_n(p) \textit{ is repulsive.}\\
&\dot{r}|_{r=\bar{u}} = 0 && \Rightarrow\ r = \bar{u} \textit{ is invariant.}
\end{aligned}$$

as well as Γ_C

$$\begin{aligned}
&\dot{r}|_{r=r_n(p)} = (\bar{u} - r_n(p))w_R(p)r_n(p) > 0,\ \forall p > p_i && \Rightarrow\ r = r_n(p) \textit{ is repulsive.}\\
&\dot{r}|_{r=\bar{u}} = 0 && \Rightarrow\ r = \bar{u} \textit{ is invariant.}
\end{aligned}$$

Then, asymptotic stability of E_i can be proved using the following simple quadratic Lyapunov function

$$V_{A,C}(t) \doteq \frac{1}{2}(p_i - p)^2 + \frac{1}{2}(\bar{u} - r)^2 \ > \ 0, \quad \forall (p, r) \in (\Gamma_A \cup \Gamma_C) \setminus \{E_i\},$$

such that

$$\frac{d}{dt}V_{A,C}(t) = -(p_i - p)\dot{p} - (\bar{u} - r)^2 w_R(p) \ < \ 0, \quad \forall (p, r) \in (\Gamma_A \cup \Gamma_C) \setminus \{E_i\},$$

which verifies the inequality given that in Γ_A, $p_i > p$ and $\dot{p} > 0$, and in Γ_C, $p_i < p$ and $\dot{p} < 0$. The equilibrium E_i is the only point where the derivative $\frac{d}{dt}V_{A,C}(t) = 0$, as well as the largest invariant set, which confirms through LaSalle theorem [48] that every solution starting in Γ_A and Γ_C approaches E_i as $t \to \infty$. □

Lemma 2.3.5. *Every solution of the initial value problem starting in the sets,*

$$\begin{aligned}
\Gamma_B &\doteq \left\{(p, r) \in \mathbb{R}^2 \ :\ p > 0,\ 1 \geq r > \max\left\{\bar{u}, r_n(p)\right\}\right\},\\
\Gamma_D &\doteq \left\{(p, r) \in \mathbb{R}^2 \ :\ p > 0,\ \min\left\{\bar{u}, r_n(p)\right\} > r > 0\right\},
\end{aligned}$$

converges asymptotically to the equilibrium E_i.

Proof. Every point in Γ_B meets $r_n(p) < r$ and $r > \bar{u}$, meaning that the vector field have signs $\dot{p} < 0$, $\dot{r} < 0$. Analogously, every point in Γ_D meets $r_n(p) > r$ and $r < \bar{u}$, so the vector field has signs $\dot{p} > 0$, $\dot{r} > 0$. As stated in Lemma 2.3.3, no trajectory can converge towards E_w, so it can go whether towards the equilibrium E_i, or towards any of both nullclines, where eventually it will converge to E_i as

proven in Lemma 2.3.4. □

Lemmas 2.3.4 and 2.3.5 lead us to the following theorem.

Theorem 2.3.1. *Every solution of the initial value problem with initial conditions* $p(0) > 0$ *and* $r(0) \in (0, 1)$ *converges asymptotically to the equilibrium* E_i *as* $t \to \infty$.

Connection with the original system

As the global stability of the equilibrium E_i in subsystem S_1' has been established, we can conclude on the global behavior of the original system S_1. For that, we will state a simplified version of the theorem presented in [47] that relates the stability of a hierarchical system

$$S_z : \begin{cases} \dfrac{dz_1}{dt} = f_1(z_1) \\ \dfrac{dz_2}{dt} = f_2(z_1, z_2) \end{cases}$$

with equilibrium $\bar{z} \doteq [\bar{z}_1, \bar{z}_2]$, with the stability of its so-called isolated subsystems

$$S_{z_1} : \frac{dz_1}{dt} = f_1(z_1), \quad S_{z_2} : \frac{dz_2}{dt} = f_2(\bar{z}_1, z_2),$$

where, in the second system S_{z_2}, z_1 is set to $\bar{z}_1$.

Theorem 2.3.2. **[47]** *Given* f_1 *and* f_2 *continuous, autonomous and continuously differentiable with respect to all variables, satisfying*

$$\sup_{t>0} \sup_{z_1} ||\nabla_{z_1} f_1(z_1)|| < \infty,$$
$$\sup_{t>0} \sup_{(z_1,z_2)} ||\nabla_{(z_1,z_2)} f_2(z_1, z_2)|| < \infty,$$

and $f_1(\bar{z}) = f_2(\bar{z}) = 0$, *then* $z = \bar{z}$ *is a globally asymptotically stable equilibrium point of* S_z *if and only if* $z_1 = \bar{z}_1$ *and* $z_2 = \bar{z}_2$ *are globally asymptotically stable equilibrium points of isolated subsystems* S_{z_1} *and* S_{z_2} *respectively.*

In the present case, as the asymptotic stability of the equilibrium E_i in S_1' has been already proven, and given that all terms in the Jacobian matrix (2.12) are

bounded for every value of Γ, we proceed to evaluate the dynamical expression of x in $(p, r) = E_i$

$$\frac{dx}{dt} = w_X(p_i)(1 - \bar{u}) - w_R(p_i)\bar{u}x,$$

showing that the $x = x_i$ is globally asymptotically stable, with

$$x_i \doteq \frac{w_X(p_i)(1 - \bar{u})}{w_R(p_i)\bar{u}}.$$

The latter is formalized in the following theorem.

Theorem 2.3.3. *In every solution of the initial value problem in system S_1 starting in $p(0) > 0$, $r(0) \in (0, 1)$, $x(0) \geq 0$ and $\mathcal{V}(0) > 0$, when $t \to \infty$,*

- *(p, r) converge asymptotically to the equilibrium E_i.*
- *The biomass volume $\mathcal{V}$ grows exponentially with rate*

$$\mu(p_i, \bar{u}) = w_R(p)\bar{u} = \frac{p_i\bar{u}}{K + p_i}$$

- *The quantity x converges asymptotically to*

$$x_i = \frac{w_X(p_i)(1 - \bar{u})}{w_R(p_i)\bar{u}}$$

as the total mass of metabolite X also grows exponentially.

2.4 Optimal control problem

2.4.1 Dynamical biomass maximization in Model (2.9)

In model (2.9), the control input is the natural allocation $\alpha(t)$. Let $\mathcal{U}$ be the set of admissible controllers, which are Lebesgue measurable real-valued functions defined on the interval $[0, T]$ and satisfying the constraint $\alpha \in [0, 1]$. The maximization of

the biomass can be formulated as an OCP

$$\begin{aligned} &\textit{maximize} \quad J_\mu(\alpha) = \int_0^T \mu(p,r)\,\mathrm{d}t \\ &\textit{subject to} \quad \alpha(\cdot) \in \mathcal{U} \end{aligned}$$

with

$$\mu(p,r) = \frac{pr}{K+p}$$

being the growth rate defined in (2.5). Given the state $\varphi = (p,r)$, and according to Pontryagin Maximum Principle, the Hamiltonian is defined as

$$H_A(\varphi,\lambda,\alpha) \doteq \lambda^0\mu + \langle\lambda, F(\varphi,\alpha)\rangle, \tag{2.13}$$

where F denotes the right-hand side of (2.9) and where $\lambda = (\lambda_p, \lambda_r)$ is the adjoint state. We assume that the process (α,φ,λ) satisfying PMP conditions is a normal extremal and set $\lambda^0 = -1$, which yields

$$H_A = \lambda_p E_M(1-r) - \frac{pr^2}{K+p}[\lambda_p(1+p) + \lambda_r(r-\alpha) - 1].$$

Since the Hamiltonian depends linearly on the control α, the optimal solution is

$$\alpha_{opt}(t) = \begin{cases} 0, & \text{if } \phi(\cdot) < 0, \\ 1, & \text{if } \phi(\cdot) > 0. \end{cases} \tag{2.14}$$

being

$$\phi(\cdot) = \lambda_r \frac{pr}{K+p}$$

the switching function.

2.4.2 Product maximization in Model (S)

The problem can be formulated as

$$\begin{aligned} &\textit{maximize} \quad J_X(u) = X(T) - X_0 \\ &\textit{subject to} \quad u(\cdot) \in \mathcal{U}, \end{aligned}$$

where $X_0 \doteq X(0)$ is the initial mass of metabolite. Using the dynamical equation for X in (S), the criterion can be written as

$$J_X(u) = \int_0^T k_1 \frac{p(1-r)}{K_1 + p} \, \mathcal{V} \, \mathrm{d}t.$$

It can be proved that, if $X_0 > 0$, the cost function can be equivalently reformulated as

$$J_X(u) = \int_0^T \frac{k_1}{x} \frac{p(1-r)}{K_1 + p} \mathrm{d}t. \tag{2.15}$$

Static problem

We are interested in the steady state that maximizes the metabolite production in terms of the constant allocation $\bar{u}$. The latter is represented by the integrand of the cost function (2.15). The problem can be stated as

$$\begin{aligned} &\textit{maximize} \quad && J(\bar{u}, p_i(\bar{u})) = \frac{w_X(p)(1-\bar{u})}{x_i}, \\ &\textit{subject to} \quad && \bar{u} \in (0,1). \end{aligned} \tag{2.16}$$

While an explicit calculation of the solution of the latter problem is rather hard to obtain, we will focus on showing the existence and uniqueness of such solution. First, we notice that the cost can be equivalently rewritten as

$$J(\bar{u}, p_i(\bar{u})) = w_R(p)\bar{u} = \frac{p_i(\bar{u})\bar{u}}{K + p_i(\bar{u})},$$

which shows that maximizing the metabolite production at steady state is equivalent to maximizing the biomass production. Recalling the function $r_n(p)$ defined in (2.10), we can express $\bar{u}$ in terms of p and maximize $J(p, \bar{u}(p))$ with respect to p (as the steady-state value $p_i(\bar{u})$ is strictly decreasing w.r.t. $\bar{u}$). Moreover, for the sake of simplicity, rather than maximizing the growth rate $w_R(p)$, we will minimize its inverse, which yields the optimization problem

$$\begin{aligned} &\textit{minimize} \quad && \hat{J}(p, \bar{u}(p)) \doteq \frac{1}{J(p, \bar{u}(p))} = \frac{K + p}{p \, \bar{u}(p)}, \\ &\textit{subject to} \quad && p \in (0, \infty). \end{aligned} \tag{2.17}$$

The objective can be further developed by replacing $\bar{u}(p)$ by its expression, which yields

$$\hat{J}(p) = \frac{K}{p} + 1 + \frac{(p+1)(K_1+p)}{E_M K_1 + p(E_M - k_1)}.$$

The following assumption is required to ensure the well-posedness of the problem:

Assumption 2.4.1. $E_M > k_1$.

The latter assumption implies that the maximal substrate intake is strictly greater than the maximal metabolite production. Note that this assumption also includes model (2.9), where $k_1 = 0$, as a particular case of the current analysis. For notation purposes, we define the constants

$$c_1 \doteq \frac{E_M K_1}{E_M - k_1} > 0, \quad c_2 \doteq \frac{(1-c_1)(K_1 - c_1)}{E_M - k_1}, \tag{2.18}$$

such that

$$K_1 - c_1 = K_1 \left(1 - \frac{E_M}{E_M - k_1}\right) < 0.$$

Then, the objective function becomes

$$\hat{J}(p) = \frac{K}{p} + 1 + \frac{p+1}{E_M - k_1} + \frac{K_1 - c_1}{E_M - k_1} + \frac{c_2}{p + c_1}. \tag{2.19}$$

It can be seen from (2.19) that, at the boundaries, the function $\hat{J}(p)$ verifies

$$\lim_{p \to 0} \hat{J}(p) = \infty, \quad \lim_{p \to \infty} \hat{J}(p) = \infty,$$

which is coherent with the fact that the growth rate $\mu(p, r) = 0$ at the boundaries. The latter, and the continuity of $\hat{J}(p)$, implies the existence of, at least, one solution. Additionally, it means that the optimal value p is interior, so in order to prove uniqueness of the solution, it suffices to show that the derivative is cancelled in only one value of p by solving

$$\frac{\partial}{\partial p}\hat{J}(p) = -\frac{K}{p^2} + \frac{1}{E_M - k_1} - \frac{c_2}{(p + c_1)^2} = 0. \tag{2.20}$$

For that, there are two possible cases to explore in terms of the sign of c_2.

Case $c_2 \geq 0$

It can be seen from (2.18) that if $1 - c_1 \leq 0$, then $c_2 \geq 0$ and so we can rewrite (2.20) as

$$\frac{1}{E_M - k_1} = \underbrace{\frac{K}{p^2} + \frac{c_2}{(p + c_1)^2}}_{f_1(p)}$$

where

$$\lim_{p \to 0} f_1(p) = \infty, \quad \lim_{p \to \infty} f_1(p) = 0, \quad \frac{\partial}{\partial p} f_1(p) = -\frac{2K}{p^3} - \frac{2c_2}{(p + c_1)^3} < 0,$$

which means that the equation $f_1(p) = \frac{1}{E_M - k_1}$ has a unique solution. Then, there is a single value of p that cancels (2.20), and so the minimization problem (2.17) has a unique solution.

Case $c_2 < 0$

In the case where $c_2 < 0$ due to $1 - c_1 > 0$, we can rewrite the equality (2.20) as

$$K = \underbrace{\frac{p^2}{E_M - k_1} - c_2 \left(\frac{p}{p + c_1} \right)^2}_{f_2(p)}$$

where

$$f_2(0) = 0, \quad \lim_{p \to \infty} f_2(p) = \infty, \quad \frac{\partial}{\partial p} f_2(p) = \frac{2p}{E_M - k_1} - \frac{2c_2\, p}{(p + c_1)^2} \left(1 - \frac{p}{p + c_1} \right) > 0,$$

showing that the equation $f_2(p) = K$ has a unique solution, which again means that the minimization problem (2.17) has a unique solution. The latter can be formalized in the following theorem:

Theorem 2.4.1. *Under assumption 2.4.1, there exists a unique constant allocation* $\bar{u} \in (0, 1)$ *that maximizes the steady-state metabolite production of Problem* (2.16) *(equivalent to the steady-state biomass production).*

Dynamic problem

For the sake of convenience, we perform a change of variables $y \doteq \ln x$, so that the criterion becomes

$$J_X(u) = \int_0^T k_1 e^{-y} \frac{p(1-r)}{K_1 + p} \mathrm{d}t,$$

with new initial condition $y(0) = y_0 = \ln x(0)$. Given the state $\varphi = (p, r, y)$ and adjoint state $\lambda = (\lambda_p, \lambda_r, \lambda_y)$, we again assume that the extremal triple (φ, u, λ) satisfies PMP in normal form, and set $\lambda^0 = -1$. The Hamiltonian becomes

$$H_B = E_M \lambda_p (1-r) - \frac{((\lambda_y + 1)e^{-y} + \lambda_p) k_1 p (1-r)}{K_1 + p} - \frac{(\lambda_p(p+1) + \lambda_r(r-u) + \lambda_y) pr}{K + p}, \tag{2.21}$$

and the adjoint system $\dot{\lambda} = -\frac{\partial H_B}{\partial \varphi}$ writes

$$\begin{cases} \dot{\lambda_p} = \dfrac{((\lambda_y + 1)e^{-y} - \lambda_p) k_1 (1-r)\left(\dfrac{p}{K_1 + p} - 1\right)}{K_1 + p} + \dfrac{\lambda_p p r}{K + p} \\ \qquad\qquad - \dfrac{(\lambda_p(p+1) + \lambda_r(r-u) + \lambda_y) r \left(\dfrac{p}{K + p} - 1\right)}{K + p}, \\ \dot{\lambda_r} = E_m \lambda_p + \dfrac{((\lambda_y + 1)e^{-y} - \lambda_p) k_1 p}{K_1 + p} \\ \qquad\qquad + \dfrac{(\lambda_p(p+1) + \lambda_r(2\,r - u) + \lambda_y) p}{K + p}, \\ \dot{\lambda_y} = -\dfrac{k_1 (\lambda_y + 1) p (r-1) e^{-y}}{K_1 + p}. \end{cases}$$

As in the previous model, the Hamiltonian for this OCP still depends linearly on u and so the solution is again as (2.14), with the same switching function $\phi(\cdot) = \lambda_r \frac{pr}{K+p}$.

2.4.3 Characterization of singular regimes

Singular trajectories play a major role in optimal control theory [49, 50], and their characterization is a necessary step towards a complete description of the optimal solution. In both explored OCPs, and as expected in linear optimization problems,

maximization of the Hamiltonian gives no information when the switching function vanishes during a whole interval of time. In this case, the OCP has a *singular arc*, that can be obtained by computing the successive time derivatives of the switching function until it is possible to obtain an explicit expression of the control. In the case of the simple resource allocation problem solved in 2.4.1, vanishing of the switching function implies than $\lambda_r = 0$, so $\dot{\lambda}_r = 0$ along the singular arc. Since the Hamiltonian is preserved along an extremal trajectory, λ_p is constant along a singular arc, which means that such arc corresponds to a steady state of the system. Moreover, it has been shown in [30] that the singular arc is in fact the optimal steady state solution of the static problem, and so the optimal control is $\alpha(t) = \alpha_{opt}$. Additionally, the arc is of order two, which implies that it has to be entered and exited through a *chattering arc* [51], i.e. an arc with an infinite number of switchings. We now focus on the model (S) and provide a detailed computation of the corresponding singular arcs. The Hamiltonian (2.21) can be expressed as

$$H_B = H_0 + u\, H_1$$

where

$$\begin{aligned} H_0 &= E_m \lambda_p (1-r) + \frac{((\lambda_y+1)e^{-y} - \lambda_p) k_1 p (1-r)}{K_1 + p} \\ &\quad - \frac{(\lambda_p(p+1) + \lambda_r r + \lambda_y) p r}{K + p}\,, \\ H_1 &= \phi(t) = \frac{\lambda_r p r}{K+p}. \end{aligned}$$

Assume H_1 vanishes on a whole sub-interval $\mathcal{I} = [t_1, t_2] \subset [0, T]$. The switching surface is the set (here $n = 3$)

$$\Sigma \doteq \left\{ (\varphi, \lambda) \in \mathbb{R}^{2n} \,|\, H_1 = 0 \right\}.$$

The time derivative of H_1 along the extremal is equal to the Poisson bracket $\{H_0, H_1\}$, denoted as H_{01}, since

$$\begin{aligned} \dot{H}_1 &= \frac{\partial H_1}{\partial \varphi}\dot{\varphi} + \frac{\partial H_1}{\partial \lambda}\dot{\lambda} = \sum_{i=1}^{n} \left(\frac{\partial H_B}{\partial \lambda_i}\frac{\partial H_1}{\partial \varphi_i} - \frac{\partial H_B}{\partial \varphi_i}\frac{\partial H_1}{\partial \lambda_i} \right) \\ &= \{H_B, H_1\} = \{H_0 + uH_1, H_1\} = \{H_0, H_1\} = H_{01}. \end{aligned}$$

Using the latter, we proceed to compute the subsequent derivatives of H_1 with respect to time, which should also vanish along the singular arc, until it is possible to compute the singular control u:

$$\begin{aligned} 0 = H_{01} = &-\tfrac{pr^2}{(K+p)^2}\lambda_r p + \tfrac{\lambda_r p^2 r^2}{(K+p)^2} \\ &-\left(E_m(r-1) - \tfrac{k_1 p(r-1)}{K_1+p} + \tfrac{(p+1)pr}{K+p}\right)\left(1 - \tfrac{pr}{K+p}\right)\tfrac{\lambda_r r}{K+p} \\ &+\left(E_m\lambda_p + \tfrac{\left((\lambda_y+1)e^{-y}-\lambda_p\right)k_1 p}{K_1+p} + \tfrac{(\lambda_p(p+1)+\lambda_r r+\lambda_y)p}{K+p}\right)\tfrac{pr}{K+p}, \end{aligned} \tag{2.22}$$

and evaluating (2.22) for $z \in \Sigma$ (which, in this case, is just setting $\lambda_r = 0$) yields

$$0 = H_{01}(z) = \frac{\left(E_m\lambda_p + \dfrac{((\lambda_y+1)e^{-y}-\lambda_p)k_1 p}{K_1+p} + \dfrac{(\lambda_p(p+1)+\lambda_y)p}{K+p}\right)pr}{K+p}. \tag{2.23}$$

From (2.23) we have the following new condition:

$$E_m\lambda_p + \frac{((\lambda_y+1)e^{-y}-\lambda_p)k_1 p}{K_1+p} + \frac{(\lambda_p(p+1)+\lambda_y)p}{K+p} = 0,$$

which defines a subset of the switching surface Σ given by

$$\Sigma' \doteq \left\{(\varphi,\lambda) \in \mathbb{R}^{2n} \mid H_{01} = 0\right\} \cap \Sigma.$$

Similarly, the second derivative can be computed as

$$\ddot{H}_1 = \dot{H}_{01} = \{H_B, H_{01}\} = \{H_0, H_{01}\} + u\{H_1, H_{01}\} = H_{001} + uH_{101},$$

and so one has

$$\begin{aligned} H_{101} = &-pr\phi_7\frac{\lambda_r r}{(K+p)^2} - \frac{\phi_3 r\left(\dfrac{p}{K+p}-1\right)}{(K+p)^2}\lambda_r p + \frac{\left(E_m\lambda_p + \dfrac{k_1 p\phi_2}{K_1+p} + \dfrac{p\phi_1}{K+p}\right)}{(K+p)^2}p^2 r \\ &+\lambda_r p\frac{p^2r^2}{(K+p)^3} + pr\phi_7\frac{\lambda_r pr}{(K+p)^3} + \frac{\lambda_r\phi_3\left(\frac{p}{K+p}-1\right)}{(K+p)^2}pr + \frac{\lambda_r\phi_7\left(\dfrac{p}{K+p}-1\right)}{(K+p)^2}pr^2 \end{aligned}$$

where

$$\begin{aligned}
\phi_1 &= \lambda_p(p+1) + 2\,\lambda_r r + \lambda_y,\\
\phi_2 &= (\lambda_y+1)e^{-y} - \lambda_p,\\
\phi_3 &= E_m(r-1) - \frac{k_1 p(r-1)}{K_1+p} + \frac{(p+1)pr}{K+p},\\
\phi_7 &= \frac{\phi_3 + E_m - \dfrac{k_1 p}{K_1+p}}{r}.
\end{aligned}$$

Evaluating in $z \in \Sigma'$ yields

$$\begin{aligned}
H_{101}(z) &= \frac{\left(E_m\lambda_p + \dfrac{((\lambda_y+1)e^{-y}-\lambda_p)k_1 p}{K_1+p} + \dfrac{(\lambda_p(p+1)+\lambda_y)p}{K+p}\right)p^2 r}{(K+p)^2}\\
&= H_{01}(z)\frac{p}{K+p} = 0
\end{aligned}$$

on Σ', entailing that singular arcs must be at least of order two. Since $\ddot{H}_1$ should vanish along the singular arc, we calculate H_{001} and evaluate it in $z \in \Sigma'$:

$$\begin{aligned}
0 = H_{001}(z) = {} & \frac{k_1 p\phi_2\phi_7\left(\dfrac{p}{K_1+p}-1\right)}{(K+p)(K_1+p)} + \frac{k_1(\lambda_y+1)p^3 e^{-y}}{(K+p)^2(K_1+p)}\\
& - \left(\frac{k_1\phi_2\left(\dfrac{p}{K_1+p}-1\right)}{K_1+p} - \frac{\lambda_p p}{K+p} + \frac{\phi_1\left(\dfrac{p}{K+p}-1\right)}{K+p}\right)\left(E_m - \frac{k_1 p}{K_1+p}\right)\frac{p}{K+p},
\end{aligned}$$

which defines the set

$$\Sigma'' \doteq \left\{(\varphi,\lambda) \in \mathbb{R}^{2n} \,|\, H_{001} = 0\right\} \cap \Sigma'.$$

As previously stated, the procedure should be repeated until u appears explicitly. For this problem, going up to the fourth derivative of H_1, one is able to retrieve the control using Poisson brackets of length five,

$$0 = H_{00001} + uH_{10001},$$

provided H_{10001} is not zero. This turns out to be the case, and one has to check the Kelley (or *generalized Legendre-Clebsch*) condition (here for $k = 2$)

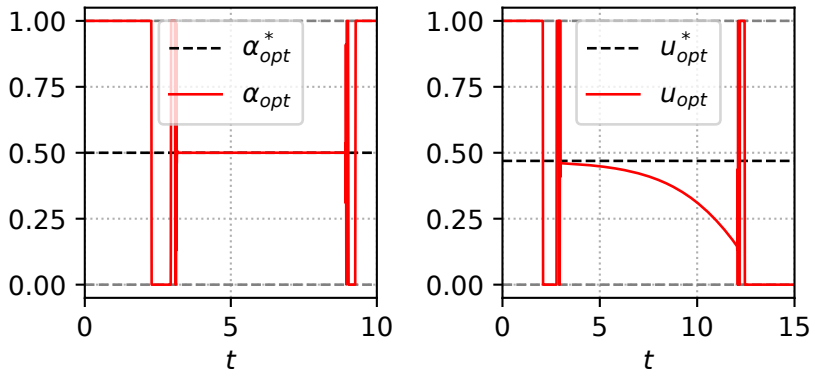
$$(-1)^k \frac{\partial}{\partial u}\left[\frac{d^{2k}}{dt^{2k}}\left(\frac{\partial H_B}{\partial u}\right)\right] < 0, \quad t \in \mathcal{I},$$

along the singular arc. This condition is necessary for optimality and we devise a numerical check in next section. Having a singular control of *intrinsic* order two (that is such that H_{101} vanishes identically on the whole cotangent space) implies that singular arcs can only be entered and exited through chattering (Fuller phenomenon); see [69]. Although the singular is only of *local* order two in our case (H_{101} vanishes when H_1 and H_{01} do, not identically), the numerical simulations below show that there is indeed chattering in and out to enter and exit the singular arc.

2.5 Numerical solution

The numerical verifications were performed with *Bocop* [52], using the following parameters: $\beta = 0.003$ L g^{-1}, $E_M = 1$, $k_R = 1.6$ h^{-1}, $K_R = 1$ h^{-1}, $k_X = 1$ h^{-1}, $K_X = 1$ h^{-1} and $\mathcal{V}_{ext} = 8$ L. The number of time steps was fixed to 5000, and the tolerance to $1e-14$. The discretization method used was the sixth-order Lobatto III C[1]. In both cases, initial conditions were set to $p(0) = 0.024$, $r(0) = 0.18$ and $\mathcal{V}(0) = 1$; and in the product maximization problem $y(0) = 0$. As expected, both optimal control solutions are characterized by a singular arc of order two that is entered and exited through chattering (Figure 2.3). Moreover, the solution for the biomass maximization problem α_{opt} matches the optimal steady-state input α^*_{opt} along the circular arc (Figure 2.3a), as predicted in the computations. However, it is not the case for the product maximization problem, where the singular solution moves away from the optimal steady-state over time (Figure 2.3b) showing that a constant input is a sub-optimal control strategy. This result suggests that, in order to maximize the production of the metabolite X, it is necessary to induce the microbial population to increasingly allocate more resources to the metabolic machinery as time passes by. Thus, the external control I should act on the natural allocation $\alpha(t)$ along the singular arc to match the lower value $u(t)$ depicted in Figure

[1]Bocop definition files are available from the authors on request.



(a) Biomass maximization (2.4.1) (b) Product maximization (2.4.2)

Figure 2.3: Solutions of the OCPs computed using *Bocop*.

2.3b. In both cases, the optimal steady-state inputs α^*_{opt} and u^*_{opt} were computed numerically by selecting the steady-state input that maximizes the integrand of each criterion, which corresponds to the solution of the static problem, already shown to exist and to be unique in previous section. The fact that, for the product maximization problem, the singular arc is of order two, can be verified by evaluating the derivatives of H_1 over the optimal trajectories (Figure 2.4). In this figure, all subsequent derivatives vanish along the sub-interval $\mathcal{I}$, except for the fourth one. We eventually provide a numerical check of the generalized Legendre-Clebsch condition.

2.6 Conclusions

Self-replicator models of bacterial growth are capable of reproducing the growth laws of certain organisms at steady state. However, in [30], it has been shown that these models can also account for dynamical environments. In this context, natural biomass maximization was achieved through a bang-singular strategy. In [43], authors showed that the same kind of strategy is necessary in order to maximize the synthetic production of a metabolite. In line with this work, we have proposed a metabolite production scheme in a CSTR Bioreactor, with the particularity that it encompasses previously studied models in the field. We have focused on the most relevant particular cases to emphasize the importance of singular regimes and chattering arcs on optimal control solutions. We have provided an analytical derivation of the results, as well as a numerical characterization of the singular regimes. Con-

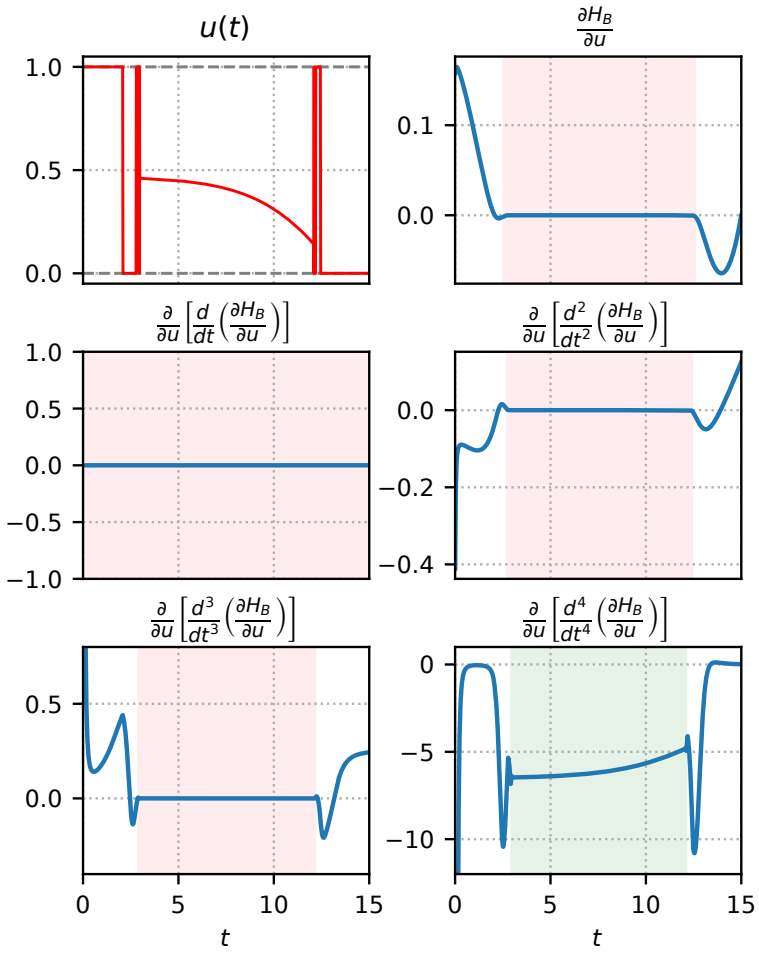


Figure 2.4: Simulation of the OCP 2.4.2 (Product maximization) with *Bocop*. The intervals where the functions vanish are highlighted in light red. All functions vanish along the singular arc but $\frac{\partial}{\partial u}\left[\frac{d^4}{dt^4}\left(\frac{\partial}{\partial u}H_B\right)\right]$, highlighted in green, which is negative as expected.

trary to the natural biomass maximization process, maximizing the production of a metabolite is accomplished by a singular solution that does not contain any steady state, showing that a time-varying action is indeed required to achieve maximization. For this particular problem, an optimal external control should increasingly allocate more resources to the metabolic machinery, while reducing the allocation to the gene expression machinery. Results raise interesting questions from the optimal control point of view as it can provide ideal control scenarios, but also about how to implement these open-loop strategies in real environments. Indeed, in order to do so, further analysis is required for the newly presented CSTR Bioreactor model,

which is our current objective.

Chapter 3

Substrate depletion

This chapter describes unpublished work in progress.

3.1 Introduction

The study of living microorganisms through resource allocation models has increasingly become relevant for its capacity to elucidate natural behaviors of microbia through very simple dynamical models. The core idea is to represent the distribution of internal resources through optimal control strategies, based on the assumption that evolutionary processes have tuned these endogenous allocation strategies to attain nearly-optimal levels. Numerous problems arise in this context, one of them being the optimal production of metabolites regulated by an external control capable of arresting the bacterial growth. To this end, a resource allocation approach can help understand how to modify the naturally-evolved allocation strategies so as to efficiently produce such metabolite.

This is the subject of this paper, which tackles the problem of batch processing from a resource allocation perspective. A simple coarse-grained self-replicator model is introduced, based on [3] and [AYT5], with minimal biological assumptions. Using mass conservation laws, it is possible to analyse the asymptotic behavior and stability of the dynamical system, showing that for every possible allocation strategy, all component of the system are transformed either into proteins or into metabolites, a condition later defined as *Full depletion*. Then, two main studies are performed: the biomass maximization case, representing the natural objective of wild-type microbial cultures; and the metabolite maximization case, using a bacterial model that

includes a pathway for metabolite synthesis for industrial purposes. Both problems are analyzed in infinite time and in finite time, the latter stated as an Optimal Control Problem (OCP) which is solved through the application of the Pontryagin's Maximum Principle (PMP). The problem presented here is similar to the one studied in [3], but approached from an analytical perspective, and including the case with no metabolite synthesis as a starting point.

3.2 Model definition

3.2.1 Self-replicator model

We define a self-replicator model describing the dynamics of a microbial population growing inside a closed bioreactor of fixed volume $\mathcal{V}_{ext} > 0$ liters. At the beginning of the experience, there is an initial mass of substrate S inside the bioreactor, that is gradually consumed by the bacterial population, transforming it into precursor metabolites P. Precursors are transformed into components of the gene expression machinery R, into enzymes that makes up the metabolic machinery M, and into a metabolite of interest X that is excreted from the cell. While the production of proteins M and R is catalyzed by the enzymes R, the absorption of S and synthesis of X are both catalyzed by M (dashed arrows in Figure 3.1). The proportion of

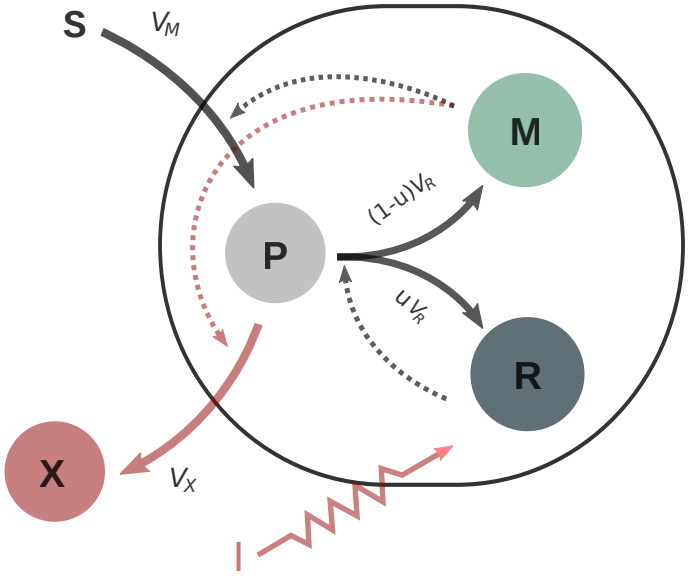


Figure 3.1: Self-replicator model

precursors dedicated to the production of M and R is decided by the allocation control $u(t) \in [0,1]$: $u = 0$ means no production of R, while $u = 1$ means no production of M. The control u can represent the natural allocation used by bacteria, as modelled in [30], as well as the artificially modified allocation modelled in [3], depending on the objective to be analyzed. The dynamics of the system are described by

$$\begin{cases} \dot{S} = -V_M, \\ \dot{P} = V_M - V_X - V_R, \\ \dot{M} = (1-u)\, V_R, \\ \dot{R} = uV_R, \\ \dot{X} = V_X, \end{cases}$$

where the variables $S(t)$, $P(t)$, $M(t)$, $R(t)$ and $X(t)$ represent the masses (in grams) of substrate, precursors metabolites, the metabolic machinery, the gene expression machinery and the metabolites of interest at time t, respectively; V_M $[g/h]$, V_R $[g/h]$ and V_X $[g/h]$ are the reaction rates of the system, and $u(t)$ is the allocation control previously defined. We define the structural volume of the bacterial population inside the bioreactor $\mathcal{V}_L(t)$ measured in liters as

$$\mathcal{V}_L \doteq \beta(M+R), \tag{3.1}$$

where β is a constant relating density and volume. This allows us to define concentrations with respect to this volume

$$p_{\mathcal{V}} \doteq \frac{P}{\mathcal{V}_L}\left[\frac{g}{L}\right], \qquad m_{\mathcal{V}} \doteq \frac{M}{\mathcal{V}_L}\left[\frac{g}{L}\right], \qquad r_{\mathcal{V}} \doteq \frac{R}{\mathcal{V}_L}\left[\frac{g}{L}\right], \tag{3.2}$$

and define relative rates involved in the processes as

$$v_M(s_{\mathcal{V}}, m_{\mathcal{V}}) \doteq \frac{V_M}{\mathcal{V}_L}\left[\frac{g}{Lh}\right], \quad v_R(p_{\mathcal{V}}, r_{\mathcal{V}}) \doteq \frac{V_R}{\mathcal{V}_L}\left[\frac{g}{Lh}\right], \quad v_X(p_{\mathcal{V}}, m_{\mathcal{V}}) \doteq \frac{V_X}{\mathcal{V}_L}\left[\frac{g}{Lh}\right].$$

Likewise, we define the concentrations related to the external volume

$$s_{\mathcal{V}} = \frac{S}{\mathcal{V}_{ext}}\left[\frac{g}{L}\right], \quad x_{\mathcal{V}} = \frac{X}{\mathcal{V}_{ext}}\left[\frac{g}{L}\right].$$

From (3.1) and (3.2) we have that

$$m_{\mathcal{V}} + r_{\mathcal{V}} = \frac{M+R}{\mathcal{V}_L} = \frac{1}{\beta}. \tag{3.3}$$

We define the growth rate $\mu = \mu(t)$ as

$$\mu \doteq \frac{\dot{\mathcal{V}}_L}{\mathcal{V}_L} = \beta v_R(p_{\mathcal{V}}, r_{\mathcal{V}}).$$

and that the dynamical system can be expressed in terms of the concentrations as

$$\begin{cases} \dot{s}_{\mathcal{V}} = -v_M(s_{\mathcal{V}}, m_{\mathcal{V}})\dfrac{\mathcal{V}_L}{\mathcal{V}_{ext}}, \\ \dot{p}_{\mathcal{V}} = v_M(s_{\mathcal{V}}, m_{\mathcal{V}}) - v_X(p_{\mathcal{V}}, m_{\mathcal{V}}) - v_R(p_{\mathcal{V}}, r_{\mathcal{V}})(\beta p_{\mathcal{V}} + 1), \\ \dot{r}_{\mathcal{V}} = (u - \beta r_{\mathcal{V}})v_R(p_{\mathcal{V}}, r_{\mathcal{V}}), \\ \dot{m}_{\mathcal{V}} = (1 - u - \beta m_{\mathcal{V}})v_R(p_{\mathcal{V}}, r_{\mathcal{V}}), \\ \dot{\mathcal{V}}_L = \beta v_R(p_{\mathcal{V}}, r_{\mathcal{V}})\mathcal{V}_L, \\ \dot{x}_{\mathcal{V}} = v_X(p_{\mathcal{V}}, m_{\mathcal{V}})\dfrac{\mathcal{V}_L}{\mathcal{V}_{ext}}. \end{cases}$$

3.2.2 Kinetics definition

We model the kinetics of the system by supposing they are linear in $m_{\mathcal{V}}$ and $r_{\mathcal{V}}$ [2]. Thus, they can be expressed as

$$\begin{aligned} v_M(s_{\mathcal{V}}, m_{\mathcal{V}}) &= w_M(s_{\mathcal{V}})m_{\mathcal{V}}, \\ v_R(p_{\mathcal{V}}, r_{\mathcal{V}}) &= w_R(p_{\mathcal{V}})r_{\mathcal{V}}, \\ v_X(p_{\mathcal{V}}, m_{\mathcal{V}}) &= \gamma w_R(p_{\mathcal{V}})m_{\mathcal{V}}, \end{aligned}$$

where $\gamma > 0$ is a proportionality constant, under the assumption that $v_X(p_{\mathcal{V}}, m_{\mathcal{V}})/v_R(p_{\mathcal{V}}, r_{\mathcal{V}})$ does not depend on $p_{\mathcal{V}}$ but on the ratio $m_{\mathcal{V}}/r_{\mathcal{V}}$. The functions w_i are assumed to have the following behaviour:

Assumption 3.2.1. *Function* $w_i(x) : \mathbb{R}_+ \to \mathbb{R}_+$ *is*

- *Continuously differentiable w.r.t. x,*
- *Null at the origin:* $w_i(0) = 0$,
- *Strictly monotonically increasing:* $w_i'(x) > 0, \forall x \geq 0$,
- *Strictly concave downwards:* $w_i''(x) < 0, \forall x \geq 0$,
- *Upper bounded:* $\lim_{x\to\infty} w_i(x) = k_i > 0$.

For numerical simulations, we will resort to the particular case where the functions follow Michalis-Menten kinetics. For that case, we define

$$w_R(p_\mathcal{V}) \doteq k_R \frac{p_\mathcal{V}}{K_R + p_\mathcal{V}}, \quad w_M(s_\mathcal{V}) \doteq k_M \frac{s_\mathcal{V}}{K_M + s_\mathcal{V}},$$

where the values of the constants k_R, K_R, k_M and K_M are based on the literature [30]. For the general case introduced in Hypothesis 3.2.1 we will define

$$k_R \doteq \lim_{p_\mathcal{V}\to\infty} w_R(p_\mathcal{V}), \quad k_M \doteq \lim_{s_\mathcal{V}\to\infty} w_M(s_\mathcal{V}).$$

3.2.3 Mass fraction formulation and non-dimensionalization

We define mass fractions of the total bacterial mass as

$$p \doteq \beta p_\mathcal{V}, \quad r \doteq \beta r_\mathcal{V}, \quad m \doteq \beta m_\mathcal{V}, \quad s \doteq \beta s_\mathcal{V}, \quad x \doteq \beta x_\mathcal{V},$$

and the biomass fraction of the bioreactor

$$\mathcal{V} \doteq \frac{\mathcal{V}_L}{\mathcal{V}_{ext}},$$

which, replacing in (3.3), yields

$$m + r = 1. \tag{3.4}$$

We define the non-dimensional time $\hat{t} \doteq k_R t$ and the non-dimensional functions

$$\hat{w}_R(p) = \frac{w_R(p_\mathcal{V})}{k_R}, \quad \hat{w}_M(s) = \frac{w_M(s_\mathcal{V})}{k_R}$$

so that $lim_{p\to\infty}\hat{w}_R(p) = 1$. Using (3.4), m can be expressed in terms of r, and so the dynamical equation of m can be removed from the system. For the sake of simplicity, let us drop all hats from the current notation. Thus, system becomes

$$\begin{cases} \dot{s} = -w_M(s)(1-r)\mathcal{V}, \\ \dot{p} = w_M(s)(1-r) - \gamma w_R(p)(1-r) - w_R(p)r(p+1), \\ \dot{r} = (u-r)w_R(p)r, \\ \dot{\mathcal{V}} = w_R(p)r\mathcal{V}, \\ \dot{x} = \gamma w_R(p)(1-r)\mathcal{V}. \end{cases} \tag{S}$$

3.3 Model analysis

Lemma 3.3.1. *The set,*

$$\Gamma = \left\{(s,p,r,\mathcal{V},x) \in \mathbb{R}^5 \,:\, s \geq 0,\, p \geq 0,\, 1 \geq r \geq 0,\, \mathcal{V} \geq 0,\, x \geq 0\right\}$$

is positively invariant for the initial value problem.

Thus, we fix initial conditions

$$s(0) = s_0 > 0, \quad p(0) = p_0 > 0, \quad r(0) = r_0 \in (0,1), \quad \mathcal{V}(0) = \mathcal{V}_0 > 0, \quad x(0) = 0. \tag{IC}$$

Some relations are immediate from the dynamics: as $\dot{s} \leq 0$ and $\dot{\mathcal{V}} \geq 0$ for all t, we have

$$s(t) \leq s_0, \quad \mathcal{V}(t) \geq \mathcal{V}_0, \tag{3.5}$$

representing the fact that the substrate can only be consumed (and not replenished), and the biomass can only grow.

3.3.1 Total available mass

As typically occurs in batch processes, there is neither inflow nor outflow of mass in the bioreactor, which is reflected in the dynamics of the system though a mass

conservation law. We define the constant

$$\Sigma \doteq s_0 + (p_0 + 1)\mathcal{V}_0,$$

representing the initial mass concentration in the system. It can be seen that the total mass concentration

$$z \doteq s + (p + 1)\mathcal{V} + x$$

is constant for all t (as $\dot{z} = 0$). This means that

$$s(t) + (p(t) + 1)\mathcal{V}(t) + x(t) = \Sigma, \qquad \text{(First integral 1)}$$

for all t. Variables $\mathcal{V}$ and x are maximal when the remaining variables are equal to 0, and so they are upper bounded. In particular, both $\mathcal{V}(t)$ and $x(t)$ are decreasing w.r.t. $s(t)$ and $p(t)$. As neither s nor p can be negative, we have that

$$\mathcal{V}(t) + x(t) = \Sigma \tag{3.6}$$

when $s(t) = p(t) = 0$. This condition means that all the available substrate and precursor metabolites have been depleted and transformed into biomass and metabolites, which is intuitively what one would expect from system (S) for t sufficiently large. Additionally, using (3.5) and (First integral 1), we can obtain the following result.

Proposition 3.3.2. *$\mathcal{V}(t) \in [\mathcal{V}_0, \Sigma]$, $x(t) \in [0, \Sigma - \mathcal{V}_0]$ and $p(t) \in [0, p_{max}]$ for all t, with $p_{max} = \Sigma/\mathcal{V}_0 - 1$.*

3.3.2 Infinite-time full depletion

Dynamics (S) shows that, under initial conditions (IC), $s(t)$ and $p(t)$ can only vanish asymptotically, that is, when $t \to \infty$. The latter can be proved by seeing that the dynamical equations of s and p can be bounded to

$$\dot{s} \geq -w_M(s)\Sigma, \quad \dot{p} \geq -w_R(p)p_{max}(p_{max} + 1 + \gamma),$$

which means that, at worst, s and p decay exponentially (as functions $w_i(x)$ can be upper bounded by linear functions $w_i(x) \leq c_i x$), and thus $s(t) = p(t) = 0$ cannot be

attained in finite time. Thus, we define the depletion of s and p in an infinite time horizon.

Definition 3.3.3. *System* (S) *achieves* Full depletion *when all the substrate and the precursors are asymptotically depleted, i.e.*

$$\lim_{t\to\infty} s(t) = \lim_{t\to\infty} p(t) = 0, \qquad (\textit{Full depletion})$$

3.3.3 Asymptotic behaviour

Now, we will study the system dynamics for an infinite time $t \to \infty$.

Constant allocation u^*

Theorem 3.3.1. *For any trajectory of system* (S) *with initial conditions* (IC) *and constant allocation $u(t) = u^*$, it follows that*

$$r(t) = u^* - (u^* - r_0)\frac{\mathcal{V}_0}{\mathcal{V}(t)}. \qquad \text{(First integral 2)}$$

Proof. Under a constant allocation $u(t) = u^*$, the dynamics of r becomes

$$\dot{r} = (u^* - r)\mu(p)r.$$

Using dynamics (S), it is possible to see that the quantity $R_u = (u^* - r)\mathcal{V}$ is constant (as $\dot{R}_u = 0$), meaning that $(u^* - r(t))\mathcal{V}(t) = (u^* - r_0)\mathcal{V}_0$ for all t, which yields (First integral 2). □

General allocation $u(t)$

Due to the boundedness of $\mathcal{V}$ stated in Lemma 3.3.2, and the relation between $\mathcal{V}$ and r shown in (First integral 1), we can see that any constant control u^* yields a bounded ribosomal concentration r. We will extend this notion to any function $u(t)$.

Lemma 3.3.4. *For any trajectory of system* (S) *with initial conditions* (IC) *and any control $u(t)$, the ribosomal concentration has bounds $r(t) \in [r_{\min}, r_{\max}]$ for all t, with*

$$r_{\min} \doteq r_0\frac{\mathcal{V}_0}{\Sigma} > 0, \qquad r_{\max} \doteq 1 - (1 - r_0)\frac{\mathcal{V}_0}{\Sigma} < 1.$$

Proof. Let us extend system (S) by defining variables $r_{\text{low}}(t)$ and $r_{\text{up}}(t)$ with dynamics

$$\begin{aligned} &\dot{r}_{\text{low}} = -r_{\text{low}} w_R(p) r \leq 0, \quad \dot{r}_{\text{up}} = (1 - r_{\text{up}}) w_R(p) r \geq 0, \\ &r_{\text{low}}(0) = r_0, \qquad\qquad\qquad r_{\text{up}}(0) = r_0, \end{aligned}$$

which correspond to the dynamics of r with $u = 0$ and $u = 1$ respectively, and which satisfy

$$r_{\text{low}}(t) \leq r(t) \leq r_{\text{up}}(t)$$

for all t. The latter can be easily proved by showing that the time-varying differences

$$\Delta_{\text{low}}(t) = r(t) - r_{\text{low}}(t), \quad \Delta_{\text{up}}(t) = r_{\text{up}}(t) - r(t)$$

with dynamics

$$\dot{\Delta}_{\text{low}} = (u - \Delta_{\text{low}}) w_R(p) r, \quad \dot{\Delta}_{\text{up}} = (1 - u - \Delta_{\text{up}}) w_R(p) r$$

are always non-negative: they satisfy $\Delta_{\text{low}}(0) = \Delta_{\text{up}}(0) = 0$ and are repulsive or (at worst) invariant at 0. Then, based on the same principle used to obtain (First integral 2), we define the quantities $R_{\text{low}} = r_{\text{low}} \mathcal{V}$ and $R_{\text{up}} = (1 - r_{\text{up}}) \mathcal{V}$ which are constant (as $\dot{R}_{\text{low}} = \dot{R}_{\text{up}} = 0$), and so

$$r_{\text{low}}(t) = r_0 \frac{\mathcal{V}_0}{\mathcal{V}(t)}, \qquad r_{\text{up}}(t) = 1 - (1 - r_0) \frac{\mathcal{V}_0}{\mathcal{V}(t)},$$

for all t. As $\mathcal{V}_0 \leq \mathcal{V}(t) \leq \Sigma$ for all t, we have

$$r_{\text{low}}(t) \in \left[r_0 \frac{\mathcal{V}_0}{\Sigma}, r_0 \right], \qquad r_{\text{up}}(t) \in \left[r_0, 1 - (1 - r_0) \frac{\mathcal{V}_0}{\Sigma} \right]$$

which shows that $r_{\min} \leq r(t) \leq r_{\max}$ for all t. □

Lemma 3.3.4 states that, for any control $u(t)$, the ribosomal concentration never reaches the bounds $r = 0$ and $r = 1$, and thus neither the substrate intake nor the protein synthesis is arrested. We will see that this feature produces (*Full depletion*) in the following theorem.

Theorem 3.3.2. *Any trajectory of system* (S) *with initial conditions* (IC) *and any control* $u(t)$ *achieves* (*Full depletion*).

Proof. Using Lemma 3.3.4, it is easy to see that

$$\dot{s} \leq -w_M(s)(1 - r_{\max})\mathcal{V}_0,$$

which means that $s(t)$ converges to 0 as $t \to \infty$. Then, this means that

$$\dot{p} \leq -\gamma w_R(p)(1 - r_{\max}) - w_R(p)r_{\min},$$

and so $p(t)$ also converges to 0 as $t \to \infty$. □

3.4 The biomass maximization case

For this section, we will write the problem of maximizing the biomass for both infinite time and finite time. The latter is a mathematical representation of the naturally-evolved resource allocation strategy used by bacteria in nature. For that, we will assume that no metabolite is produced, as the pathway responsible for its production is artificially engineered, and thus not present in wild-type bacteria. This is simply modeled through $\gamma = 0$. The resulting system is

$$\begin{cases} \dot{s} = -w_M(s)(1 - r)\mathcal{V}, \\ \dot{p} = w_M(s)(1 - r) - w_R(p)r(p + 1), \\ \dot{r} = (u - r)w_R(p)r, \\ \dot{\mathcal{V}} = w_R(p)r\mathcal{V}, \end{cases} \tag{$S_\mathcal{V}$}$$

3.4.1 Infinite-time problem

Problem formulation

We first write the biomass maximization problem for an infinite time horizon. The problem can be expressed as

$$\max_{u(t)} \lim_{t \to \infty} \mathcal{V}(t).$$

Since $\mathcal{V} \in [\mathcal{V}_0, \Sigma]$, applying (*Full depletion*) in (First integral 1) yields the condition

$$\lim_{t\to\infty} \mathcal{V}(t) = \Sigma.$$

meaning that, in infinite time, the biomass is maximized for every control $u(t)$. As a consequence, using Lemma 3.3.1, we have the following result for constant allocations.

Corollary 3.4.1. *For any trajectory of system* ($\mathrm{S}_\mathcal{V}$) *with initial conditions* (IC) *and constant control* $u(t) = u^*$,

$$\lim_{t\to\infty} r(t) = u^* - (u^* - r_0)\frac{\mathcal{V}_0}{\Sigma}$$

Numerical simulations

Examples of trajectories confirming the analytical results are shown in Figure 3.2 and Figure 3.3, where we see that the system approaches (*Full depletion*) asymptotically in every case, thus approaching the maximal biomass value $\mathcal{V}(t) = \Sigma$. Figure 3.2 shows the resulting trajectories associated to the same initial conditions, when varying the allocation parameter u. On the other hand, Figure 3.3 illustrates the trajectories for different values of r_0.

3.4.2 Finite-time problem

Problem formulation

Let us fix a final time $t_f > 0$, and write the OCP maximizing the final bacterial volume $\mathcal{V}(t_f)$ with initial conditions (IC). The latter writes

$$\begin{cases} \textit{maximize} & \mathcal{V}(t_f), \\ \textit{subject to} & \textit{dynamics of } (\mathrm{S}_\mathcal{V}), \\ & \textit{initial conditions } (\mathrm{IC}), \\ & u(\cdot) \in \mathcal{U}. \end{cases} \qquad (\mathrm{OCP}_\mathcal{V})$$

For this class of optimal control problem, where there are no terminal constraints, there is no controllability issues. Additionally, the dynamics is affine in the control,

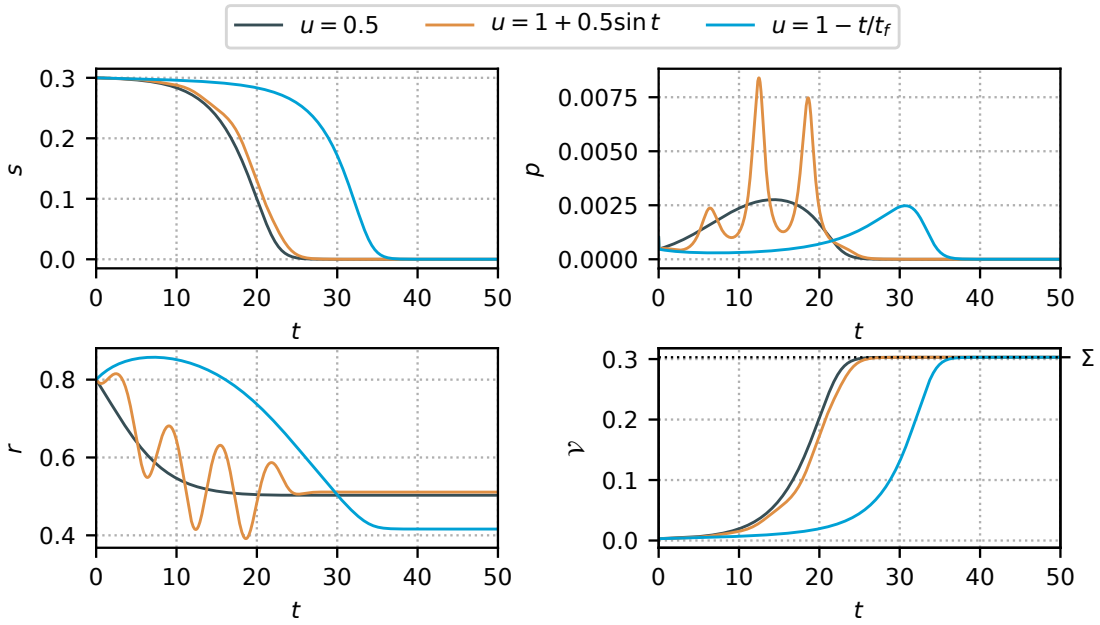


Figure 3.2: Simulation of $(\mathrm{S}_\mathcal{V})$ with $s_0 = 0.3$, $p_0 = 0.001$, $r_0 = 0.8$, $\mathcal{V}_0 = 0.003$ and different allocation functions u.

with the latter included in a compact and convex set (a closed interval), and it can be checked that every finite-time trajectory remains bounded. Thus, existence of a solution is guaranteed by Filippov's theorem [53]. Then, for a problem $(\mathrm{OCP}_\mathcal{V})$ with state $\varphi \in \mathbb{R}^n$, PMP ensures that there exist $\lambda^0 \leq 0$ and a piecewise absolutely continuous mapping $\lambda(\cdot) : [0, t_f] \to \mathbb{R}^n$, with $(\lambda(\cdot), \lambda^0) \neq (0, 0)$, such that the extremal $(\varphi, \lambda, \lambda^0, u)$ satisfies the generalized Hamiltonian system

$$\begin{cases} \dot{\varphi} = \dfrac{\partial}{\partial \lambda} H(\varphi, \lambda, \lambda^0, u), \\ \dot{\lambda} = -\dfrac{\partial}{\partial \varphi} H(\varphi, \lambda, \lambda^0, u), \\ H(\varphi, \lambda, \lambda^0, u) = \max_{u \in [0,1]} H(\varphi, \lambda, \lambda^0, u), \end{cases} \tag{PMP}$$

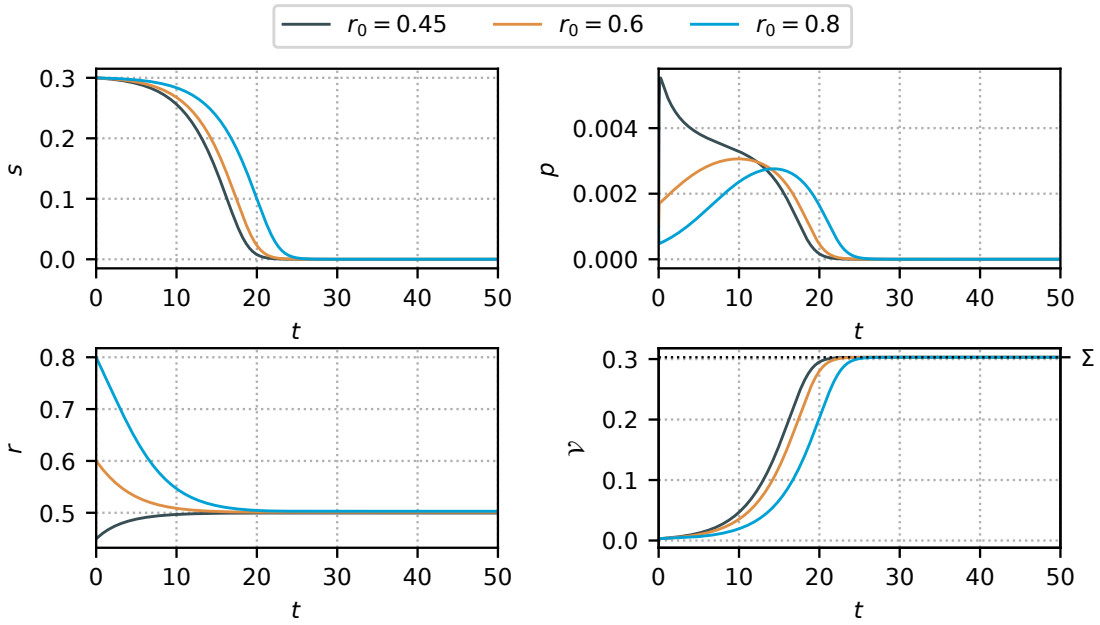


Figure 3.3: Simulation of ($S_{\mathcal{V}}$) with $s_0 = 0.3$, $p_0 = 0.001$, $\mathcal{V}_0 = 0.003$, $u = 0.5$ and different values of r_0.

for almost every $t \in [0, t_f]$. We define the adjoint states for this particular case as $\lambda = (\lambda_s, \lambda_p, \lambda_r, \lambda_{\mathcal{V}})$, and we write the Hamiltonian

$$\begin{aligned} H = &- w_M(s)(1-r)\mathcal{V}\lambda_s + \Big(w_M(s)(1-r) - w_R(p)r(p+1)\Big)\lambda_p + w_R(p)r\mathcal{V}\lambda_{\mathcal{V}} \\ &+ (u-r)w_R(p)r\lambda_r, \end{aligned}$$

and the adjoint system as

$$\begin{cases} \dfrac{d\lambda_s}{dt} = w_M'(s)(1-r)\left(\mathcal{V}\lambda_s - \lambda_p\right), \\ \dfrac{d\lambda_p}{dt} = \Big(w_R'(p)r(p+1) + w_R(p)r\Big)\lambda_p - w_R'(p)r\mathcal{V}\lambda_{\mathcal{V}} - (u-r)w_R'(p)r\lambda_r, \\ \dfrac{d\lambda_r}{dt} = -w_M(s)(\mathcal{V}\lambda_s - \lambda_p) + w_R(p)(p+1)\lambda_p - w_R(p)\mathcal{V}\lambda_{\mathcal{V}} - (u-2r)w_R(p)\lambda_r, \\ \dfrac{d\lambda_{\mathcal{V}}}{dt} = w_M(s)(1-r)\lambda_s - w_R(p)r\lambda_{\mathcal{V}}. \end{cases}$$

Given that there are no terminal conditions on the state, the transversality conditions for the adjoint state are $\lambda(t_f) = (0, 0, 0, \lambda_0)$, where the presence of λ_0 is due to the cost function being $\mathcal{V}(t_f)$. Since the Hamiltonian is linear in the control u, we rewrite it in the input-affine form $H = H_0 + uH_1$ with

$$\begin{aligned} H_0 =& - w_M(s)(1-r)\mathcal{V}\lambda_s + \Big(w_M(s)(1-r) - w_R(p)r(p+1)\Big)\lambda_p \\ & + w_R(p)r\mathcal{V}\lambda_{\mathcal{V}} - w_R(p)r^2\lambda_r, \\ H_1 =& w_R(p)r\lambda_r. \end{aligned}$$

The constrained optimal control u should maximize the Hamiltonian, so the solution of $(\mathrm{OCP}_{\mathcal{V}})$ is

$$u(t) = \begin{cases} 0 & \textit{if } H_1 < 0, \\ 1 & \textit{if } H_1 > 0, \\ u_{\mathrm{sing}}(t) & \textit{if } H_1 = 0, \end{cases} \tag{3.7}$$

where $u_{\mathrm{sing}}(t)$ is called a singular control, showing that any optimal control is a concatenation of bangs ($u = 0$ and $u = 1$) and singular arcs, depending on the sign of the switching function H_1.

Singular arcs

The singular arc is produced when H_1 vanished (as well as its successive derivatives w.r.t. time) on a whole sub-interval $[t_1, t_2] \subset [0, t_f]$. As r is bounded and p cannot vanish in finite time, along the singular arc one has

$$\phi_0 = 0, \quad \textit{with } \phi_0 = \lambda_r, \tag{C1}$$

and thus, along a singular arc, the Hamiltonian becomes

$$H = -w_M(s)(1-r)\mathcal{V}\lambda_s + \Big(w_M(s)(1-r) - w_R(p)r(p+1)\Big)\lambda_p + w_R(p)r\mathcal{V}\lambda_{\mathcal{V}}.$$

We differentiate $(C1)$ and we get

$$\phi_1 = 0, \quad \textit{with } \phi_1 = -w_M(s)(\mathcal{V}\lambda_s - \lambda_p) + w_R(p)(p+1)\lambda_p - w_R(p)\mathcal{V}\lambda_{\mathcal{V}}. \tag{C2}$$

Then, the Hamiltonian can be expressed as

$$H = -w_M(s)(\mathcal{V}\lambda_s - \lambda_p) - r\phi_1 + (u - r)H_1 = c$$

and so ϕ_1 becomes

$$\phi_1 = c + w_R(p)\Big((p+1)\lambda_p - \mathcal{V}\lambda_{\mathcal{V}}\Big).$$

We differentiate $(C2)$ and we get

$$\phi_2 = 0, \quad \textit{with } \phi_2 = \frac{w_R^2(p)}{w_M(s)w_R'(p)} - 1. \tag{C3}$$

Given that u does not appear in $(C3)$, we have the following result.

Lemma 3.4.2. *Any singular arc is of order two and, along it, s can be expressed in terms of p (provided that $w_M(s)$ is invertible) through the equation $w_M(s) = x(p)$, where $x(p) : \mathbb{R}^+ \to \mathbb{R}^+$ is defined as $x(p) \doteq w_R^2(p)/w_R'(p)$.*

We differentiate $(C3)$ and we get

$$\phi_3 = 0, \quad \textit{with } \phi_3 = w_M'(s)w_M(s)(1-r)\mathcal{V} + x'(p)\Big(w_M(s)(1-r) - w_R(p)r(p+1)\Big), \tag{C4}$$

which shows another relation between states. The latter implies:

Lemma 3.4.3. *Along the singular arc, r can be expressed in terms of p and $\mathcal{V}$ as $r = \phi(p, \mathcal{V})$ with $\phi(p, \mathcal{V}) : \mathbb{R}^+ \times \mathbb{R}^+ \to (0, 1)$ defined as*

$$\phi(p, \mathcal{V}) \doteq \frac{x(p)\Big(w_M'(s)\mathcal{V} + x'(p)\Big)}{x(p)\Big(w_M'(s)\mathcal{V} + x'(p)\Big) + x'(p)w_R(p)(p+1)}$$

Now we have:

Theorem 3.4.1. *The singular optimal control $u(t) = u(p, \mathcal{V})$ (i.e. is in feedback form).*

Proof. Derivating the expression $r = \phi(p, \mathcal{V})$ and solving for u yields

$$u = \frac{1}{w_R(p)r}\frac{d\phi(p, \mathcal{V})}{dt} + r.$$

As ϕ only depends on p and $\mathcal{V}$, its derivative will only depend on the state. Then, since s and r can be expressed in terms of p and $\mathcal{V}$, the control u becomes a function of these variables. □

Numerical simulations

The optimal trajectories were computed with Bocop [4], which solves the OCP through a direct method. The time discretization algorithm used is Lobato IIIC (implicit, 4-stage, order 6) with 2000 time steps. Figures 3.4 and 3.5 show optimal trajectories for the same set of initial conditions and different values of t_f. Using the mass conservation law (First integral 1), the quantities are represented in the plots as fractions of the total mass in the bioreactor Σ. The optimal control is characterized by the presence of the chattering phenomenon after and before the singular arc, as expected in singular arcs of order two. From a biological point of view, both allocation strategies prioritize the synthesis of proteins of the metabolic machinery M (red in both Figures): the singular arc takes rather small values, and, depending on the choice of t_f, a large proportion of the optimal control at the end of the trajectory is composed of a bang arc $u = 0$. The latter strategy promotes nutrient uptake, which results in a faster depletion of the substrate. By assumption, the concentration of precursor metabolites in the bioreactor remains negligible in comparison with the other quantities. Figure 3.6 illustrates the trajectory represented in Figure 3.4 in the $sp-$plane, showing that the trajectory approaches the curve $\phi_2(p, s) = 0$ obtained from the singular surface, slides along it, and then follows a trajectory obtained from the $u = 0$ arc that goes towards (*Full depletion*).

3.5 The product maximization case

As done in the previous section, we approach the product maximization objective in infinite time and finite time using the full model (S) where $\gamma \in \mathbb{R}^+$.

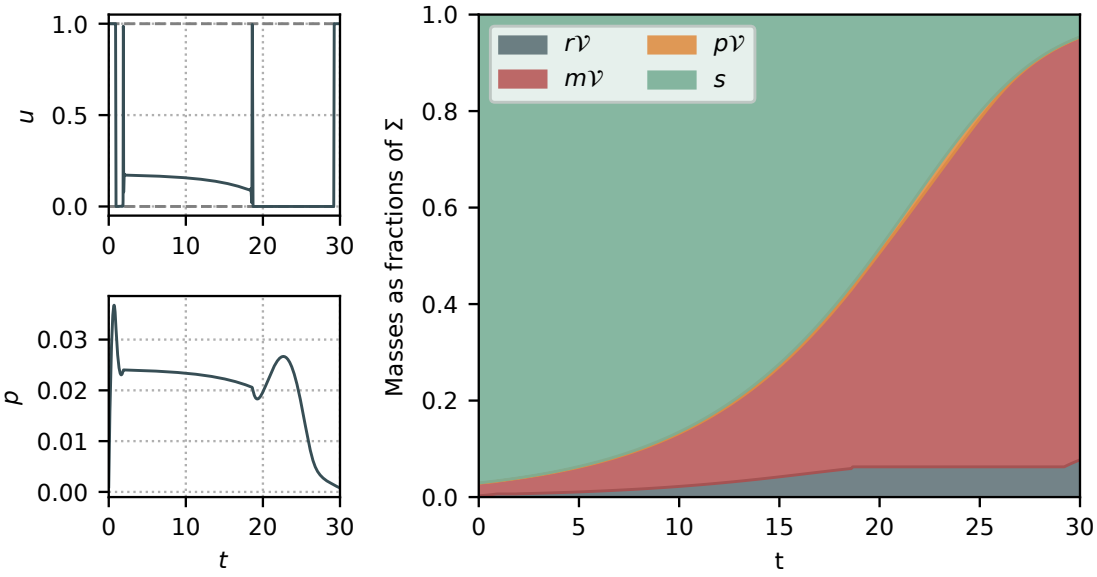


Figure 3.4: Simulation with $s_0 = 0.1$, $p_0 = 0.003$, $r_0 = 0.1$, $\mathcal{V}_0 = 0.003$ and $t_f = 30$. The final volume $\mathcal{V}(t_f)$ is at 95% of Σ (knowing that $m\mathcal{V} + r\mathcal{V} = \mathcal{V}$).

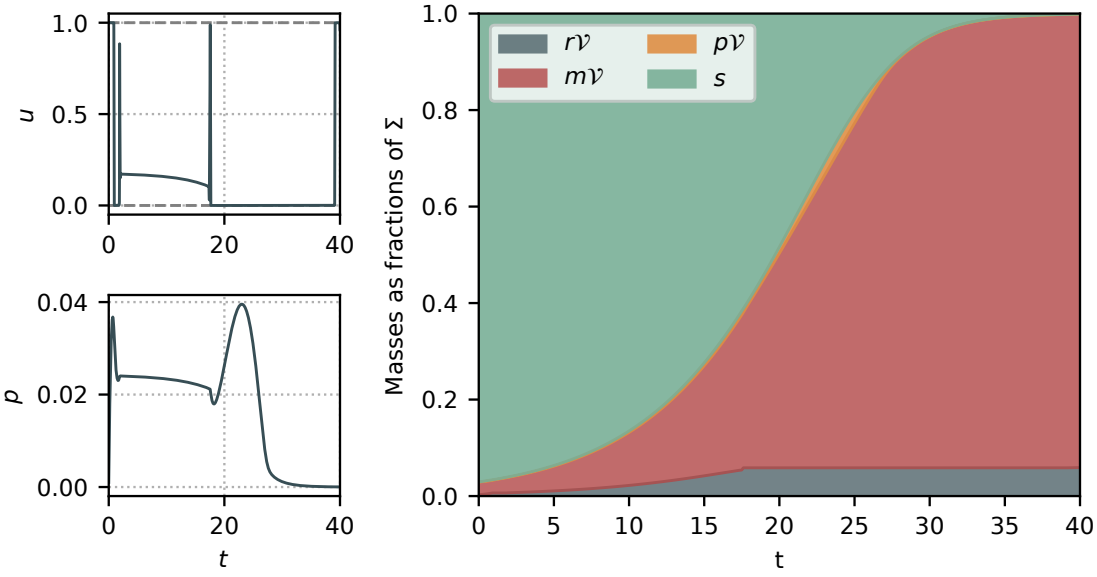


Figure 3.5: Simulation with $s_0 = 0.1$, $p_0 = 0.003$, $r_0 = 0.1$, $\mathcal{V}_0 = 0.003$ and $t_f = 40$. The final volume $\mathcal{V}(t_f)$ is at 99.8% of Σ.

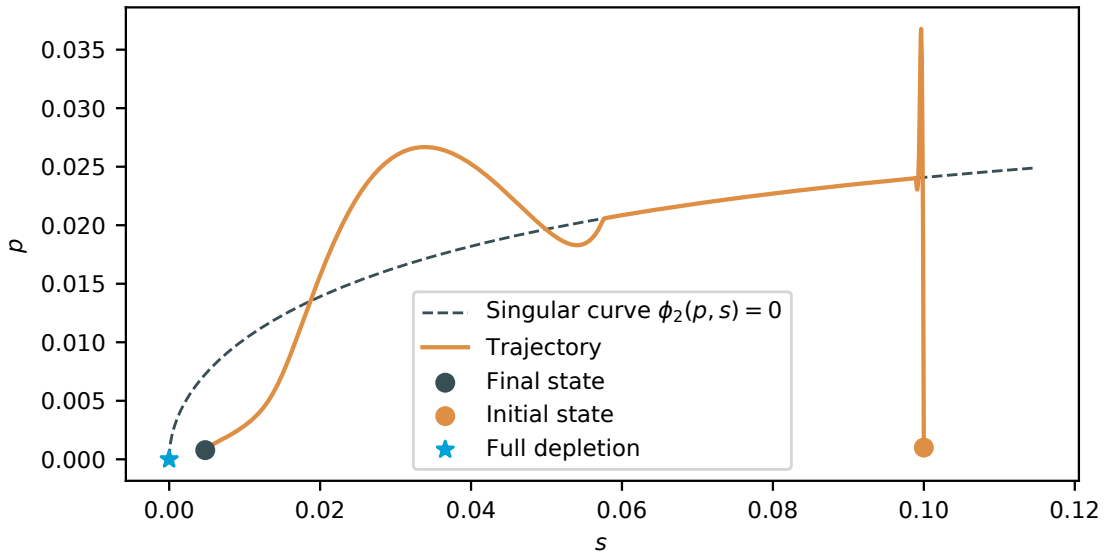


Figure 3.6: Same trajectory as Figure 3.4, in the *sp*-plane. The state approaches the curve $\phi_2 = 0$ and slides along it.

3.5.1 Infinite-time problem

The problem of maximizing the product concentration at infinite time is given by the expression

$$\max_{u^*} \lim_{t\to\infty} x(t),$$

which, using (3.6), can be rewritten as

$$\min_{u^*} \lim_{t\to\infty} \mathcal{V}(t)$$

indicating that maximizing the metabolite concentration at infinite time equates to minimizing the biomass. While, in the previous section, the conditions (*Full depletion*) and (First integral 1) were sufficient to determine the asymptotic behavior of the system, the presence of x in this particular problem does not allow a similar resolution. Thus, we write a relaxed version of the problem in terms of a constant ribosomal concentration r^*, which can provide an insight into the optimization problem.

Constant ribosomal concentration

The system with constant ribosomal concentration r^* writes

$$\begin{cases} \dot{s} = -w_M(s)(1-r^*)\mathcal{V}, \\ \dot{p} = w_M(s)(1-r^*) - \gamma w_R(p)(1-r^*) - w_R(p)r^*(p+1), \\ \dot{\mathcal{V}} = w_R(p)r^*\mathcal{V}. \end{cases} \tag{S_r}$$

It can be seen that the study of the asymptotic behavior of system (S) applies to (S_r) as the latter is a particular case of the original one (S) where $r(0) = u^* = r^*$. We then maximize the final product x^* in terms of the constant ribosomal concentration $r^* \in [r_{min}, r_{max}]$. We can see that the quantity

$$z = s + (p+1)\mathcal{V} + \gamma\frac{1-r^*}{r^*}\mathcal{V}$$

is constant. Thus,

$$\mathcal{V}^* + \gamma\frac{1-r^*}{r^*}(\mathcal{V}^* - \mathcal{V}_0) = \Sigma$$

which, using (3.6), yields

$$x^* = \gamma\frac{1-r^*}{r^*}(\mathcal{V}^* - \mathcal{V}_0).$$

Using the fact that $\mathcal{V}^* + x^* = \Sigma$ from (3.6), we see that x^* is monotone decreasing w.r.t. r^*, and so the value maximizing the product is $r^* = r_{\min}$. This is what one would expect intuitively in an infinite time horizon, as $r^* = r_{\min}$ favours the production of M, which catalyzes the production of X without arresting the production of biomass (given by the case $r^* = 0$, which cannot be attained in trajectories starting in Γ). However, in a finite horizon, a first phase dedicated to bacterial growth could also foster the production of X, which depends directly on the concentration of bacteria in the bioreactor.

3.5.2 Finite-time problem

Problem formulation

In current section, we study the metabolite production objective, in which the final concentration of metabolite in the bioreactor $x(t_f)$ is maximized. Maximizing the final biomass $\mathcal{V}(t_f)$ was already done for model ($S_{\mathcal{V}}$) representing a wild-type bacteria. However, it is likely that the presence of the heterologous pathway responsible for the production of x might affect the results already obtained. Thus, the two objectives will be compared in model (S) from a numerical perspective. Given a fixed final time $t_f > 0$, the OCP maximizing $c\mathcal{V}(t_f) + (1-c)x(t_f)$ (with $c = 0$ or $c = 1$ depending on the objective) with initial conditions (IC) writes

$$\begin{cases} \textit{maximize} \quad c\mathcal{V}(t_f) + (1-c)x(t_f), \\ \textit{subject to} \quad \textit{dynamics of (S)}, \\ \qquad \textit{and} \quad u(\cdot) \in \mathcal{U}, \end{cases} \tag{OCP$_x$}$$

One can easily see that, given the dynamics of the system, applying PMP would yield a Hamiltonian linear in the control for both values of c, which means that the solution of (OCP$_x$) is similar to that of (OCP$_{\mathcal{V}}$), given by expression (3.7). A numerical analysis of this results is provided in next section.

Numerical simulations

The optimal trajectories were obtained following the same procedure as in the biomass maximization case. Figures 3.7 and 3.8 illustrate optimal trajectories for the same set of initial conditions and same final time t_f. Figure 3.7 is the solution of (OCP$_x$) where the objective is the product maximization $x(t_f)$. As expected, and similar to the results obtained for (OCP$_{\mathcal{V}}$), the optimal control takes the value $u = 0$ for most of the interval, which promotes the synthesis of proteins of the metabolic machinery M, catalyzing the production of x. Additionally, according to our simulations, these results do not depend on the final time t_f: the final bang $u = 0$ of the optimal control is always predominant in the control strategy. The finite-time numerical results match the results obtained for the infinite-time case, in which the ribosomal sector of the cell should be minimized to maximize the production of x. Figure 3.8 shows an optimal trajectory solution of (OCP$_x$) with cost function $\mathcal{V}(t_f)$.

In this case, the allocation strategy is described by an initial bang $u = 1$ followed by a singular arc that takes up most of the optimal solution, with values near to an intermediate strategy $u = 0.5$; and a small bang $u = 0$ at the end. Such strategy leads to a bacterial composition much more balanced between ribosomal and enzymatic proteins, in opposition to the metabolite production case, where most of the bacterial proteins were dedicated to the metabolic machinery. The latter behavior illustrates a natural trade-off between two opposed strategies: maximizing the number of ribosomes to prioritize the synthesis of macromolecules over the production of x and, at the same time, maximizing the enzymatic activity in order to consume the substrate in the medium as fast as possible.

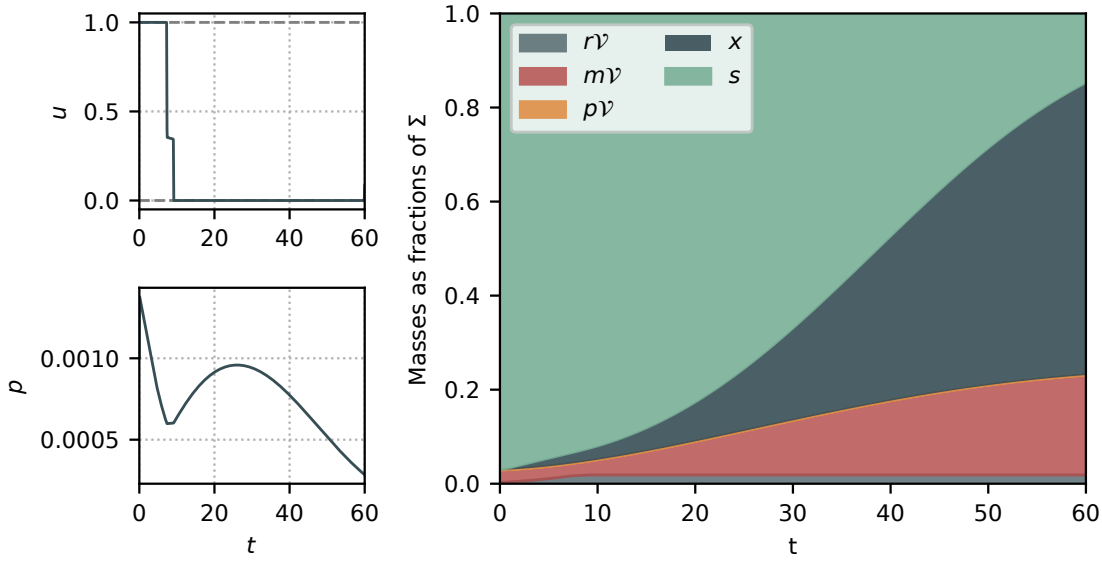


Figure 3.7: Solution of (OCP_x) for the metabolite maximization case $x(t_f)$, with $s_0 = 0.1$, $p_0 = 0.001$, $r_0 = 0.1$, $\mathcal{V}_0 = 0.003$ and $t_f = 60$. The final product concentration $x(t_f)$ is at 62% of Σ, while the final volume $\mathcal{V}(t_f)$ is only at 23% of Σ.

3.6 Discussion

This chapter described a preliminary study of resource allocation in bacteria for batch processing. A model considering the production of a value-added chemical compound is proposed, and a dynamical study of the system based on mass con-

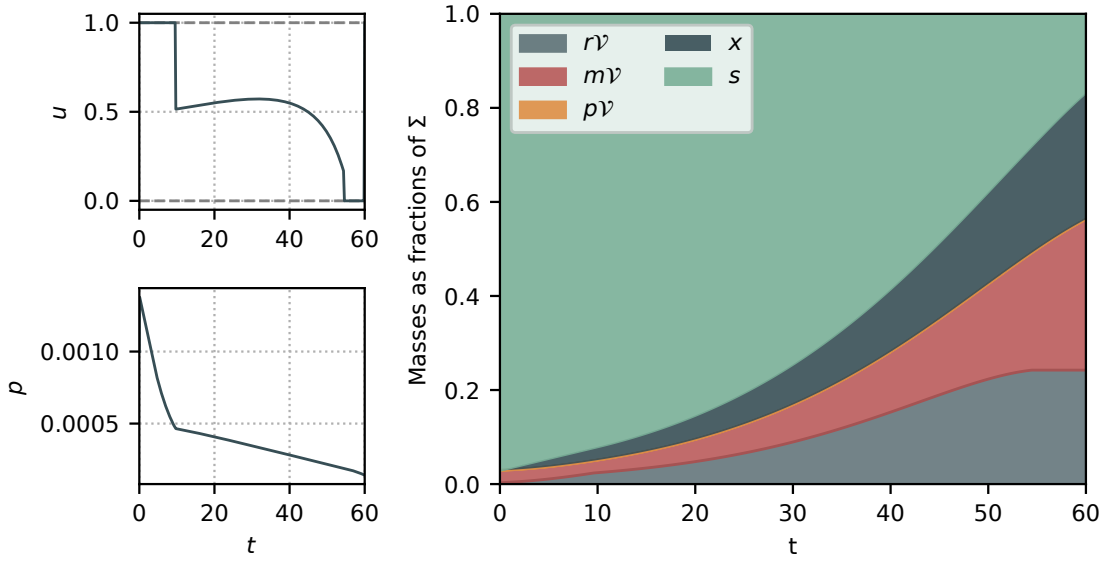


Figure 3.8: Solution of (OCP_x) for the biomass maximization case $\mathcal{V}(t_f)$, with $s_0 = 0.1$, $p_0 = 0.001$, $r_0 = 0.1$, $\mathcal{V}_0 = 0.003$ and $t_f = 60$. The final product concentration $x(t_f)$ is at 27% of Σ, while the final volume $\mathcal{V}(t_f)$ is at 56% of Σ.

servation laws shows that, under all possible resource allocation strategies, all the substrate in the medium is consumed. Then, the particular case of a wild-type bacteria with no metabolite production is analyzed, showing that the optimal allocation propitious for biomass production should yield a very high m/r ratio in the cell. Paradoxically, for the metabolite production case, these kind of strategies would rather maximize the production of x, while for maximizing the biomass under the presence of the heterologous pathway, a more balanced allocation is required. This work will be readily completed with a more detailed analysis of the impact of the final time in the allocation decision.

Acknowledgments

We would like to acknowledge the help of Sacha Psalmon and Baptiste Schall from Polytech Nice Sophia for the numerical simulations and the production of the online example in the control toolbox gallery.

Chapter 4

Continuous bioreactor

This chapter reproduces [AYT4] and [AYT2], published in the 21th IFAC World Congress and the Mathematical Biosciences and Engineering journal respectively.

4.1 Introduction

Microorganisms continuously have to contend with environmental changes in nature, and so they have evolved to accordingly adapt their physiology to cope with this unsteadiness. This is done by reorganizing the gene expression machinery, which is accomplished by dynamically allocating resources to different cellular functions. Microorganisms like bacteria exhibit great genetic variability, which is the main driver of natural selection, a phenomenon that depends on the continuous mutations in living populations. For this reason, they have achieved highly optimized resource management strategies. Such is the case of *E. Coli*: For specific environmental conditions, they seek to maximize their growth rate, which is a naturally evolved characteristic of this bacterium, and well known in the scientific community [14]. In this regard, optimality of bacterial organisms has been a central hypothesis in several research allocation studies [21].

Nevertheless, most of previous works in the literature consider the resource allocation problem in steady-state, without taking into account the changing conditions of natural environments [54]. This motivated a series of works with a dynamical-oriented approach to the problem [30, 8], that aimed to shed light on how bacteria tune their allocation mechanisms when facing changes on the nutrient concentration of the medium. Using Optimal Control theory, they investigate how their inter-

nal pathways can be dynamically readjusted so as to maximize their growth rate, obtaining different strategies that can be related to well known natural regulation mechanisms (such as the ppGpp-mediated sensing of the pool of aminoacids). The approach is based on a widely used modeling technique in systems biology: The so-called coarse-grained self-replicator models, used in bacterial growth representations for their simplicity and their capacity to reproduce observed experimental behaviors [45]. Although this kind of single-cell models are somewhat limited when predicting complex phenomena, they can accurately account for bacterial culture growth laws under the right assumptions (e.g., homogeneity of the culture) [55].

From an industry point of view, a natural question triggered by these studies is: How can we divert the natural allocation strategies of bacteria to improve current biotechnological production schemes? Such is the case of the artificial synthesis of metabolites or proteins of interest. The synthesis of such compounds is highly relevant for its wide range of applications: Production of antitumor agents, insulin, antibiotics, immunosuppressive agents and insecticides, among others [56, 7]. Motivated by the increasing understanding of the biosynthetic properties of certain microorganisms, research on this area can potentially lead to more efficient and sustainable production schemes. This is the matter addressed in recent work using a strain of *Escherichia coli* that includes an artificially engineered heterologous pathway for the production of a certain metabolite of interest [3, 74, AYT5]. In this approach, a control loop is developed through a bacterial growth switch that allows to externally modify the natural resource allocation decision [44]. The mechanism is implemented by re-engineering the transcriptional control of the expression of RNA polymerase, a key component of the gene expression machinery. This way, it is possible to optimize the productivity of the bioprocess by channeling resources into this new heterologous pathway.

At the same time, the synthesis of these metabolites draws resources from the native pathways used for producing biomass in bacteria, thus leading to an inherent compromise between these two objectives. One possible approach to this trade-off is to model it through different cost functions, thus obtaining multi-objective optimization problems. This is the case of [41], where the authors aim to maximize the production of a metabolite while minimizing the genetic burden caused by pathway expression. In contrast to this method, the work of [3] models the main trade-offs behind the process through a single decision parameter, which considerably reduces

the complexity of the optimization problem.

In a similar vein, we address a classic production scheme: the Continuous Stirred-Tank Reactor (CSTR), also known as chemostat [57]. While resource allocation in bacteria has been vastly studied in constant environments (e.g., under the assumption that there is always enough substrate in the medium), how this goes in continuous bioreactors is not trivial, since a feedback occurs from the physiology of the cell to the environmental conditions given by the interaction bacteria-medium [26]. Examples of self-replicator models in continuous bioreactors can be found in some recent works: In [58], authors use a coarse-grained self-replicator kinetic model of *Saccharomyces cerevisiae's* in a continuous bioreactor to investigate the trade-off between respiratory and fermentative metabolism, showing that optimal strategies are 'pure' metabolic strategies (e.g., either respiration of fermentation, but not respiro-fermentation). Likewise, it is rather classical in the continuous bioreactor scheme to maximize a certain performance measure (i.e., biomass production) in terms of the operational parameters related to the setup, such as dilution rate and/or concentration of the substrate inflow [59, 60, 61, 62, 63]. An example can be seen in [64], where the infinite-time optimal control problem of maximizing the average biogas production in terms of the dilution rate is studied. However, incorporating the aforementioned external control—that can disrupt the natural allocation process of the whole culture—provides an extra degree of freedom, which can, in turn, contribute to further improve the classical production scheme.

Based on the presented works, we show novel results addressing the problem of bacterial resource allocation in the CSTR framework. Our approach is based on a coarse-grained self-replicator dynamical model that accounts for the microbial culture growth inside the continuous bioreactor, and incorporates the external allocation control previously described. The novelty of the approach lies in the combination of the resource allocation control scheme, and the capacity to regulate the bacterial growth rate through the dilution rate of the continuous bioreactor. Further on, we study its asymptotic behavior using dynamical systems theory, and we provide conditions for the persistence of the bacterial population. The analysis is carried out by studying the local and global behavior of its limiting system, and relating its convergence to the original model through the theory of asymptotically autonomous systems [65]. Then, we pose the problem of maximizing the synthesis of the metabolite of interest during a fixed interval of time in terms of the resource

allocation decision. The latter is expressed as an OCP (Optimal Control Problem) which is then explored using PMP (Pontryagin's Maximum Principle) [31]. We analyze the solution of the problem and propose a sub-optimal control strategy given by a constant allocation decision, which eventually takes the system to the optimal steady-state production regime. On this basis, we study and compare the two most significant steady-state production objectives in CSTRs: Biomass production and metabolite production. For this last purpose, and in addition to the allocation parameter, we control the constant volumetric flow of the bioreactor (or dilution rate), and we analyze the results through a numerical approach. The resulting two-dimensional optimization problem is defined in terms of Michaelis-Menten kinetics with the parameter values of [30], and taking into account the constraints for the existence of the equilibrium of interest.

4.2 Model definition

4.2.1 Self-replicator model

In this section, we define the coarse-grained self-replicator model including the continuous bioreactor scheme and the allocation parameter. We consider a growing bacterial population in a CSTR bioreactor of constant volume $\mathcal{V}_{\text{ext}}$ [L]. The self-replicating system that models the culture is composed of the metabolic machinery M (transporters, enzymes...) and the gene expression machinery R (RNA polymerase, ribosomes...), both responsible of the cell growth. The validity of this single-cell model as a representation of a growing bacterial culture depends on a number of simplifying assumptions, one of them being that individual cells share the same macromolecular composition. For the production of the metabolite of interest X, we consider the artificially engineered pathway introduced in [3] (Figure 4.1).

The model represents three chemical macroreactions,

$$\mathrm{S} \xrightarrow{\mathrm{V_M}} \mathrm{P}$$
$$\mathrm{P} \xrightarrow{\mathrm{V_R}} \alpha\mathrm{R} + (1\text{-}\alpha)\,\mathrm{M}$$
$$\mathrm{P} \xrightarrow{\mathrm{V_X}} \mathrm{X}$$

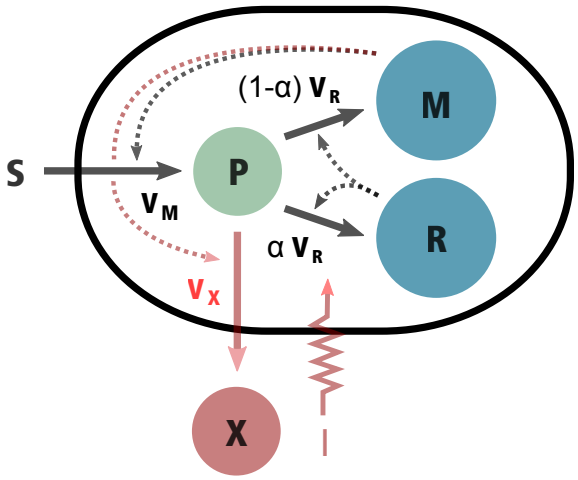


Figure 4.1: Extended coarse-grained self replicator model introduced in [3]. The substrate in the bioreactor (S_1) is consumed by the bacterial culture and transformed into precursors P through the action of the metabolic machinery M. Then, precursors are used to make macromolecules of the gene expression machinery R and the metabolic machinery M, with proportions α and $1-\alpha$ respectively; and to synthesize the metabolite X, which is excreted from the cell to the bioreactor. The external control I affects how the precursors P are distributed between both cellular functions M and R.

The first reaction is catalyzed by M and describes the transformation of external substrate S into precursor metabolites P at a rate V_M. The second one represents the conversion of precursors into macromolecules R and M and is catalyzed by R, at a rate V_R. Finally, the third reaction describes the transformation of precursors P into product X at a rate V_X, and catalyzed by M. The natural resource allocation parameter is modeled through the dimensionless parameter $\alpha \in [0,1]$, that represents the proportion of precursors allocated to the gene expression machinery R, while $1-\alpha$ indicates that of the metabolic machinery M. The model describes all mass quantities S, P, M, R and X in grams, and the rates in grams per hour. A constant volumetric flow rate F [L h^{-1}] produces an inflow of fresh medium rich in substrate, and an outflow of bacterial culture and metabolites [46]. As already stated, we include in our scheme the growth switch described in [44] to externally affect the allocation decision α by varying the inducer concentration in the medium. The mechanism is modeled as

$$u(t) = I(t)\,\alpha(t), \quad u \in [0,1] \tag{4.1}$$

where I is the external signal, which acts independently from the allocation parameter α. The aggregated form of equation (4.1) accounts for the fact that the synthetic switch is capable of adjusting bacterial growth between zero (by setting $I = 0$, which yields $u = 0$) and the maximal growth rate supported by the medium (given by $I = 1$, which yields $u = \alpha$). In this work, we limit the analysis to calculating the control input u, without decoupling both individual controls I and α (refer to Section 5 in [3] for more details on the implementation of this external controller). Then, we write the system of differential equations describing the evolution of the mass of each component,

$$\begin{cases} \dot{S} = V_{S_{in}} - V_M - V_{S_{out}}, \\ \dot{P} = V_M - V_R - V_X - V_{P_{out}}, \\ \dot{M} = (1-u)\, V_R - V_{M_{out}}, \\ \dot{R} = uV_R - V_{R_{out}}, \\ \dot{X} = V_X - V_{X_{out}}. \end{cases}$$

The inflow/outflow rates are defined as

$$V_{S_{out}} = DS, \quad V_{P_{out}} = DP, \quad V_{M_{out}} = DM, \quad V_{R_{out}} = DR, \quad V_{X_{out}} = DX, \quad V_{S_{in}} = Fs_{in},$$

where s_{in} [g L^{-1}] is the nutrient concentration of the inflow of fresh medium, and D [h^{-1}] the dilution rate defined as

$$D \doteq \frac{F}{\mathcal{V}_{\text{ext}}}.$$

We define the volume of the cell population $\mathcal{V}$ [L] as

$$\mathcal{V} \doteq \beta(M + R), \tag{4.2}$$

where β [L g^{-1}] is the inverse of the cytoplasmic density. The above definition is based on the assumption that the cytoplasmic density of the cells is constant for the whole culture, and it also takes into account the experimental results showing that macromolecules are responsible for most of the biomass in microbial cells [6]. Thus, the mass of precursors P is excluded from the volume $\mathcal{V}$ as a simplifying assumption.

Then, we express the quantities of the system as concentrations with respect to the volumes,

$$p \doteq \frac{P}{\mathcal{V}}, \quad r \doteq \frac{R}{\mathcal{V}}, \quad m \doteq \frac{M}{\mathcal{V}}, \quad s \doteq \frac{S}{\mathcal{V}_{\text{ext}}}, \quad x \doteq \frac{X}{\mathcal{V}_{\text{ext}}}, \tag{4.3}$$

where p, r and m [g L^{-1}] are intracellular concentrations of precursor metabolites, components of the gene expression machinery and metabolic enzymes respectively; and s and x [g L^{-1}] the extracellular concentrations of substrate and metabolite. It is worth stressing that intracellular concentrations are defined with respect to the bacterial volume $\mathcal{V}$, while extracellular concentrations with respect to the volume of the bioreactor $\mathcal{V}_{\text{ext}}$. Then, using definition (4.2), we obtain that $m + r = 1/\beta$. We define the growth rate of the bacterial population μ [h^{-1}] as the relative variation of cell volume $\dot{\mathcal{V}}/\mathcal{V}$ without considering the effect of the outflowing biomass. Replacing with concentrations leads to the system

$$\begin{cases} \dot{s} = D(s_{in} - s) - v_M(s,m)\dfrac{\mathcal{V}}{\mathcal{V}_{\text{ext}}}, \\ \dot{p} = v_M(s,m) - v_R(p,r) - v_X(p,m) - \mu(p,r)p, \\ \dot{r} = u\, v_R(p,r) - \mu(p,r)r, \\ \dot{m} = (1-u)\, v_R(p,r) - \mu(p,r)m, \\ \dot{x} = v_X(p,m)\dfrac{\mathcal{V}}{\mathcal{V}_{\text{ext}}} - Dx, \\ \dot{\mathcal{V}} = (\mu(p,r) - D)\, \mathcal{V}, \end{cases} \tag{S}$$

where $v_M(s,m)$, $v_R(p,r)$ and $v_X(p,m)$ [g L^{-1} h^{-1}] are the mass fluxes per unit volume obtained from dividing the rates V_M, V_R and V_X by $\mathcal{V}$, function of the concentrations of system (S_1). In this new system, the growth rate becomes

$$\mu(p,r) \doteq \left.\frac{\dot{\mathcal{V}}}{\mathcal{V}}\right|_{F=0} = \left.\frac{\dot{M} + \dot{R}}{M + R}\right|_{F=0} = \beta v_R(p,r)$$

showing that the bacterial growth rate is proportional to the macromolecule synthesis rate. We propose a change of variables that simplifies the expressions of the

system

$$\hat{s} = \beta s, \quad \hat{p} = \beta p, \quad \hat{r} = \beta r, \quad \hat{m} = \beta m, \quad \hat{x} = \beta x, \quad \hat{\mathcal{V}} = \frac{\mathcal{V}}{\mathcal{V}_{\text{ext}}},$$

which yields the relation

$$\hat{r} + \hat{m} = 1, \tag{4.4}$$

and we define the non-dimensional synthesis rates

$$\hat{v}_M(\hat{s}, \hat{m}) = \beta v_M(s, m), \quad \hat{v}_R(\hat{p}, \hat{r}) = \beta v_R(p, r), \quad \hat{v}_X(\hat{p}, \hat{m}) = \beta v_X(p, m), \tag{4.5}$$

and the non-dimensional substrate concentration $\hat{s}_{in} = \beta s_{in}$. Then, dropping all hats yields the following system

$$\begin{cases} \dot{s} = D(s_{in} - s) - v_M(s, 1-r)\mathcal{V}, \\ \dot{p} = v_M(s, 1-r) - v_X(p, 1-r) - \mu(p, r)(p+1), \\ \dot{r} = (u - r)\,\mu(p, r), \\ \dot{x} = v_X(p, 1-r)\mathcal{V} - Dx, \\ \dot{\mathcal{V}} = (\mu(p, r) - D)\,\mathcal{V}, \end{cases} \tag{S$_1$}$$

where the dynamical expression of m has been omitted since it can be computed from r, as shown in equation (4.4). It can also be seen that both concentrations m and r are limited to the interval $[0, 1]$ due to physical constraints from the relation in equation (4.4). For this latter, and due to the nature of the reactions involved in the studied problem, we will make some assumptions on the synthesis rates $v_M(s, m)$, $v_R(p, r)$ and $v_X(p, m)$:

Assumption 1. *Functions $v_M(s, m)$, $v_R(p, r)$ and $v_X(p, m)$ meet*

- $v_i(y, z) : \mathbb{R}^+ \times [0, 1] \to \mathbb{R}^+$
- $v_i(y, z)$ *continuously differentiable w.r.t. both variables*

- $v_i(0, z) = v_i(y, 0) = 0$
- $v_i(\cdot)$ *strictly monotonically increasing:*

$$\frac{\partial v_i}{\partial y}(y, z) > 0, \forall (y, z) \in \mathbb{R}_{>0} \times (0, 1], \qquad \frac{\partial v_i}{\partial z}(y, z) > 0, \forall (y, z) \in \mathbb{R}_{>0} \times (0, 1]$$

- $v_i(\cdot)$ *bounded w.r.t.* y*:* $\lim_{y\to\infty} v_i(y, z) = v_{i,max}(z)$.

Assumption 1 encompasses all general monotone increasing kinetics models used in biological models, such as Michaelis-Menten or Hill equation-based kinetics [66]. In this first work, we will focus on a particular kind of systems where the synthesis rate related to the metabolite production depends on the growth rate:

Assumption 2. *For* $r \in (0, 1)$*, the metabolite synthesis rate* $v_X(p, 1 - r)$ *can be expressed in terms of the macromolecule synthesis rate* $v_R(p, r)$ *(i.e., the growth rate)*

$$v_X(p, 1 - r) = c(r)\, v_R(p, r),$$

where $c(r) : (0, 1) \to \mathbb{R}^+$ *is a positive continuously differentiable function.*

As previously described, the reaction $v_X(p, m)$ is catalyzed by m, and the reaction $v_R(p, r)$ is catalyzed by r, meaning that the ratio between M and R in the microbial culture determines whether the resources are being allocated to the production of biomass or metabolite. This represents the trade-off described in the introduction of this paper, which is here modeled through the function $c(r)$. The fact that c does not depend on the concentration of precursors implies that the host cell has the same affinity to synthesize both biomass and metabolite from the precursors, even when the reactions are not expected to consume the precursors in the same proportion. For the particular case of Michaelis-Menten kinetics, this phenomenon is captured by the half-saturation constant [67]. Notably, for fixed values of r, both

synthesis rates are simply proportional. This assumption reduces (S_1) to

$$\begin{cases} \dot{s} = D(s_{in} - s) - v_M(s, 1-r)\mathcal{V}, \\ \dot{p} = v_M(s, 1-r) - \mu(p, r)(p + c(r) + 1), \\ \dot{r} = (u - r)\,\mu(p, r), \\ \dot{x} = c(r)\mu(p, r)\mathcal{V} - Dx, \\ \dot{\mathcal{V}} = (\mu(p, r) - D)\,\mathcal{V}. \end{cases} \tag{S_1}$$

We note that both Assumptions 1 and 2 are formulated for the non-dimensional synthesis rates given in definition (4.5), but they also hold for the original functions v_M, v_R and v_X (taking into account that, for these functions, the domain is defined as $\mathbb{R}^+ \times [0, \frac{1}{\beta}] \to \mathbb{R}^+$).

4.2.2 Asymptotic behavior

The asymptotic behavior of system (S_1) describes the "open-loop" operation mode of the continuous bioreactor, where the resource allocation control $u(t)$ is fixed to $\bar{u} \in (0, 1)$. In the present section we propose a series of mass conservation laws that allow to reduce (S_1) to a 3-dimensional limiting system. Then, we study the local stability of their equilibria, and we show, using the theory of asymptotically autonomous systems, that the full system (S_1) converges to the equilibria of its limiting system. Let us start the analysis of system (S_1) by defining its invariant region in the following lemma.

Lemma 1. *The set*

$$\Gamma = \left\{(s, p, r, x, \mathcal{V}) \in \mathbb{R}^5 \,:\, s_{in} \geq s > 0,\, p \geq 0,\, x \geq 0, 1 \geq r \geq 0,\, \mathcal{V} \geq 0\right\}$$

is positively invariant for the initial value problem.

Proof. Let us analyze the boundaries of Γ,

$$\begin{aligned}
&\dot{s}|_{s=0} = Ds_{in} > 0 && \Rightarrow \quad s = 0 \textit{ is repulsive.}\\
&\dot{s}|_{s=s_{in}} = -v_M(s, 1-r)\mathcal{V} \leq 0 && \Rightarrow \quad s = s_{in} \textit{ is repulsive/invariant.}\\
&\dot{p}|_{p=0} = v_M(s, 1-r) \geq 0 && \Rightarrow \quad p = 0 \textit{ is repulsive/invariant.}\\
&\dot{r}|_{r=0} = 0 && \Rightarrow \quad r = 0 \textit{ is invariant.}\\
&\dot{r}|_{r=1} = (u-1)\,\mu(p, 1) < 0 && \Rightarrow \quad r = 1 \textit{ is repulsive.}\\
&\dot{x}|_{x=0} = v_X(p, m)\mathcal{V} \geq 0 && \Rightarrow \quad x = 0 \textit{ is repulsive/invariant.}\\
&\dot{\mathcal{V}}|_{\mathcal{V}=0} = 0 && \Rightarrow \quad \mathcal{V} = 0 \textit{ is invariant.}
\end{aligned}$$

□

We will study the initial value problem of system (S_1) with initial conditions

$$s_{in} \geq s(0) \geq 0, \quad p(0) \geq 0, \quad 1 \geq r(0) > 0, \quad x(0) \geq 0, \quad \mathcal{V}(0) > 0, \tag{4.6}$$

where two cases have been excluded for being trivial to the analysis: $\mathcal{V}(0) = 0$, since it is necessary to have an initial amount of biomass to have bacterial growth; and $r(0) = 0$, since an empty gene expression machinery pool implies null growth rate $\mu(p, 0) = 0$, and therefore it is not possible to self-replicate from that point.

Mass conservation

System (S_1) can be rewritten as

$$\begin{cases} \dot{\varphi} = D\,(s_{in}\vec{v_{in}} - \vec{v_{out}}) + N\vec{v_i} - \vec{v_\mu}\,\mu(p, r), \\ \dot{\mathcal{V}} = (\mu(p, r) - D)\mathcal{V}, \end{cases}$$

where

- $\varphi \doteq [s, p, r, m, x]^T$ is the state vector of concentrations in the bioreactor.
- N is the stoichiometry matrix of the macroreactions.
- v_i is the vector of internal synthesis rates.
- v_{in} and v_{out} are the vectors of inflows and outflows respectively, associated to the continuous bioreactor setup.

- v_μ is the vector modeling the dilution effect due to variation of the bacterial volume.

which are defined as

$$N \doteq \begin{bmatrix} -\mathcal{V} & 0 & 0 \\ 1 & -1 & -1 \\ 0 & u & 0 \\ 0 & 1-u & 0 \\ 0 & 0 & \mathcal{V} \end{bmatrix}, \qquad \vec{v_i} \doteq \begin{bmatrix} v_M(s, 1-r) \\ v_R(p, r) \\ v_X(p, 1-r) \end{bmatrix}, \qquad \begin{aligned} \vec{v_{in}} &\doteq [1,0,0,0,0]^T, \\ \vec{v_{out}} &\doteq \operatorname{diag}(\varphi)\,[1,0,0,0,1]^T, \\ \vec{v_\mu} &\doteq \operatorname{diag}(\varphi)\,[0,1,1,1,0]^T. \end{aligned}$$

By analyzing the left null space of N, it can be seen that there are two mass conservation laws related to the total mass inside the bioreactor.

Definition 1. *We define the quantities*

$$\begin{aligned} w_1 &\doteq s + (p + m + r)\,\mathcal{V} + x = s + (p+1)\,\mathcal{V} + x, \\ w_2 &\doteq s + \left(p + \frac{r}{\bar{u}}\right)\mathcal{V} + x. \end{aligned}$$

The first quantity tends asymptotically to $w_1 = s_{in}$ as $t \to \infty$ for every input $u(t)$, as it obeys the dynamical equation

$$\dot{w}_1 = D\,(s_{in} - w_1)\,. \tag{4.7}$$

Moreover, when fixing $u(t)$ to $\bar{u} \in (0,1)$, the quantity w_2 also obeys the same equation (4.7), meaning that this second quantity also converges to s_{in}, which greatly simplifies the analysis of the asymptotic behavior of the system.

Lemma 2. *The ω-limit set of any solution of system* (S_1) *lies in the hyperplane*

$$\Omega_1 \doteq \left\{(s,p,r,x,\mathcal{V}) \in \mathbb{R}^5 \,:\, s + (p+1)\,\mathcal{V} + x = s_{in}\right\}.$$

Moreover, under constant input $u(t) = \bar{u}$, this is also true for the hyperplane

$$\Omega_2 \doteq \left\{(s,p,r,x,\mathcal{V}) \in \mathbb{R}^5 \,:\, s + \left(p + \frac{r}{\bar{u}}\right)\mathcal{V} + x = s_{in}\right\}.$$

Further on, we will use Lemma 2 to analyze system (S_1) through its limiting system.

Limiting systems

Lemma 2 presents two mass conservation laws that can be used to reduce subsystem (S_1) by two dimensions. We will first analyze the asymptotic behavior of concentration r: when $t \to \infty$, quantities $w_1 = w_2 = s_{in}$, so

$$s + (p+1)\,\mathcal{V} + x = s + \left(p + \frac{r}{\bar{u}}\right)\mathcal{V} + x \quad \Rightarrow \quad r = \bar{u}$$

meaning that, as $t \to \infty$, r will converge to the value $\bar{u}$. We can also express $x = s_{in} - s - (p+1)\mathcal{V}$, so that the limiting system of (S_1) becomes

$$\begin{cases} \dot{s} = D(s_{in} - s) - \bar{v}_M(s)\mathcal{V}, \\ \dot{p} = \bar{v}_M(s) - \bar{\mu}(p)(p + \bar{c} + 1), \\ \dot{\mathcal{V}} = (\bar{\mu}(p) - D)\,\mathcal{V}, \end{cases} \tag{S_1'}$$

where

$$\bar{v}_M(s) \doteq v_M(s, 1-\bar{u}), \quad \bar{v}_R(p) \doteq v_R(p, \bar{u}), \quad \bar{v}_X(p) \doteq v_X(p, 1-\bar{u}),$$
$$\bar{\mu}(p) \doteq \mu(p, \bar{u}), \quad \bar{c} \doteq c(\bar{u}).$$

Details on the convergence of the limiting system (S_1') to the original one (S_1) will be addressed later in the article. In next section, we will fully describe the asymptotic behavior of (S_1').

Local stability

In the interest of simplifying the notation, we define the following function.

Definition 2. *We define the function*

$$\bar{f}(p) \doteq \bar{v}_R(p) + \bar{v}_X(p) + \bar{\mu}(p)p = \bar{\mu}(p)(p + \bar{c} + 1),$$

Function f meets $\bar{f}(p) > 0, \bar{f}'(p) > 0, \forall p \in \Gamma$ (positive and monotonically in-

creasing).

The main result of local stability study is summarized in Theorem 1.

Theorem 1. *The local stability of equilibria is given by the following criterion.*

- *If* $\bar{\mu}(p_w) \geq D$*:*
  - *The interior equilibrium* E_i *exists, is unique and locally stable.*
  - *The washout equilibrium* E_w *exists, is unique and locally unstable.*
- *If* $\bar{\mu}(p_w) < D$*:*
  - *The interior equilibrium* E_i *does not exist.*
  - *The washout equilibrium* E_w *exists, is unique and locally stable.*

In the above criterion, the equilibria are defined as follows:

- *The interior equilibrium* $E_i \doteq (s_i, p_i, \mathcal{V}_i)$*, with*

$$p_i : \left\{ p \in \mathbb{R}^+ \,:\, \bar{\mu}(p) = D \right\}, \tag{4.8}$$

$$s_i : \left\{ s \in \mathbb{R}^+ \,:\, \bar{v}_M(s) = \bar{f}(p_i) \right\}, \tag{4.9}$$

$$\mathcal{V}_i \doteq \frac{D(s_{in} - s_i)}{\bar{v}_M(s_i)}. \tag{4.10}$$

- *The washout equilibrium* $E_w \doteq (s_{in}, p_w, 0)$*, with*

$$p_w : \left\{ p \in \mathbb{R}^+ \,:\, \bar{f}(p) = \bar{v}_M(s_{in}) \right\}. \tag{4.11}$$

Proof. First, we will prove the boundedness of p. Using definition (4.11), it is possible to define a constant upper-bound on p by analyzing the time-varying upper bound $p_{up}(t)$ with dynamical equation

$$\dot{p}_{up} \doteq \bar{v}_M(s_{in}) - \bar{f}(p_{up}). \tag{4.12}$$

In equation (4.12), p_{up} converges to the value p_w satisfying equation (4.11). Additionally, the vector field at $p = p_w$ is always negative (or null when $s = s_{in}$)

$$\dot{p}|_{p=p_w} = \bar{v}_M(s) - \bar{f}(p_w) \leq 0$$

meaning that $p = p_w$ is either repulsive or invariant, so a new invariant set $\Gamma' \subset \Gamma$ can be defined,

$$\Gamma' \doteq \left\{(s, p, \mathcal{V}) \in \mathbb{R}^3 \,:\, s_{in} \geq s > 0,\, p_w \geq p > 0,\, \mathcal{V} \geq 0\right\}. \tag{4.13}$$

Equilibrium E_i The existence of the interior equilibrium E_i is given by the boundedness of the flows,

$$\max \bar{\mu}(p) \geq D, \tag{4.14}$$

$$\max \bar{v}_M(s) \geq \bar{f}(p_i), \tag{4.15}$$

and its uniqueness can be proved through monotonicity arguments: In $\bar{\mu}(p) = D$, $\bar{\mu}(p)$ is strictly monotonically increasing w.r.t. p so, if inequality (4.14) is met, p_i should be unique. Similarly, in $\bar{v}_M(s) = \bar{f}(p_i)$, $\bar{v}_M(s)$ is again strictly monotonically increasing so, if inequality (4.15) is met, s_i should be unique. As is standard in continuous bioreactors, inequality (4.14) implies that the maximal growth rate of the bacterial population should be bigger than the dilution rate D. At the same time, the inequality (4.15) requires the maximal uptake flow to be bigger than the flows responsible for bacterial growth and metabolite production. In Γ', these conditions become

$$\bar{\mu}(p_w) \geq D, \tag{4.16}$$

where the second condition (4.15) is included in the inequality (4.16), as $max\ \bar{v}_M(s) = \bar{v}_M(s_{in}) = \bar{f}(p_w)$, and the condition $\bar{f}(p_w) \geq \bar{f}(p_i)$ is true if and only if $\mu(p_w) \geq D$. The Jacobian matrix is given by

$$J_i = \begin{bmatrix} -D - \bar{v}'_M(s)\mathcal{V}_i & 0 & -\bar{v}_M(s) \\ \bar{v}'_M(s) & \bar{f}'(p) & 0 \\ 0 & \bar{\mu}'(p)\mathcal{V}_i & 0 \end{bmatrix},$$

and the characteristic polynomial is

$$\begin{aligned} P_i(\lambda) &= (\lambda + D + \bar{v}'_M(s)\mathcal{V}_i)(\lambda + \bar{f}'(p))\lambda + \bar{v}_M(s)\bar{v}'_M(s)\bar{\mu}'(p)\mathcal{V}_i \\ &= \lambda^3 + \lambda^2(D + \bar{v}'_M(s)\mathcal{V}_i + \bar{f}'(p)) + \lambda(D + \bar{v}'_M(s)\mathcal{V}_i)\bar{f}'(p) + \bar{v}_M(s)\bar{v}'_M(s)\bar{\mu}'(p)\mathcal{V}_i. \end{aligned} \tag{4.17}$$

Using Definition 2, it can be seen that $-D$ is an eigenvalue by replacing in expression 4.17

$$P_i(-D) = -D\bar{v}'_M(s)\mathcal{V}_i\left(p+\bar{c}+1\right)\bar{\mu}'(p) + \bar{v}_M(s)\bar{v}'_M(s)\bar{\mu}'(p)\mathcal{V}_i = \bar{v}'_M(s)\mathcal{V}_i\bar{\mu}'(p)\dot{p} = 0.$$

Then, dividing $P_i(\lambda)$ by $\lambda + D$ yields

$$P_i(\lambda)\frac{1}{\lambda+D} = \lambda^2 + \lambda\underbrace{\left(D + \bar{v}'_M(s)\mathcal{V}_i + \left(p+\bar{c}+1\right)\bar{\mu}'(p)\right)}_{>0} + \underbrace{\bar{v}'_M(s)\mathcal{V}_i\left(p+\bar{c}+1\right)\bar{\mu}'(p)}_{>0},$$

which, by Routh-Hurwitz criteria, implies that all eigenvalues have negative real part. Then, if it exists, the equilibrium is always stable.

Equilibrium E_w The washout equilibrium E_w exists for all values of Γ', since the only condition for existence is given by

$$max\ \bar{f}(p) \geq \bar{v}_M(s_{in}),$$

which is always true in Γ' since $\bar{v}_M(s_{in}) = \bar{f}(p_w)$ and so the inequality becomes $\bar{f}(p_w) \geq \bar{f}(p_w)$. Again, as $\bar{f}(p)$ is strictly monotonically increasing, there is a unique solution p_w. Its stability is given by

$$J_w = \begin{bmatrix} -D & 0 & -\bar{c} \\ \bar{v}'_M(s) & -\bar{f}'(p) & 0 \\ 0 & 0 & \bar{\mu}(p_w) - D \end{bmatrix}$$

with characteristic polynomial

$$P_w(\lambda) = (\lambda + D)(\lambda + \bar{f}'(p))(\lambda - \bar{\mu}(p_w) + D).$$

Eigenvalues $\lambda = -D$ and $\lambda = -\bar{f}'(p)$ are always real and negative, and so the stability criterion becomes

$$E_w\ stable \quad \text{if and only if} \quad \bar{\mu}(p_w) < D.$$

□

Global analysis

In the current section, we show how the limiting system (S_1') can be further reduced using another mass conservation law given by the fact that $\bar{c}$ is constant. This is formalized in the following lemma.

Lemma 3. *The ω-limit set of any solution of the limiting system* (S_1') *lies in the hyperplane*

$$\Omega_3 \doteq \left\{(s, p, \mathcal{V}) \in \mathbb{R}^3 \,:\, s + (p + \bar{c} + 1)\,\mathcal{V} = s_{in}\right\}$$

Proof. The quantity

$$w_3 \doteq s + (p + \bar{c} + 1)\,\mathcal{V} \tag{4.18}$$

obeys the dynamical equations $\dot{w}_3 = D(s_{in} - w_3)$, which means that the system converges asymptotically to the limit set Ω_3. □

Using Lemma 2, we can further reduce system (S_1') by expressing

$$s = s_{in} - \mathcal{V}(p + \bar{c} + 1), \tag{4.19}$$

for the values $(p, \mathcal{V})$ that meet $s_{in} - \mathcal{V}(p + \bar{c} + 1) > 0$, and so the new limiting system becomes

$$\begin{cases} \dot{p} = \bar{v}_M\left(s(\cdot)\right) - \bar{\mu}(p)(p + \bar{c} + 1), \\ \dot{\mathcal{V}} = (\bar{\mu}(p) - D)\mathcal{V}. \end{cases} \tag{S_1''}$$

Since (S_1'') is a 2-dimensional continuous system, its global behavior (illustrated in Figure 4.2) can be studied through the Poincaré-Bendixson trichotomy.

Lemma 4. *Every solution of the limiting system* (S_1'') *with initial conditions* (4.6) *converges to*

- $E_i^* \doteq (p_i, \mathcal{V}_i)$ *if* $\bar{\mu}(p_w) \geq D$
- $E_w^* \doteq (p_w, 0)$ *if* $\bar{\mu}(p_w) < D$

Proof. The Poincaré-Bendixson trichotomy ensures that every non-empty compact ω-limit set of (S_1'') is either a fixed point, a periodic orbit or a cycle of equilibria. Then, by applying Poincaré-Bendixson theorem through the Dulac criterion, we can discard periodic orbits and cycles of equilibria:

$$\begin{aligned}\frac{\partial}{\partial p}\dot{p}+\frac{\partial}{\partial \mathcal{V}}\dot{\mathcal{V}} =&\frac{\partial}{\partial p}\bar{v}_M\left(s(\cdot)\right)-\bar{\mu}(p)-\bar{\mu}'(p)\left(p+\bar{c}+1\right)+\bar{\mu}(p)-D\\ =&\bar{v}_M'\left(s(\cdot)\right)\frac{\partial s(\cdot)}{\partial p}-\bar{\mu}'(p)\left(p+\bar{c}+1\right)-D<0,\end{aligned} \tag{4.20}$$

as $\partial s(\cdot)/\partial p = -\mathcal{V}$ from equation (4.19). This ensures that the new limiting system should converge to one of the stable equilibria as $t \to \infty$ which, according to Theorem 1, is known to be unique. □

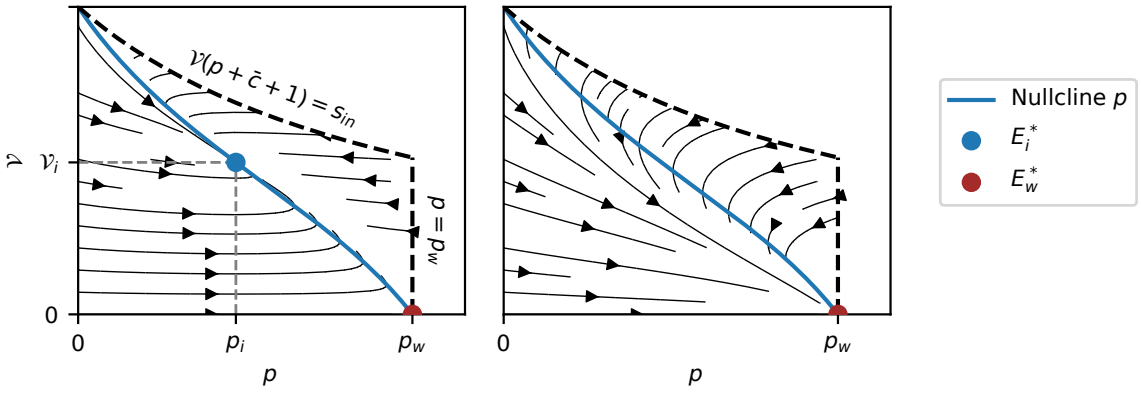


Figure 4.2: Phase plane of the limiting system S_1'' showing: The case where $\bar{\mu}(p_w) \geq D$ (left), so that the equilibrium E_i^* exists and attracts all solutions; and the case where $\bar{\mu}(p_w) < D$ (right) and E_w^* attracts all solutions.

Through the theory of asymptotically autonomous systems [65], we can relate the asymptotic behavior of the 2-dimensional limiting system S_1'' to that of the full 5-dimensional system (S_1). The main result is established in Theorem 2.

Theorem 2. *Every solution of system* (S_1) *with initial conditions* (4.6) *converges to,*

- *The extended interior equilibrium* $\hat{E}_i \doteq (s_i, p_i, \bar{u}, x_i, \mathcal{V}_i)$ *if it exists, with* $x_i \doteq \bar{c}\mathcal{V}_i$.

- *The extended washout equilibrium* $\hat{E}_w \doteq (s_{in}, p_w, \bar{u}, 0, 0)$ *if* $\hat{E}_i$ *does not exist.*

where the condition for the existence of the extended interior equilibrium is $\bar{\mu}(p_w) \geq D$.

Proof. We resort to a more particular case of the general theory of asymptotically autonomous systems (introduced in Appendix F of [46]), that requires a certain number of hypotheses to be met related to the limiting system S_1'', its equilibria and the mass conservation equations defined in Lemma 2:

(A1) The dynamical equations (4.7) and (4.18) of quantities w_1, w_2 and w_3 are stable.

(A2) S_1'' has 2 rest points E_i^* and E_w^*, which are hyperbolic.

(A3) If E_i^* exists, $dim(M^+(E_i^*)) = 2$ and $dim(M^+(E_w^*)) = 1$. If E_i^* does not exist, $dim(M^+(E_w^*)) = 2$.

(A4) S_1'' has no cycles of rest points, as shown in equation (4.20).

(A5) S_1'' has no periodic orbits, as shown in equation (4.20).

Then, almost all trajectories of the original system (S_1) converge to one of the asymptotically stable equilibria of the limiting system, which is always unique. □

4.3 Metabolite production

In this section, we aim to maximize the production of the metabolite of interest X. We start by modeling the synthesis rates as explicit functions of the concentrations of the system. While this latter is not required for most of the results, it enables the numerical simulations both for the dynamical and for the static optimization problems. In particular, we resort to the Michaelis-Menten kinetics defined in previous works [30] and the same biological constants. Then, we pose the problem of maximizing the production of X over a fixed period of time, and we approximate the optimal solution with a constant allocation strategy. This finally motivates the static optimization approach, that is solved in terms of the tuple $(D, \bar{u})$.

4.3.1 Kinetic's definition

Michaelis-Menten kinetics are defined as

$$\begin{aligned} v_M(s,m) &\doteq k_M\, m \frac{s}{K_M + s} = k_M\,(1-r)\frac{s}{K_M + s}, \\ v_R(p,r) &\doteq k_R\, r \frac{p}{K_R + p}, \\ v_X(p,m) &\doteq k_X\, m \frac{p}{K_X + p} = k_X\,(1-r)\frac{p}{K_X + p}, \end{aligned}$$

where the biological constants and their values [30, 3] are described in Table 4.1.

Table 4.1: Relevant biological system variables. RT and HSC stand for "Rate constant" and "Half-saturation constant" respectively.

Parameter	Description	Unit	Value
β	Inverse of the cytoplasmic density	$L\,g^{-1}$	0.003
k_M	RT of the metabolic macroreaction	h^{-1}	4.32
K_M	HSC of the metabolic macroreaction	$g\,L^{-1}$	33.33
k_R	RT of the protein synthesis reaction	h^{-1}	3.6
K_R	HSC of the protein synthesis reaction	$g\,L^{-1}$	1
k_X	RT of the metabolite synthesis reaction	h^{-1}	0.5
K_X	HSC of the metabolite synthesis reaction	$g\,L^{-1}$	1

Assumption 2 is implemented by setting the half-saturation constant $K_X = K_R$, such that the function $c(r)$ becomes

$$c(r) = \frac{v_X(p, 1-r)}{v_R(p,r)} = \frac{k_X}{k_R}\frac{1-r}{r}.$$

The substrate concentration of the inflow s_{in} is set to 0.4 g/L.

4.3.2 Dynamic optimization problem

The problem of maximizing the production of the metabolite X during a fixed interval of time T is explored in this section. As it is classical in the continuous bioreactor framework, the instantaneous production of metabolite is described by the quantity

$$DX \quad [g\,h^{-1}] \tag{4.21}$$

which, using definitions (4.3), can be expressed as $D\mathcal{V}_{\text{ext}}x$. Then, the total metabolite production over an interval T amounts to

$$\int_0^T D\,\mathcal{V}_{\text{ext}}\,x\,\mathrm{d}t \quad [g]. \tag{4.22}$$

We will be first interested in dynamically adjusting the aggregated control $u(t)$ so as to maximize the quantity (4.22) for a fixed dilution rate D. The problem is tackled through an optimal control approach, formulated as

$$\begin{cases} \textit{maximize} & J_x(u) = D\,\mathcal{V}_{\text{ext}} \displaystyle\int_0^T x(t)\,\mathrm{d}t \\ \textit{subject to} & \textit{dynamics of } (\mathrm{S}_1), \\ & u(\cdot) \in \mathcal{U}. \end{cases} \tag{OCP}$$

with $\mathcal{U}$ the set of admissible controllers, which are Lebesgue measurable real-valued functions defined on the interval $[0, T]$ and satisfying the constraint $u(t) \in [0, 1]$. We first note that, since D and $\mathcal{V}_{\text{ext}}$ are constants, the problem reduces to maximizing the integral of the concentration x. Thus, the nature of the OCP allows us to obtain a first result on the existence of the solutions.

Proposition 4.3.1. *The dynamic maximization problem has at least one solution.*

Proof. Since there are no terminal conditions, the set of admissible controls is not empty (any constant control within the prescribed bounds is admissible). For bounded controls ($u(t)$ belongs to $[0, 1]$) and for a fixed final time $T > 0$, the dynamics (S_1) cannot blow up in finite time so all trajectories remain in a fixed compact. Indeed, all state variables but $\mathcal{V}$ are bounded (as set (4.13) is invariant), and since

$$\dot{\mathcal{V}} = (\mu(p, r) - D)\mathcal{V}$$

with $\mu(p, r)$ bounded, the volume has at worst an exponential rate and is also bounded. As the dynamics is affine in the control, the set of velocities is convex and existence holds by Filippov's theorem [68]. □

In order to further explore the solution of (OCP), we define the adjoint state vector $\lambda \doteq (\lambda_s, \lambda_p, \lambda_r, \lambda_x, \lambda_{\mathcal{V}})$ associated to the state vector $\varphi \doteq (s, p, r, x, \mathcal{V})$, and

we write the Hamiltonian

$$H(\varphi, \delta, u) = \lambda^0 x + \langle \lambda, F(\varphi, u) \rangle,$$

where F represents the right-hand side of system (S_1) given by equations (S_1). Developing the expression we obtain

$$\begin{aligned} H = &\lambda_s \left(D(s_{in} - s) - v_M(s, 1-r)\mathcal{V}\right) + \lambda_p \Big(v_M(s, 1-r) - v_X(p, 1-r) - \mu(p, r)(p+1)\Big) \\ &+ \lambda_r \left(u - r\right) \mu(p, r) + \lambda_x \Big(v_X(p, 1-r)\mathcal{V} - Dx\Big) + \lambda_{\mathcal{V}}(\mu(p, r) - D)\mathcal{V} - \lambda_0 x, \end{aligned}$$

which shows that the Hamiltonian is linear in the control u, so it can be rewritten in the affine form

$$H = H_0 + uH_1,$$

where

$$\begin{aligned} H_0 = &\lambda_s \left(D(s_{in} - s) - v_M(s, 1-r)\mathcal{V}\right) + \lambda_p \Big(v_M(s, 1-r) - v_X(p, 1-r) - \mu(p, r)(p+1)\Big) \\ &- r\lambda_r \mu(p, r) + \lambda_x (v_X(p, 1-r)\mathcal{V} - Dx) + \lambda_{\mathcal{V}}(\mu(p, r) - D)\mathcal{V} - \lambda_0 x, \\ H_1 = &\lambda_r \mu(p, r). \end{aligned}$$

In the absence of terminal constraints, there are no abnormal extremals and λ_0 can be set to -1. Since the constrained optimal control u should maximize the Hamiltonian, one has

$$u_{OCP}(t) = \begin{cases} 0 & \text{if } H_1 < 0, \\ 1 & \text{if } H_1 > 0, \\ u_s(t) & \text{if } H_1 = 0. \end{cases}$$

These alternatives account for the possibility of bang and singular arcs, an extremal being in general an arbitrary concatenation of these. Bang controls correspond to pure allocation strategies $u = 0$ and $u = 1$ (i.e., purely geared towards either the metabolic machinery M or the gene expression machinery R respectively), while singular control $u_s(t)$ occurs when the function H_1 identically vanishes over some subinterval. Such singular arcs can be further described by successively differentiat-

ing the switching function H_1 until the singular control u can be explicitly computed. For the present case, this occurs when differentiating four times ("order two" singular arc). More precisely, the singular arc is of *local* (not *intrinsic*) order two (see [69]). In this case, although chattering (a.k.a. Fuller phenomenon) in and out is not necessary to enter and exit the singular arc, it is indeed what will be observed on the numerical simulations.

Theorem 3. *Any singular arc $u_s(t)$ of a normal extremal process solution of* (OCP) *is at least of order two.*

Proof. We assume that the process (φ, u, λ) that satisfies PMP is a normal extremal, and we set $\lambda_0 = -1$. In order to further describe the singular arc, we assume H_1 vanishes on a whole sub-interval $\tau = [t_1, t_2] \subset [0, T]$. Then, the switching surface is the set

$$\Sigma = \left\{ (\varphi, \lambda) \in \mathbb{R}^{2n} \,|\, H_1 = 0 \right\},$$

where, for this case, $n = 5$. The first derivative is

$$\dot{H}_1 = \dot{\lambda}_r \mu(p, r) + \lambda_r \Big(\mu_r(p, r)\dot{r} + \mu_p(p, r)\dot{p} \Big), \tag{4.23}$$

where

$$\mu_r(p, r) = \frac{\partial \mu(p, r)}{\partial r}, \quad \mu_p(p, r) = \frac{\partial \mu(p, r)}{\partial p}.$$

Along the singular arc H_1 is identically zero, so expression (4.23) also vanishes. In order to compute the successive derivatives of H_1, we will resort to the Poisson bracket operator [70],

$$\{f, g\} = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial \lambda_i} \frac{\partial g}{\partial \varphi_i} - \frac{\partial f}{\partial \varphi_i} \frac{\partial g}{\partial \lambda_i} \right).$$

Applying the latter definition to the derivative of H_1 we obtain

$$\begin{aligned} \dot{H}_1 = \frac{\partial H_1}{\partial \varphi}\dot{\varphi} + \frac{\partial H_1}{\partial \lambda}\dot{\lambda} &= \sum_{i=1}^{n} \left(\frac{\partial H}{\partial \lambda_i} \frac{\partial H_1}{\partial \varphi_i} - \frac{\partial H}{\partial \varphi_i} \frac{\partial H_1}{\partial \lambda_i} \right) \\ &= \{H, H_1\} = \{H_0 + uH_1, H_1\} = \{H_0, H_1\} \end{aligned}$$

as $\{uH_1, H_1\} = u\{H_1, H_1\} = 0$. In this context, we will use the notation H_{01} to refer to $\{H_0, H_1\}$, and so forth. The second derivative of H_1 is

$$\ddot{H}_1 = \dot{H}_{01} = \{H, H_{01}\} = \{H_0, H_{01}\} + u\{H_1, H_{01}\} = H_{001} + uH_{101}.$$

Showing that $H_{101} = 0$ along the switching surface Σ ensures that the singular arc is at least of order 2 (otherwise, a singular arc of order 1 could be computed as $u = -H_{001}/H_{101}$). Since H_1 depends only on λ_r, p and r, the computation of the Poisson bracket reduces to

$$\begin{aligned} H_{101} = \{H_1, H_{01}\} &= \frac{\partial H_1}{\partial \lambda_r}\frac{\partial H_{01}}{\partial r} - \frac{\partial H_1}{\partial r}\frac{\partial H_{01}}{\partial \lambda_r} + \frac{\partial H_1}{\partial \lambda_p}\frac{\partial H_{01}}{\partial p} - \frac{\partial H_1}{\partial p}\frac{\partial H_{01}}{\partial \lambda_p} \\ &= \mu(p,r)\frac{\partial H_{01}}{\partial r} - \lambda_r\mu_r(p,r)\frac{\partial H_{01}}{\partial \lambda_r} - \lambda_r\mu_p(p,r)\frac{\partial H_{01}}{\partial \lambda_p} \end{aligned} \tag{4.24}$$

where

$$\frac{\partial H_{01}}{\partial r} = \dot{\lambda}_r\mu_r(p,r) + \lambda_r\frac{\partial}{\partial r}\Big(\mu_r(p,r)\dot{r} + \mu_p(p,r)\dot{p}\Big). \tag{4.25}$$

Replacing the expression (4.25) in equation (4.24) yields

$$\begin{aligned} H_{101} =& \mu(p,r)\left[\dot{\lambda}_r\mu_r(p,r) + \lambda_r\frac{\partial}{\partial r}\Big(\mu_r(p,r)\dot{r} + \mu_p(p,r)\dot{p}\Big)\right] \\ &- \lambda_r\mu_r(p,r)\frac{\partial H_{01}}{\partial \lambda_r} - \lambda_r\mu_p(p,r)\frac{\partial H_{01}}{\partial \lambda_p} \end{aligned}$$

which is also equal to 0 along the switching surface since every term is multiplied either by λ_r or $\dot{\lambda}_r$, both identically zero along a singular arc (see H_1 expression), showing that the singular arc is at least of order 2. □

It is noteworthy that the above proof holds for all general flows $v_M(s, 1-r)$, $v_X(p, 1-r)$ and $v_R(p, r)$ considered in Assumption 1, with no need to use the defined Michaelis-Menten kinetics. Numerical results shown in Figure 4.3 confirm that, when the conditions for the existence of the interior equilibrium E_i are met, the optimal control strategy consists on a series of bang-bang arcs, and a singular arc that is entered and left through the chattering phenomenon. While, as already mentioned, it is compulsory to enter and leave order two singular arcs through chattering, a more precise description of the structure of the extremal would require a

deeper analysis (for instance to prove that there is only one singular arc). Moreover, when the simulation time is long enough, we observe that the singular control $u_s(t)$ converges to a constant value, eventually taking the system to steady-state. This is related to the so-called turnpike phenomenon [71] that relates the singular arc of the dynamic optimization problem with the solution of the static one. See, e.g., [72] for a preliminary analysis on a similar case.

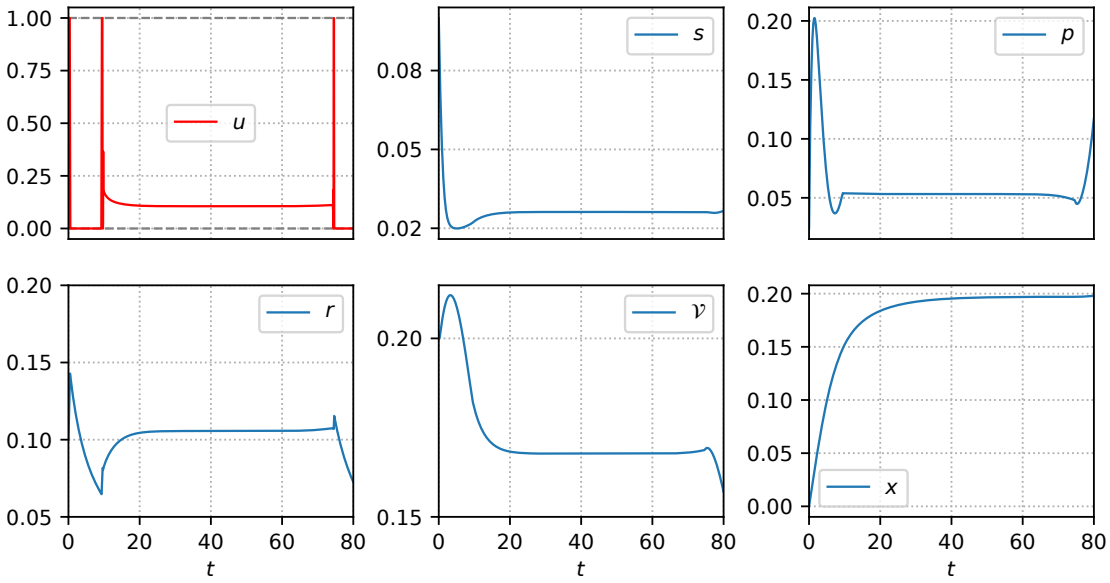


Figure 4.3: Results of the numerical simulations using BOCOP [4], showing the optimal control input u and the state variables. The simulation has been done in a time horizon $T = 80$, with 5000 time steps and *Midpoint* discretization method. The Ipopt (interior point nonlinear optimizer) arguments are: max_iter = 1000, tol $= 1.0e-14$, The initial conditions are fixed to: $s(0) = 0.1$, $p(0) = 0.024$, $r(0) = 0.1$, $x(0) = 0$, $\mathcal{V}(0) = 0.2$.

Additionally, we observe that the steady-state approached by the system during the singular regime maximizes the integrand of the cost function, which is in fact the instantaneous production of metabolite described in expression (4.21) (otherwise, it would not be the optimal strategy). Therefore, the static optimal control to which the dynamic one converges along the singular arc is the solution of the optimization

problem

$$\begin{cases} \textit{maximize} & J_x(\bar{u}) \doteq x \\ \textit{subject to} & \textit{dynamics of } (\mathrm{S}_1), \\ & \dot{\varphi} = 0, \\ & 0 \leq \bar{u} \leq 1. \end{cases} \tag{SP}$$

The dynamical optimal solution acts as the gold-standard in terms of what can be achieved through control techniques. However, implementing such strategy is, in practice, unfeasible. Consequently, it is possible to design a static suboptimal strategy consisting of a constant allocation u_{sp} that takes the system to the same steady-state which maximizes the integrand of the cost function. Numerical simulations of such strategy are shown in Figure 4.4, where it can be seen that the area below curves x_{sp} and x in OCP only differ slightly ($< 1\%$ for this particular simulation). The constant control u_{sp} basically disregards the initial and final bang arcs, as well as the chattering phenomena, which constitutes a small fraction of the complete time horizon T. Moreover, this fraction gets smaller as T becomes large, which is a typical feature of the turnpike effect for a specific class of optimal control problems [71]. This way, the difference between strategies becomes marginal in long-term production schemes.

Additionally, we provide a numerical computation of the successive derivatives of the switching function H_1 in Figure 4.5. It can be seen that the fourth derivative $H_{10001} \neq 0$ over the interval where $H_1 = 0$, which shows that the singular arc is of order 2. Furthermore, we verify that the Kelley (or *generalized Legendre-Clebsch* [73]) condition

$$(-1)^k \frac{\partial}{\partial u} \left[\frac{d^{2k}}{dt^{2k}} \left(\frac{\partial H}{\partial u} \right) \right] < 0, \quad \forall t \in \mathcal{I} \tag{4.26}$$

is met along the singular arc, which in this case is equivalent to $H_{10001} < 0$, $\forall t \in \mathcal{I}$. A check of this condition, necessary for optimality, is also shown in Figure 4.5. Although there is no available sufficient condition to test local optimality of extremals with Fuller arcs, verifying the Legendre-Clebsch condition along the singular arc only ensures that we do not compute a too crude local minimizer. Besides, numerically only a small finite number of bang arcs are retrieved by the optimizer for the chattering parts before and after the singular arc. This is usually sufficient to give

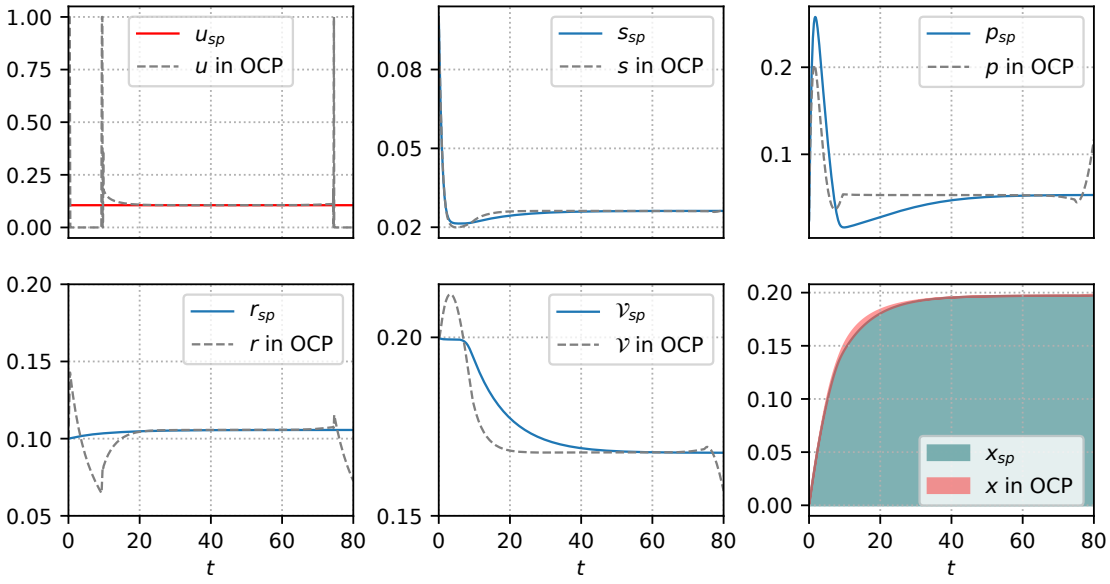


Figure 4.4: Results of the numerical simulations using BOCOP, comparing the control u solution of the OCP to the solution of the static optimization problem u_{sp} (both with same initial conditions). The parameters for the numerical simulation match the ones used in Figure 4.3. The last plot emphasizes the area below the curves x_{sp} and x in OCP, as they are proportional to the total mass of metabolite produced, which is the quantity to be maximized.

a very good approximation of the solution.

4.3.3 Static optimization problem

We have shown that a constant allocation decision $\bar{u}$ represents a simplified alternative to the optimal control solution, a strategy composed of bang arcs, a time-varying singular arc and the chattering artifact. In this section, we will further explore the static optimization problem by adding a second degree of freedom to the problem: the dilution rate D. In addition to that, we investigate two objectives: the production of biomass $\mathcal{V}$ and of metabolite X. The static biomass maximization problem

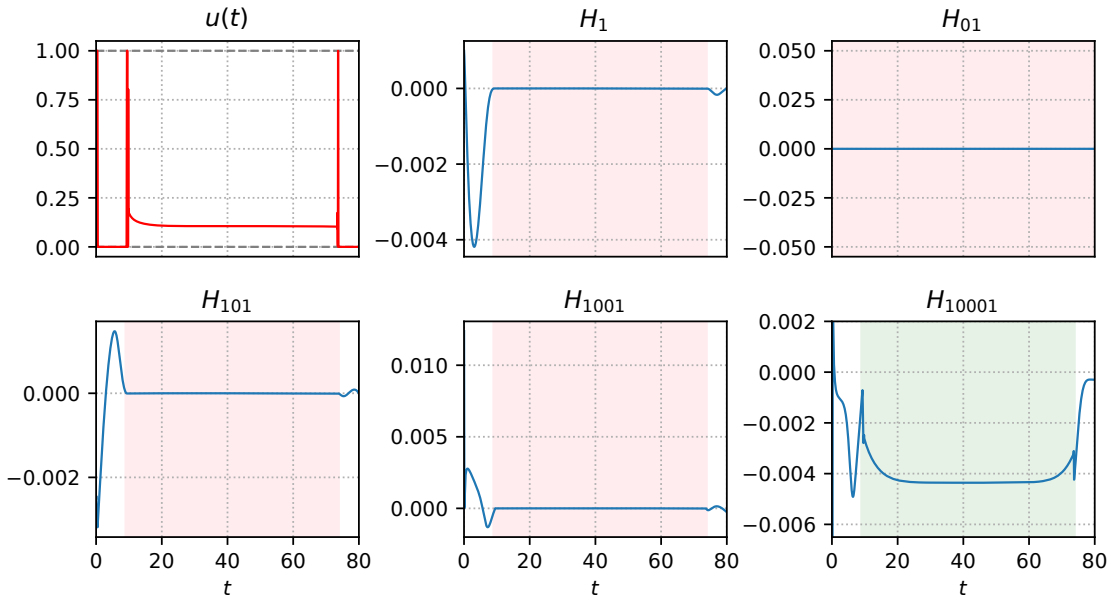


Figure 4.5: Successive derivatives of H_1 obtained with BOCOP. The intervals where the functions vanish are highlighted in light red. All functions vanish along the singular arc except for H_{10001}, highlighted in green, which is negative as required by Kelley condition (4.26).

($\mathrm{SP}_{\mathcal{V}}$) can be written as

$$\begin{cases} \textit{maximize} & J_{\mathcal{V}}(\bar{u}, D) \doteq D\mathcal{V} \\ \textit{subject to} & \textit{dynamics of } (\mathrm{S}_1), \\ & \dot{\varphi} = 0, \\ & 0 \leq \bar{u} \leq 1. \end{cases} \tag{$\mathrm{SP}_{\mathcal{V}}$}$$

Analogously, the product maximization problem can be defined as

$$\begin{cases} \textit{maximize} & J_X(\bar{u}, D) \doteq DX = D\mathcal{V}_{\mathrm{ext}} x \\ \textit{subject to} & \textit{dynamics of } (\mathrm{S}_1), \\ & \dot{\varphi} = 0, \\ & 0 \leq \bar{u} \leq 1. \end{cases} \tag{SP_X}$$

Since we look for the steady-states that maximize each objective, the washout equilibrium E_w can be excluded from the analysis since, as shown in Theorem 2, the

equilibrium corresponds to the steady-state values $\mathcal{V} = 0$ and $x = 0$. Therefore, the static problems are reduced to finding the equilibria E_i in terms of the pair $(D, \bar{u})$ that maximize each objective function. Moreover, it can be shown that the optimal solution cannot belong to the boundary of the equilibrium E_i

$$\Theta \doteq \left\{(\bar{u}, D) \in \mathbb{R}^2 \,:\, \bar{\mu}(p_w) = D\right\}.$$

Indeed, using the definitions (4.8), (4.9) and (4.10), it can be seen that on Θ there is no bacterial population, as $s_i = s_{in}$, which means that $\mathcal{V}_i = x_i = 0$. We formalize the latter reasoning in the following proposition.

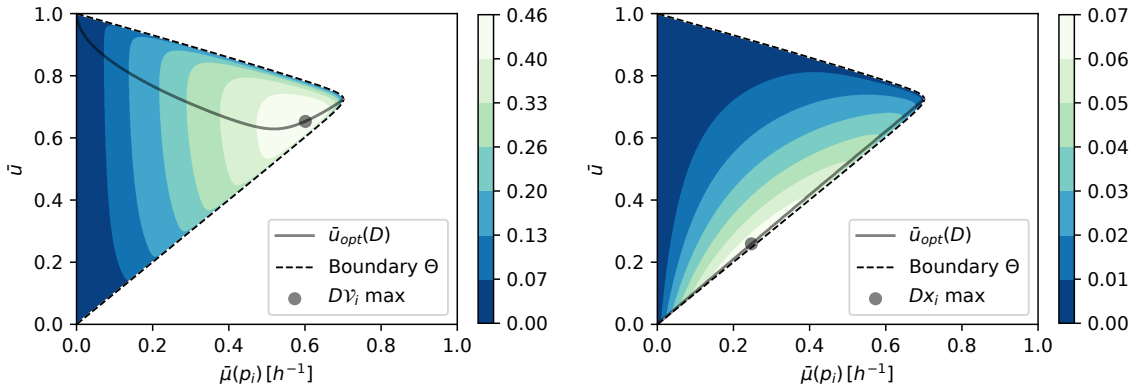
Proposition 4. *Every solution $(D, \bar{u})$ of the static optimization problems $(\mathrm{SP}_{\mathcal{V}})$ and (SP_X) is in the region of existence of E_i given by the condition $\bar{\mu}(p_w) > D$.*

When considering the same self-replicator scheme under constant environmental conditions (which could describe fed-batch cultivation) the solution for both static problems corresponds to the steady-state with maximal growth rate [3]. Proposition 4 shows that this is not the case in continuous bioreactors, where maximal growth rate is attained at the boundary set of existence of the equilibrium E_i. Two interesting particular cases within Θ are the pure static allocation strategies $\bar{u} = 0$ and $\bar{u} = 1$: A pure metabolic strategy $\bar{u} = 0$ will lead to a bacterial population with no RNA polymerase (i.e., $r = 0$), which will eventually stop the production of biomass (as $v_R(p, 0) = 0$), leading to washout in the bioreactor. Analogously, allocating all resources to the gene expression machinery will finally empty the metabolic machinery m, halting the absorption of substrate from the environment (as $v_M(p, 0) = 0$) and depleting the bacteria from precursors, which will also lead to washout. From this analysis, we can conclude that any optimal steady-state allocation $\bar{u}$ should belong to $(0, 1)$.

We recall that, in continuous bioreactors, the growth rate $\bar{\mu}(p_i)$ is fixed by the dilution rate D, as shown in equation (4.8). Figure 4.6 illustrates a numerical analysis of the static problems. It is interesting to notice in both subfigures 4.6a and 4.6b that the model accounts for the classical quasi-linear relation between the maximal growth rate (that lies on the boundary Θ), and the control $\bar{u}$ which regulates the RNA/protein mass ratio of the bacterial population [2]. The latter is a phenomenon which has been first observed experimentally, and later used to develop dynamical self-replicator models for natural and biotechnological purposes

[30, 3, 74]. However, in the present case, the growth rate is fixed through D, so it is not a result of the nutrient quality in the environment, which enables the multivariate approach.

Biomass maximization objective Results for this problem are shown in Figure 4.6a. For this case, the curve $\bar{u}_{opt}$ remains over 0.6 for all values of the growth rate, tending to $\bar{u}_{opt} = 1$ as the growth rate goes to 0. This shows that, in order to maximize $\mathcal{V}$, the allocation strategy should prioritize the synthesis of macromolecules of the gene expression machinery R over the metabolic machinery M, which catalyzes the production of biomass and not the synthesis of metabolites. Moreover, the maximum $D\mathcal{V}_i$ is attained through a fairly high growth rate ($\approx 85\%$ of the maximal growth rate).



(a) Results associated to ($\text{SP}_{\mathcal{V}}$) with $J_{\mathcal{V}} = D\mathcal{V}$.

(b) Results associated to (SP_X) with $J_X = DX$.

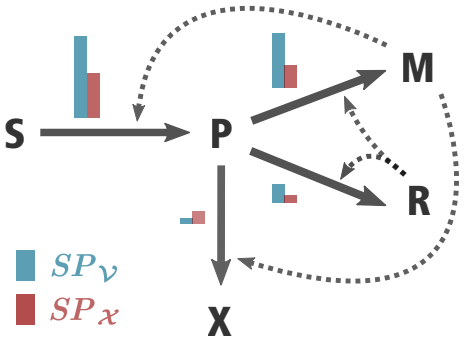
Figure 4.6: Numerical results for both static problems. The values for the objective functions $J_{\mathcal{V}}$ and J_X are represented through a qualitative colormap. The set Θ of maximum growth rate delimits the region of existence of the interior equilibrium E_i. Additionally, curves $\bar{u}_{opt}(D)$ show the optimal allocation $\bar{u}$ in terms of the dilution rate D.

Product maximization objective Results for this case are shown in Figure 4.6b. We can see that, in opposition to the first case, the optimal solution requires allocating as much precursors as possible into the metabolic machinery M, no matter the value of the dilution rate. In other words, the allocation control $\bar{u}$ should be as low as possible within the region of existence of the equilibrium E_i (but not

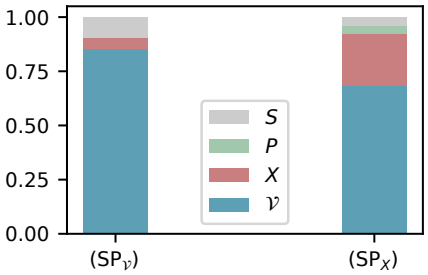
in Θ). This result is consistent with the fact that the metabolic machinery M catalyzes the synthesis of metabolite X. It is noteworthy that the optimal point DX_i is accomplished through a rather lower dilution rate D in comparison to the biomass production case, resulting in a continuous production at a rate of about 35% of the maximal growth rate. The latter might appear counter-intuitive, as it is well established in the literature that high dilution rates in continuous bioreactors imply high production rates. This characteristic can be attributed to the compromise between allocating resources to the metabolic machinery M and increasing the dilution rate, linked to the maximum value of the function $c(r)$. In other words, the lower the dilution rate of operation, the wider the interval of existence of the equilibrium $(u_{\min}, u_{\max})$, which enables the possibility of further promoting the synthesis of compounds of the metabolic machinery (by artificially lowering the allocation parameter $\bar{u}$) without going to washout. Figure 4.7a illustrates the resulting synthesis rates for each solution. In (SP_X), the synthesis rate of the metabolite v_X is increased at the expense of reducing both precursor and biomass synthesis rates v_M and v_R, which is in large part due to the reduction of the dilution rate D. This difference in the flux distribution impacts directly on the mass quantities inside the bioreactor at steady-state—depicted in Figure 4.7b—for each problem: for the (SP_X) case, there is a reduction in the biomass $\mathcal{V}$ of only 20% w.r.t. the ($\mathrm{SP}_{\mathcal{V}}$) case. However, we can see an increase on the amount of metabolite X of about 5 times that of the ($\mathrm{SP}_{\mathcal{V}}$) case. As already said, the allocation parameter $\bar{u}$ becomes quite lower in the metabolite synthesis objective. In turn, this inhibits the synthesis of macromolecules, which translates into a decrease of the bacterial population's growth rate. To compensate this effect, the steady-state pool of precursors P required for producing the metabolite X becomes considerably bigger than that of the biomass production objective (about 10 times), therefore increasing the rate of fabrication of biomass $v_R(p, r)$.

4.4 Discussion

The objective in this paper was to synthesize a certain metabolite of interest by re-adjusting the natural allocation of resources in bacteria. This is done by drawing resources from the native pathways of the cell originally used for producing biomass, and allocating them into the production of the compound X. We addressed the problem in the continuous bioreactor framework, which allows a steady-state pro-



(a) Synthesis rates magnitudes



(b) Mass quantities as percentages of $s_{in}\mathcal{V}_{\text{ext}}$ [g], knowing that $w_1\mathcal{V}_{\text{ext}} = S + P + \mathcal{V} + X = s_{in}\mathcal{V}_{\text{ext}}$ at equilibrium (Lemma 2)

Figure 4.7: Numerical results for both static problems.

duction regime. Based on previous dynamical systems approaches [30, 3, AYT5], we proposed a coarse-grained self-replicator model capable of accounting for well-studied bacterial growth laws in a simplified way, and we studied its asymptotic behavior. We tackled two different production objectives: Biomass maximization $\mathcal{V}$ and metabolite maximization X, and we compared results in order to better understand the potential control strategies required to that effect. We rely on novel bio-engineering techniques capable of delivering groundbreaking control schemes: A synthetic growth switch that allows to control the transcription of RNA polymerase through an inducible promoter. In addition to this regulation mechanism, we include the dilution rate as a control input, which yields a multi-variable optimization problem. We concluded by showing very contrasting results, but in accordance with our previous understanding of microbial resource allocation. The biomass-oriented strategy involves an almost maximal dilution rate, and prioritizes investing resources into the gene expression machinery. Conversely, producing the compound of interest requires a rather low value of the dilution rate, which allows an allocation strategy more geared towards the synthesis of components of the metabolic machinery (i.e., a lower value of $\bar{u}$). The latter shows there exist a compromise between augmenting the dilution rate and artificially boosting the production of enzymes and transporters.

The capacity to interfere with the natural bacterial behaviors at molecular level represents a promising tool to improve biotechnological processes. We are interested in further exploring the details of the implementation of such control schemes, by considering different models of the external growth switch. A next step in this regard

would involve taking into account the nature of the external signal I and its physical constraints, in order to adapt our current results towards a more implementable control loop. Additionally, our approach is based on very simplified representations of bacterial growth. Such biological processes usually involve numerous reactions and variables. In our case, we clustered all macromolecules into only 2 different classes, and we purposely omitted a number of known phenomena in bacteria, such as cell division, protein degradation and the influence of temperature. However, some of these effects have been proven to affect only marginally the results regarding the allocation problem [75]. Thus, in our case, a simplified representation becomes useful to emphasize the effects of optimally dealing with the internal resource distribution of bacteria in industrial frameworks. Eventually, such optimal strategies could provide guidance for developing online feedback solutions based on real-time measuring of the process.

Chapter 5

Model-predictive control schemes

This chapter reproduces [AYT3], accepted for publication in the 60th Conference on Decision and Control (IEEE CDC 2021).

5.1 Introduction

Microorganisms have evolved over millions of years under natural selection, submitted to a continuous optimisation process that has improved their capacity to proliferate in nature. Thus, they have developed highly optimized distribution mechanisms of their internal resources to cellular functions enabling them to face changing environments. Unraveling these internal mechanisms has always been of great interest for the scientific community, not only from a pure biological point of view, but also for biotechnological purposes. In this context, being able to understand and control the growth process is key for several industrial applications, such as in combating antibiotics resistance, food preservation, and biofuel production [7].

Considering the microbial self-replication process as a resource allocation problem is a novel approach that has successfully answered some of the underlying question in the field [30]. The latter has also motivated numerous applications to the artificial production of metabolites of interest [3, AYT5, AYT4, AYT2]. These studies aim to find how to divert the cell internal resources into a heterologous pathway in order to efficiently synthesize a specific protein. This is done through an external control that is able to disrupt the cellular allocation process of a growing culture by reengineering the transcriptional control of the expression of RNA polymerase [44]. In a dynamical systems framework, the problem can be posed as an Optimal Con-

trol Problem (OCP), which can be approached through the well-known Pontrjagin's Maximum Principle (PMP).

Model-based optimal control studies are essential in understanding the overall allocation process, as they are able to provide the gold-standard strategies, i.e. the best that can be achieved from a theoretical point of view. However, in most cases, it is impossible to obtain a closed-loop control strategy: the obtained optimal control often depends on the so-called adjoint state, which hinders its implementation (as it is the case in [3]). Additionally, such approaches depend on the accuracy of the model and the precision of its parameters, which often tend to be limited for most biochemical and biological processes. At the same time, the existing industrial applications that allow a closed-loop implementation are mainly based on general schemes such as non-linear Model Predictive Controllers (NMPC), which tend to disregard the structure of each particular problem [76].

Motivated by the lack of synergy between pure theoretical approaches and very general implementations, in this work we revisit the metabolite production problem. We summarize the open-loop optimal allocation strategies found in the literature, which are characterized by sharing the same simple structure and a common parametrisation. Then, we propose a hierarchical NMPC scheme designed on the basis of these open-loop optimal controllers. In particular, we resort to the shrinking horizon NMPC (sh-NMPC) [77], an approach targeted to control processes of known time duration, such as batch processes [78]. In contrast to typical receding horizon approaches, in the sh-NMPC, the final time of the process is fixed, and so the time window considered in the optimisation problem (i.e. the prediction horizon) shrinks at each step.

Our approach uses an optimal control-based input parametrisation that takes into account the structure of the open-loop natural allocation strategy of the cell. Thus, we first implement an MPC loop which creates a closed-loop natural allocation, followed by a second MPC that computes the external control maximising the production of a metabolite of interest. Similar approaches have been proposed to control batch and semi-batch processes through the manipulation of the feedrate [77, 79], which is a standard scheme in the bioreactor framework. The novelty of this work resides in a hierarchical control scheme that aims to affect the internal pathways of the cells in a bacterial growing culture by affecting the expression of RNA polymerase.

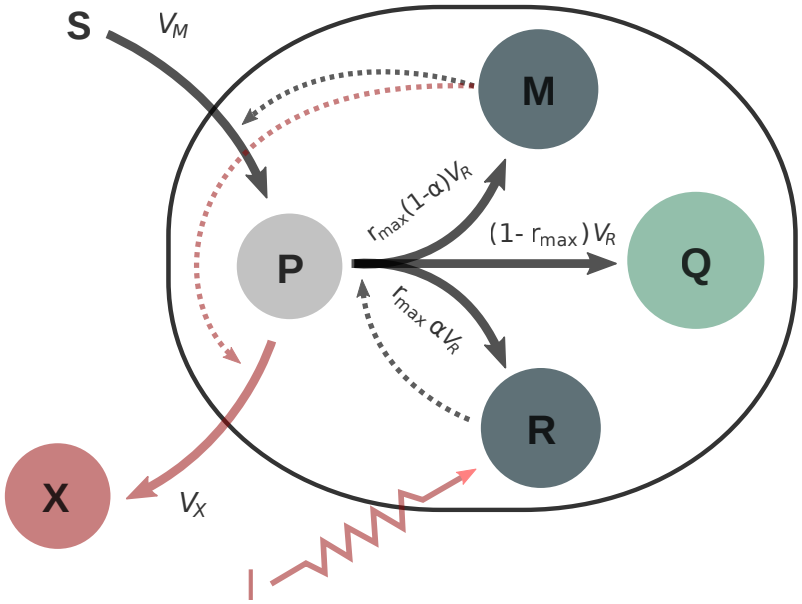


Figure 5.1: Coarse-grained self-replicator model. The external substrate S is consumed by bacteria and transformed into precursor metabolites P through the action of the metabolic machinery M. These precursors are used to produce macromolecules of the gene expression machinery R, the metabolic machinery M, the housekeeping machinery Q, and metabolites X. The external control I is able to externally affect the natural allocation parameter α in order to channel resources into the production of metabolites of interest.

We start the paper by defining the model, and the naturally-evolved resource allocation strategy used by the bacteria. Then, we propose a suboptimal parametrisation intended to emulate the open-loop strategy. In Section 5.3, we introduce the open-loop metabolite maximisation problem, and the closed-loop hierarchical scheme. Finally, we provide a numerical simulation of the approach and a comparison with the optimal case, followed by a conclusion.

5.2 Model definition

5.2.1 Self-replicator system

Based on [30], we define the self-replicator system composed of the mass (in grams) of: precursor metabolites P, the gene expression machinery R, the metabolic machinery M, the housekeeping machinery Q, and a metabolite of interest X. As illustrated

in Figure 5.1, substrate S is taken from the environment and transformed into P at rate V_M through a reaction catalyzed by M. Then, the precursors P are transformed into M, Q, R and X at rates $r_{\max}(1-\alpha)V_R$, $(1-r_{\max})V_R$, $r_{\max}\alpha V_R$, and V_X, respectively. While the reactions that produce M, Q and R are catalyzed by R, the reaction synthesizing X is catalyzed by M. In short, the ribosomal proteins R are responsible of producing new proteins, and the metabolic proteins M are responsible for the uptake of nutrients into the cell, and the production of metabolites X. The latter represents a classical trade off in synthetic biology, and is modeled through the parameter α, defined as a time function with bounds $\alpha(t) \in [0,1]$. The dynamical system is

$$\begin{cases} \dot{S} = V_{in} - V_M \\ \dot{P} = V_M - V_X - V_R, \\ \dot{R} = r_{\max}\alpha V_R, \\ \dot{M} = r_{\max}(1-\alpha)V_R, \\ \dot{Q} = (1-r_{\max})V_R, \\ \dot{X} = V_X. \end{cases}$$

where the time variable t is measured in hours. The bacterial volume is defined as $\mathcal{V} \doteq \beta(R+M+Q)$, and the growth rate given by $\mu \doteq \dot{\mathcal{V}}/\mathcal{V}$. We define the intracellular concentrations

$$p = \frac{P}{\mathcal{V}}, \quad r = \frac{R}{\mathcal{V}}, \quad m = \frac{M}{\mathcal{V}}, \quad q = \frac{Q}{\mathcal{V}}$$

and the extracellular concentration of substrate s. Using the definition of bacterial volume, we obtain the relation $\beta(r+m+q) = 1$. Then, following [16], we assume the transcription of proteins Q to be internally autoregulated to a constant value, such that

$$\beta(r+m) = r_{\max}, \quad \beta q = q_{\max} \doteq 1 - r_{\max}. \tag{5.1}$$

We define the rates of mass flow per unit volume, which we assume to be functions of the concentrations s, m and r, as $v_M(s,m) \doteq V_M/\mathcal{V}$ and $v_R(p,r) \doteq V_R/\mathcal{V}$. In this new system, the growth rate becomes $\mu = \beta V_R/\mathcal{V} = \beta v_R(p,r)$. Taking into account that a minimal concentration of ribosomes $r_{\min}$ is required in order for bacteria to self replicate, we define the kinetics of the problem as $v_M(s,m) \doteq w_M(s)\, m$,

$v_R(p,r) \doteq w_R(p)\,(r - r_{\min})$ and $v_X\,(p,m) \doteq w_X(p)\,m$, with $w_R(p) \doteq k_R\,p/(K_R + p)$, $w_M(s) \doteq k_M s/(K_S + s)$ and $w_X(p) \doteq k_X p/(K_X + p)$. We will model a production process in which the substrate remains constant. This could be the result of an external control regulating through an inflow of fresh medium to the bioreactor, or due to high availability in the medium. Thus, we replace $w_M(s) = e_M$, with $e_M > 0$ constant. We define the non-dimensional timescale $\hat{t} = k_R t$, as well as the mass fractions of the total volume $\mathcal{V}$: $\hat{p} \doteq \beta p$, $\hat{r} \doteq \beta r$, $\hat{r}_{\min} \doteq \beta r_{\min}$, $\hat{m} \doteq \beta m = r_{\max} - r$, $\hat{q} \doteq \beta q = 1 - r_{\max}$. Additionally, we define non-dimensional synthesis rates $\hat{w}_R(p) = w_R(p)/k_R$, $\hat{w}_X(p) = w_X(p)/k_R$, and parameter $E_M \doteq e_M/k_R$. Then, dropping all hats, the model becomes

$$\begin{cases} \dot{p} = E_M\,(r_{\max} - r) - w_X(p)(r_{\max} - r) \\ \qquad\quad -(p+1)w_R(p)(r - r_{\min}), \\ \dot{r} = (r_{\max}\alpha - r)w_R(p)(r - r_{\min}), \\ \dot{X} = w_X(p)(r_{\max} - r)\mathcal{V}, \\ \dot{\mathcal{V}} = w_R(p)(r - r_{\min})\mathcal{V}, \end{cases} \tag{5.2}$$

where q and m have been removed using equations (5.1). The parameter values of the kinetics and of bounds $r_{\min}$ and $r_{\max}$ are fixed based on previous studies [30, 3, 23].

5.2.2 Naturally-evolved resource allocation strategy

A common assumption in biology is that microorganisms have evolved resource allocation strategies that maximise their growth rate, which allow them to outgrow competing organisms. Such assumption can be represented by an OCP, in which the objective is to maximise the synthesis of biomass in an interval of time T given by $\Delta\mathcal{V}(T) = \mathcal{V}(T) - \mathcal{V}(0)$. This defines the cost function

$$J_N(\alpha) = \int_0^T w_R(p)(r - r_{\min})\mathcal{V}\,dt.$$

Thus, as neither the states nor the cost function depend on variable X, we will define the OCP for the reduced state $(p, r, \mathcal{V})$ with dynamics

$$\begin{cases} \dot{p} = E_M\,(r_{\max} - r) - w_X(p)(r_{\max} - r) - (p+1)w_R(p)(r - r_{\min}), \\ \dot{r} = (r_{\max}\alpha - r)w_R(p)(r - r_{\min}), \\ \dot{\mathcal{V}} = w_R(p)(r - r_{\min})\mathcal{V}, \end{cases} \tag{$\mathrm{S_N}$}$$

and initial conditions

$$p(0) = p_0, \quad r(0) = r_0 \quad \mathcal{V}(0) = \mathcal{V}_0. \tag{IC}$$

with $p_0 > 0$, $r_0 \in (r_{\min}, r_{\max})$ and $\mathcal{V}_0 > 0$. The OCP is then defined as

$$\begin{cases} \underset{\alpha}{\textit{maximise}} & \textit{biomass production } J_N(\alpha), \\ \textit{subject to} & \textit{dynamics } (\mathrm{S_N}), \\ & \textit{initial conditions } (\mathrm{IC}), \\ & \alpha(\cdot) \in \mathcal{U}, \\ & t \in [0, T]. \end{cases} \tag{$\mathrm{OCP_N}$}$$

where $\mathcal{U}$ is the set of admissible controllers, which are Lebesgue measurable real-valued functions defined on the time interval $[0, T]$ and satisfying $\alpha(t) \in [0, 1]$. In [30], the particular case where $w_X(p) = 0$, $r_{\max} = 1$ and $r_{\min} = 0$ has been studied, and similar analyses have been carried out in [3, AYT5]. By application of the PMP, it is possible to show that the optimal control that solves $\mathrm{OCP_N}$ has bang and singular arcs, where the singular arc corresponds to the solution of the static optimal control problem obtained through the addition of the constraint $(\dot{p}, \dot{r}, \dot{\mathcal{V}}) = 0$. The solution is characterized by the presence of the Fuller phenomenon before and after the constant singular arc, which produces an infinite number of bangs (also known as *chattering*), a feature that, due to obvious physical limitations, is not possible to implement (nor expected to be found in nature). An example of this kind of structure is shown in Figure 5.2.

Remark 5.2.1. *Problem OCP_N can be further simplified by considering the cost function $\ln \mathcal{V}(T)$ instead of $\mathcal{V}(T)$. Then, neither the dynamics nor the cost function depend on $\mathcal{V}$, so the problem can be rewritten in terms of the state (p, r) [30].*

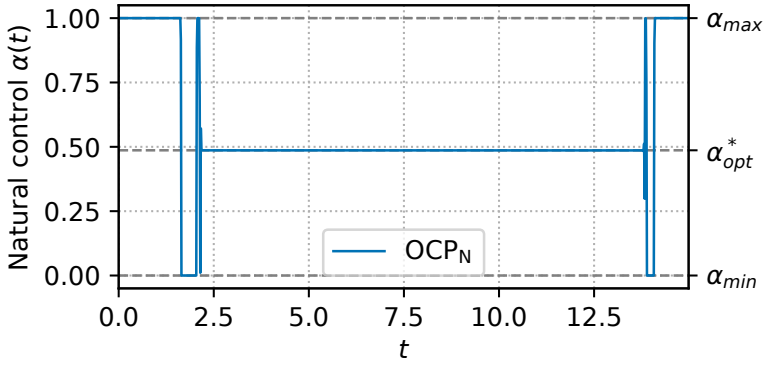


Figure 5.2: Optimal control α obtained with Bocop [4]. Simulation in a rich medium with $e_M = k_R$, meaning the medium enables the maximum growth rate. Initial conditions are $p_0 = 0.024$, $r_0 = 0.2$, and $\mathcal{V}_0 = 0.003$, and the simulation time is set to $T = 15$.

5.2.3 MPC parametrisation of the natural allocation

In order to incorporate the natural allocation strategy of the cell into the MPC loop, we propose a sub-optimal parametric form of α given by

$$\alpha_{so}(\theta, t) = \begin{cases} b_1 & \text{if } t < t_1, \\ \alpha^* & \text{if } t_1 \leq t \leq t_2, \\ b_2 & \text{if } t > t_2, \end{cases}$$

with the set of parameters $\theta \doteq (b_1, b_2, t_1, t_2, \alpha^*)$ subject to

$$\begin{aligned} b_1 \in \{0, 1\}, \qquad & b_2 \in \{0, 1\}, \\ t_f \geq t_2 \geq t_1 \geq 0, \quad & 1 \geq \alpha^* \geq 0, \end{aligned} \tag{5.3}$$

where b_1 and b_2 are Boolean parameters. The suboptimal parametric allocation α_{so} deliberately neglects the chattering artifact from the optimal control α, replacing it by pure bang controls during the intervals $[0, t_1)$ and $(t_2, T]$. In order to compare the performance of the proposed controller, we write an optimisation problem with the same biomass production objective J_N. At each time instant of the control loop, the algorithm finds the vector of parameters θ that maximises the final volume of biomass $\mathcal{V}(T)$. The latter amount to solving four optimisation problems in terms of (t_1, t_2, α^*) given by all possible combinations of Boolean parameters (b_1, b_2). At each

iteration k, the control loop starts by measuring the system and getting an estimation $(\widetilde{p}_k, \widetilde{r}_k, \widetilde{\mathcal{V}}_k)$ of the system state. Thus, the optimisation problem at iteration k is formulated with initial conditions

$$(p(k\tau), r(k\tau), \mathcal{V}(k\tau)) = (\widetilde{p}_k, \widetilde{r}_k, \widetilde{\mathcal{V}}_k). \tag{5.4}$$

This defines the optimisation problem

$$\left\{\begin{array}{ll} \underset{\theta}{\textit{maximise}} & \textit{biomass production } J_N(\theta) \\ \textit{subject to} & \textit{dynamics } (\mathrm{S_N}), \\ & \textit{initial conditions } (5.4) \\ & \alpha(\cdot) = \alpha_{so}(\theta, t), \\ & \textit{input constraints } (5.3) \\ & t \in [k\tau, T] \end{array}\right. \tag{OP^k}$$

which is solved at each instant $k\tau$. The scheme proposes a closed-loop form of the open-loop optimal control found in Subsection 5.2.2, with the purpose of implementing it in the hierarchical MPC loop.

5.2.4 Numerical example

Figure 5.3 shows a comparison of the optimal control α and the proposed suboptimal control α_{so}. The initial and final Fuller arcs are approximated by pure bang arcs (which are $\alpha = 1$ for this particular case), and the parameter α^* of the suboptimal control takes exactly the same value of the static optimal control α^*_{opt}. The difference between both control functions is minor, which translates into an imperceptible difference in the trajectories.

5.3 Artificial metabolite production

The artificial metabolite production problem is to maximise the synthesis of X over a fixed interval of time $[0, T]$, which is equal to $\Delta X(T) = X(T) - X(0)$, and can be expressed as

$$J_X(u) = \int_0^T w_X(p)(r_{\max} - r)\mathcal{V}\, dt.$$

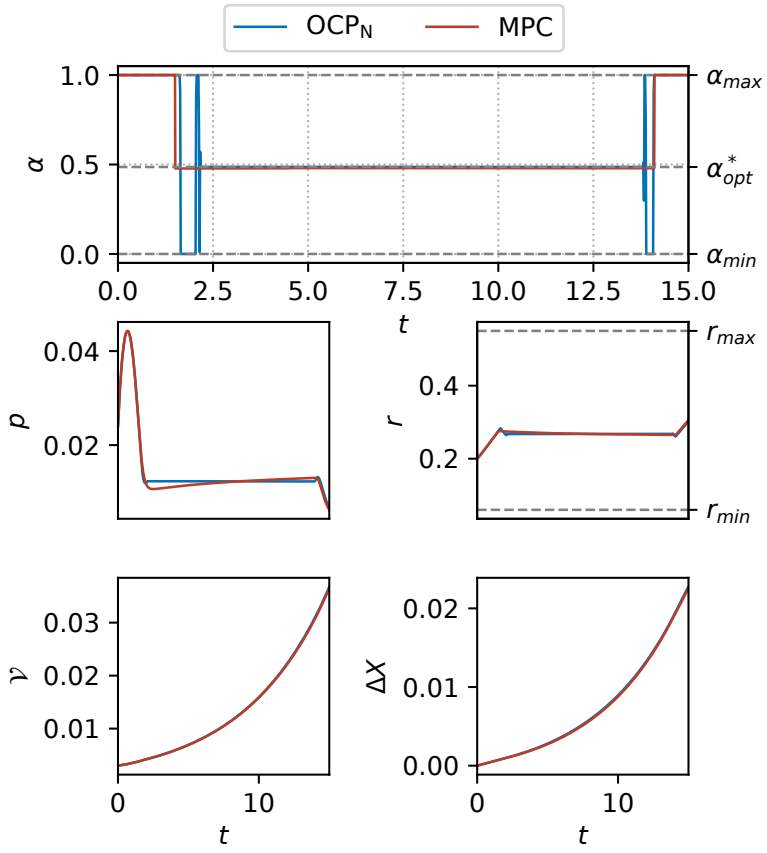


Figure 5.3: Comparison of the optimal control $\alpha(t)$ solution of $\mathrm{OCP_N}$ and the MPC scheme parametrized with the suboptimal control $\alpha_{so}(\theta, t)$. Initial conditions are set to $p_0 = 0.024$, $r_0 = 0.2$, and $\mathcal{V}_0 = 0.003$. The scheme is executed with time step $\tau = 0.3$. The quantity ΔX amounts to $X(T) - X(0)$.

As neither the states nor the cost function depend on variable X, the reduced system ($\mathrm{S_N}$) can be used.

5.3.1 Optimal Control Problem

In the original approach [3], the naturally-evolved resource allocation parameter α is overridden by the external control u, and so the dynamical equation of r becomes

$$\dot{r} = (r_{\max} u - r) w_R(p)(r - r_{\min}).$$

Then, the optimal control problem is defined as

$$\begin{cases} \underset{u}{\textit{maximise}} & \textit{metabolite production } J_X(u) \\ \textit{subject to} & \textit{dynamics } (\mathrm{S_N}) \\ & \textit{initial conditions } (\mathrm{IC}) \\ & u(\cdot) \in \mathcal{U}, \\ & t \in [0, T]. \end{cases} \tag{$\mathrm{OCP_X}$}$$

5.3.2 On the solution of the OCP

Applying PMP, we see that the Hamiltonian is affine in the control, so it has the form $H = H_0 + uH_1$, meaning that the solution is bang-singular-bang, given by

$$u(t) = \begin{cases} 0 & \textit{if } H_1 < 0, \\ 1 & \textit{if } H_1 > 0, \\ u_{sing}(t) & \textit{if } H_1 = 0. \end{cases}$$

Examples of optimal trajectories are shown in Figure 5.4 and Figure 5.5, where both problems $\mathrm{OCP_N}$ and $\mathrm{OCP_X}$ are compared for different environmental conditions representing rich and poor qualities of the nutrient in the medium. While the structures of the optimal control for both problems are similar, the optimal strategy maximizing the production of X is characterized by a non-constant singular arc, which is close to the solution u^*_{opt} of the static OCP, but deviates from it towards the end. Additionally, the times at which the junctions between bang and singular arc are produced differ, as well as the values of the bangs. In particular, in both Figures, the final bang of the natural control is $\alpha = 1$, while that of the artificial control is $u = 0$. More detailed calculations of the PMP approach can be found in [3].

5.3.3 Product maximisation including naturally-evolved allocation

The comparison between $\mathrm{OCP_N}$ and $\mathrm{OCP_X}$ proves useful to observe that the natural behavior of microbes does not necessarily match the artificial objective of producing a certain metabolite. However, the assumption made in $\mathrm{OCP_X}$ is a pure theoretical one, as α cannot be completely substituted by the external control. We then propose

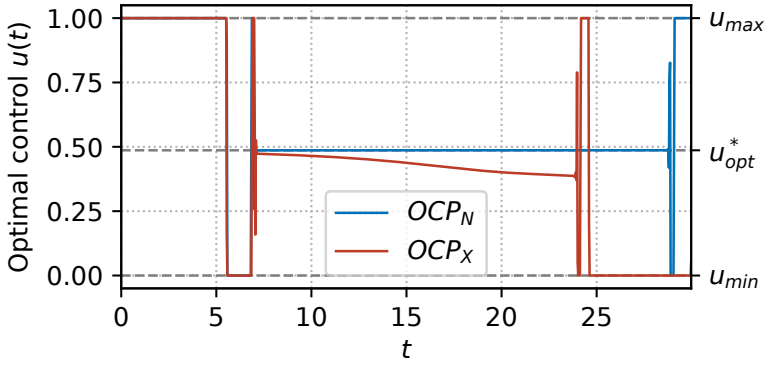


Figure 5.4: Optimal control obtained with Bocop. Simulation in a rich medium with $e_M = k_R$, meaning that the substrate enables the maximum growth rate. Initial conditions are $p_0 = 0.024$, $r_0 = 0.1$ and the simulation time is set to $T = 30$.

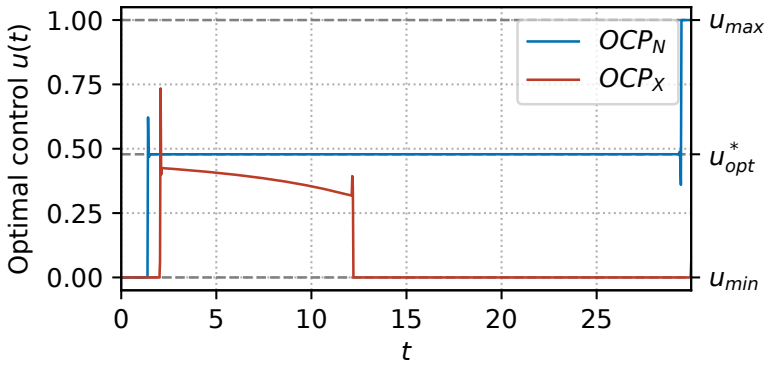


Figure 5.5: Optimal control obtained with Bocop. Simulation in a poor medium with $e_M = 0.5k_R$. Initial conditions are $p_0 = 0.024$, $r_0 = 0.3$ and the simulation time is set to $T = 30$.

an approach that takes into account two distinct processes: (i) the ability of bacteria to maximise their biomass through the optimal allocation described by the internal control problem $\mathrm{OCP_N}$, (ii) the external action of an operator intending to maximise the production of the metabolite of interest X. In practice, the new pathway associated to the dynamics of X in (5.2) is obtained by optogenetic engineering of a strain of bacteria: a light-induced control I is able to externally modulate the natural allocation process. This is modeled by replacing the control u of $\mathrm{OCP_X}$ by $u = \alpha(p, r)I$, so that the external control affects in a multiplicative way the internal allocation strategy. The crucial difference with the previous formulation [3] is that the internal

control α of the bacteria now appears in feedback form (thus depending on the two states of $\mathrm{OCP_N}$; see Remark 5.2.1), and is modulated by the external light-induced control I. It is noteworthy that no competition occur between the two objectives (biomass *vs.* metabolite production maximisation). The approach considers the internal control in feedback form, which is mitigated by an external control in relation with a global process that includes the new pathway to produce the metabolite. In this context, the proposed hierarchical approach proves to be more relevant from a biological point of view than a multicriterion one. Additionally, we assume that the feedback $\alpha(p, r)$ is known and smooth. While the latter seems to be a strong assumption from the control point of view (as the solution of $\mathrm{OCP_N}$ comprises bang and singular arcs), it is a reasonable assumption in our biological setting, where the kinetics of the involved biochemical reactions prescribe continuous behaviours (see [30] for biologically relevant approximations of the feedback). Thus, the dynamical equation of r becomes

$$\dot{r} = (r_{\max}\alpha(p, r)I - r)w_R(p)(r - r_{\min}), \tag{5.5}$$

where the new control $I(t)$ is subject to bound constraints $0 \leq I(t) \leq I_{\max}$, and the cost J_X remains unchanged. This defines problem $\overline{\mathrm{OCP}}_X$.

5.3.4 Hierarchical MPC for metabolite production

In this approach, we approximate the allocation feedback $\alpha(p, r)$ through the sh-MPC loop described in 5.2.3. At each iteration k, solving on $[k\tau, T]$ (where T is, as before, the fixed horizon) yields an approximation of $\alpha(p_k, r_k)$ (based on the suboptimal parametric form α_{so}), and of α evaluated at further steps. Then, this suboptimal feedback is injected in the dynamics (5.5) of $\overline{\mathrm{OCP}}_X$so as to find the optimal external control I maximizing $X(T)$. Thus, a second MPC is used "above" the first one. There is a quite large literature on such approaches combining several MPC loops (see, e.g., [80, 81] and references therein). Other relevant matters such as using different time grids for each MPC loop or, more generally, synchronisation issues, are not discussed here (see also the recent paper [82] on convergence of MPC methods in finite horizon). Instead, we focus on the biological application. We note, in particular, that when the feedback $\alpha(p, r)$ is zero (which would be expected for the genuine—though biologically unrealistic—feedback of $\mathrm{OCP_N}$ as zero bang arcs

can occur), the external control I is not active. In practice, when the allocation is close to zero, the MPC loop would compensate through I for this discrepancy. As the external control I is bounded, the latter can induce certain performance loss between the results of the ideal model $\mathrm{OCP_X}$ and the more realistic problem $\overline{\mathrm{OCP}}_X$. Such comparisons are provided in the next paragraph.

5.4 Numerical results

Figure 5.6 shows a comparison between the optimal trajectory solution of $\mathrm{OCP_X}$ and the hierarchical MPC proposed in this paper. Differences between both trajectories are marginal, mainly given by the approximation of the singular arc by a constant control u. In Figure 5.7 we see how, in order to match the optimal control u solution of $\mathrm{OCP_X}$, the external signal I completely arrests the natural allocation α (around $t = 12$), which implies allocating all the cellular resources to the metabolic machinery M, thus catalyzing the synthesis of X. The latter produces the suboptimal control α_{so} to increasingly compensate until it reaches the value 1. This result shows the difference between the simulated closed-loop behavior of the natural allocation strategy $\alpha_{so}(p, r)$, and the open-loop one (which remains constant almost over the whole interval $[0, T]$).

5.5 Conclusion

In this paper, we presented a hierarchical sh-NMPC approach to the problem of optimally producing a metabolite of interest in bacteria. The scheme is based on a parametric version of the naturally-evolved research allocation strategy proposed in [30], which represents a closed-loop alternative to these existing open-loop studies. A second sh-NMPC loop is applied in a hierarchical manner, in order to achieve the metabolite maximisation objective while taking into account the closed-loop natural control. Despite the approach being at an early stage, with no experimental results, it represents a step towards plausible biosynthetic real-time implementations. In future works, we are interested in comparing the natural MPC approximation with alternatives proposed in the literature, such as ppGpp regulation [30]. Other extensions include the possibility of estimating the real value of the natural allocation through online identification techniques.

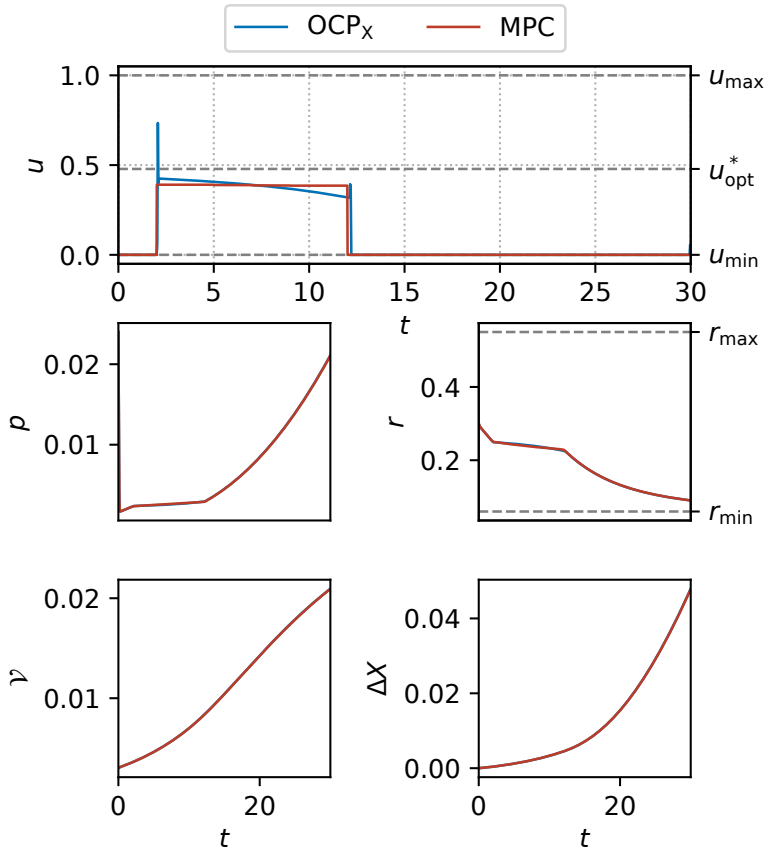


Figure 5.6: Comparison of the optimal control $u(t)$ solution of OCP_X and the hierarchical MPC scheme parametrized that considers the natural allocation as an inner MPC loop. Initial conditions are set to $p_0 = 0.024$, $r_0 = 0.3$, and $\mathcal{V}_0 = 0.003$. Final time is set to $T = 30$, the scheme is executed with time step $\tau = 1$ and the environmental constant $e_M = 0.5k_R$.

Acknowledgments

The authors thank Hidde de Jong (Inria) as well as colleagues from the ANR Maximic project for many inputs on the biological problem.

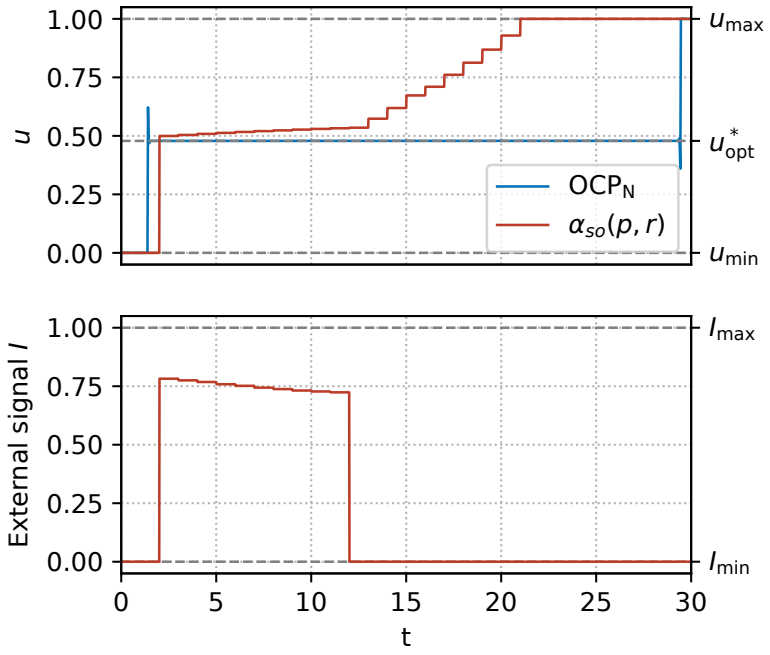


Figure 5.7: Final control u and external signal I obtained from the MPC loop simulated in Figure 5.6.

Chapter 6

A generalized resource allocation model of microbial growth

This chapter reproduces [AYT1], accepted for publication in the SIAM Journal on Applied Dynamical Systems.

6.1 Introduction

The growth of microorganisms is a paradigm example of self-replication in Nature. Microbial cells are capable of transforming nutrients from the environment into new microbial cells astonishingly fast and in a highly reproducible manner [83]. The biochemical reaction network underlying microbial growth has evolved under the pressure of natural selection, a process that has retained changes in the network structure and dynamics increasing fitness, i.e., favoring the ability of the cells to proliferate in their environment. Gaining a better comprehension of the growth of microorganisms in the context of evolution is a major scientific challenge [84], and the ability to externally control growth is critical for a wide range of applications, such as in combating antibiotics resistance, food preservation, and biofuel production [85, 86, 87].

A fruitful perspective on microbial growth is to view it as a resource allocation problem [2]. Microorganisms must assign their available resources to different cellular functions, including the uptake and conversion of nutrients into molecular building blocks of proteins and other macromolecules (metabolism), the synthesis of proteins and other macromolecules from these building blocks (gene expression),

and the detection of changes in the environment and the preparation of adequate responses (signalling and regulation). It is often assumed that microorganisms have evolved resource allocation strategies so as to maximize their growth rate, as this would allow them to outgrow competing species.

Simple mathematical models based on resource allocation principles have been surprisingly effective in accounting for experimental observations of the growth and physiology of microorganisms [2, 88, 26, 16, 30, 33, 89]. Instead of providing a detailed description of the entire biochemical reaction network, these models include a limited number of macroreactions responsible for the main growth-related functions of the cell. The models usually take the form of nonlinear ODE systems, typically 3-10 equations with parameters obtained from the experimental literature or estimated from published data. The models have been instrumental in explaining a number of steady-state relations between the growth rate and the cellular composition, in particular the concentration of ribosomes, protein complexes that are responsible for the synthesis of new proteins [26, 2, 23, 30, 90]. Moreover, they have brought out a trade-off between the rate and yield of alternative metabolic pathways that produce energy-carrying molecules, necessary for driving forward many cellular reactions, such as those involved in the synthesis of proteins and other macromolecules [26, 91, 92].

In previous work, using a three-variable resource allocation model, it was possible to predict an optimal resource allocation scheme for the response of microbial cells to a sudden nutrient change in the environment [30]. The prediction was based on the Infinite Horizon Maximum Principle, a generalization of the well-known PMP (Pontrjagin Maximum Principle) [93, 43]. A feedback control strategy inspired by a known regulatory mechanism for growth control in the bacterial cell was shown to give a quasi-optimal approximation of the optimal solution. Strategies for optimal control were also explored for an extension of the model, inspired by recent experimental work [94], which comprises a pathway for the production of a metabolite of biotechnical interest as well as an external signal allowing growth to be switched off [3, AYT5, AYT4, AYT2]. We showed by a combination of analytical and computational means that the optimal solution for the targeted metabolite production problem consists of a phase of growth maximization followed by a phase of product maximization, in agreement with strategies proposed in metabolic engineering. Optimal control approaches have also been used for studying other dynamic op-

timization problems in biology (see [95] for a review). A classical example is the determination of optimal activation patterns of metabolic pathways, such as to minimize the transition time of metabolites or minimize enzyme costs [96, 97].

The resource allocation model that lies at the basis of the above-mentioned work [30] has a number of limitations. First, the biomass of the cell was assumed to consist of two classes of proteins, enzymes catalyzing metabolic reactions and ribosomes responsible for protein synthesis, whose relative proportions vary with the growth rate. However, experimental data show that a large fraction of the total protein contents of the cell is growth rate-independent [98]. This suggests the introduction of a third protein category, dedicated mainly to basic housekeeping functions of the cell. The proportion of these proteins is independent of the growth rate and thus constrains the variations in the other two, growth rate-dependent categories [2, 23]. Second, the concentration of ribosomes and enzymes, the two protein categories included in the original model, have both a growth rate-dependent and a growth rate-independent component [2, 98]. This implies that the protein synthesis rate, and thus the growth rate, does not depend on the total ribosome concentration, as in the original model, but only on its growth rate-dependent fraction [23].

In the present manuscript, we revise the above modeling assumptions and study their impact on predicted optimal strategies for resource allocation following changes in the environment of different nature (i.e., changes in the nutrient concentration or stress responses). This leads to a number of interesting problems in mathematical analysis and control, which are addressed using tools from dynamical systems analysis and optimal control theory. A full dynamical analysis of the system shows there is a single globally attractive equilibrium, which can be related to steady-state growth conditions of bacteria observed in experiments. In spite of the simplicity of the presented model, the solutions of the associated biomass maximization problems exhibit quite interesting features. Notably, the second-order singular arc is characterized by a) the Fuller's phenomenon at its junctions, yielding an infinite set of switching points in a finite-time window, and b) the turnpike effect, which produces very particular asymptotic behaviors towards the solution of the static optimization problem. We provide a full description of the singular arc in terms of the state, as well as an explicit proof of the presence of the turnpike effect. While the predicted (optimal) control dynamics does not change much qualitatively in comparison with the previous model, the more realistic modeling assumptions offer a more general

perspective of the biological problem. For example, in contrast with the previous model where the absence of growth-rate independent protein yields a constant singular arc equal to the solution of the static optimization problem, the singular arc of the new model is not constant, but governed by a turnpike phenomenon.

In Section 2, we describe the model used in this study, followed by a global dynamical analysis of the model in Section 3. In Section 4, we calibrate the model from literature data using the equilibrium of interest for an optimal steady-state allocation parameter, and in Section 5 we formulate an optimal control problem and prove properties of the optimal solutions. In Section 6, we show that the general analysis can be applied to two different cases of environmental changes related to nutrient shifts and stress responses.

6.2 Model definition

We define a self-replicator system composed of the mass of precursor metabolites P, the gene expression machinery R (ribosomes, RNA polymerase, ...) and the metabolic machinery M (enzymes, transporters, ...), as shown in Figure 6.1. Essentially, the ribosomal proteins R are responsible for the fabrication of new proteins, and the metabolic proteins M are in charge of the uptake of nutrients for building precursor metabolites P. Following Scott *et al.* [2], we also introduce a class Q of proteins whose functions fall outside the range of tasks performed by M and R. This sector comprises mainly growth rate-independent proteins such as housekeeping proteins responsible for the maintenance of certain basic cellular functions. Needless to say, the synthesis of Q proteins draws resources away from the pathways to M and R, and consequently imposes an upper bound on the fraction of resources dedicated to self-replication and nutrient uptake. This constraint appears in the model through a constant $\gamma \in [0, 1]$, and it indicates the maximum fraction of the protein synthesis rate available for making ribosomes and metabolic enzymes. The overall allocation process can be represented by the biochemical macroreactions

$$\begin{aligned} &\mathrm{S} \xrightarrow{V_M} \mathrm{P}, \\ &\mathrm{P} \xrightarrow{V_R} \gamma\alpha\,\mathrm{R} + \gamma(1-\alpha)\,\mathrm{M} + (1-\gamma)\,\mathrm{Q}. \end{aligned}$$

The first reaction describes the transformation of external substrate S into precur-

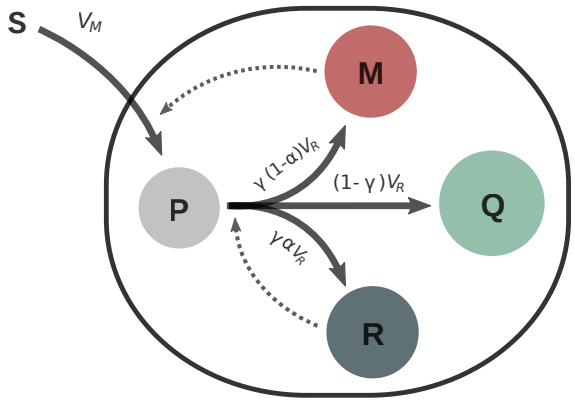


Figure 6.1: Coarse-grained self-replicator model. The external substrate S is consumed by bacteria and transformed into precursor metabolites P by the metabolic machinery M. The precursors are used to produce macromolecules of classes R, M and Q, with proportions $\gamma\alpha$, $\gamma(1-\alpha)$, and $1-\gamma$, respectively. Solid lines indicate the macroreactions with their respective synthesis rates, and dashed lines denote a catalytic effect.

sor metabolites P at a rate V_M. The second reaction represents the conversion of precursors into macromolecules R, M, and Q at a rate V_R. The roles of the enzymes M in the uptake and metabolization of nutrients and the ribosomal proteins R in the production of proteins are represented through catalytic effects, indicated with dotted arrows in Figure 6.1. In this context, protein A *catalyzes* reaction B means that the rate of reaction B is proportional to the cellular concentration of A, but the reaction itself does not consume A. The natural resource allocation strategy is modeled through the time-varying function $\alpha(t) \in [0, 1]$. Thus, the proportion of the total synthesis rate of proteins dedicated to the gene expression machinery R is $\gamma\alpha$, while that of the metabolic machinery M is $\gamma(1-\alpha)$. In particular, the allocation parameter does not influence the synthesis rate of Q, with constant proportion $1-\gamma$, as the synthesis of proteins in this class is auto-regulated through mechanisms not relevant in this study. From a biological perspective, the function $\alpha(t)$ represents the naturally-evolved allocation strategy of the cell which is, *a priori*, unknown. In the context of control theory, and throughout this paper, α is treated as the control input of the system.

6.2.1 Self-replicator system

Generalizing upon Giordano *et al.* [30], a mass balance analysis yields the dynamical system

$$\begin{cases} \dot{P} = V_M - V_R, \\ \dot{R} = \gamma\alpha V_R, \\ \dot{M} = \gamma(1-\alpha)V_R, \\ \dot{Q} = (1-\gamma)V_R. \end{cases}$$

where mass quantities P, M, R and Q are described in grams (g), the synthesis rates V_M and V_R in grams per hour, and α is the dimensionless allocation parameter. In what follows, we will assume that the proteins of classes R, M and Q are responsible for most of the bacterial mass [83], and so we define the bacterial volume $\mathcal{V}$ measured in liter units (L) as

$$\mathcal{V} = \beta(R + M + Q), \tag{6.1}$$

where β corresponds to a density constant relating mass and bacterial volume [99], such that the total biomass in grams is given by $\mathcal{V}/\beta$. The above assumption implies that the mass of precursor metabolites represents a negligible fraction of the total biomass, (in other words, $P \ll \mathcal{V}/\beta$). We define the intracellular concentrations in grams per liter as

$$p_{\mathcal{V}} \doteq \frac{P}{\mathcal{V}}, \quad r_{\mathcal{V}} \doteq \frac{R}{\mathcal{V}}, \quad m_{\mathcal{V}} \doteq \frac{M}{\mathcal{V}}, \quad q_{\mathcal{V}} \doteq \frac{Q}{\mathcal{V}}. \tag{6.2}$$

Using (6.1) and (6.2), we to obtain the relation

$$r_{\mathcal{V}} + m_{\mathcal{V}} + q_{\mathcal{V}} = \frac{1}{\beta}. \tag{6.3}$$

We also define the rates of mass flow per unit volume, which we assume to be functions of the available concentrations, as

$$v_M(s, m_{\mathcal{V}}) \doteq \frac{V_M}{\mathcal{V}}, \quad v_R(p_{\mathcal{V}}, r_{\mathcal{V}}) \doteq \frac{V_R}{\mathcal{V}},$$

where s corresponds to the extracellular concentration of substrate measured in grams per liter. The growth rate of the bacterial population is defined as the relative change of the bacterial volume

$$\mu \doteq \frac{\dot{\mathcal{V}}}{\mathcal{V}} = \frac{\beta V_R}{\mathcal{V}} = \beta v_R(p_\mathcal{V}, r_\mathcal{V}).$$

We write the system in terms of the concentrations as

$$\begin{cases} \dot{p}_\mathcal{V} = v_M(s, m_\mathcal{V}) - (1 + \beta p_\mathcal{V}) v_R(p_\mathcal{V}, r_\mathcal{V}), \\ \dot{r}_\mathcal{V} = (\gamma\alpha - \beta r_\mathcal{V}) v_R(p_\mathcal{V}, r_\mathcal{V}), \\ \dot{m}_\mathcal{V} = (\gamma(1-\alpha) - \beta m_\mathcal{V}) v_R(p_\mathcal{V}, r_\mathcal{V}), \\ \dot{q}_\mathcal{V} = ((1-\gamma) - \beta q_\mathcal{V}) v_R(p_\mathcal{V}, r_\mathcal{V}), \\ \dot{\mathcal{V}} = \beta v_R(p_\mathcal{V}, r_\mathcal{V}) \mathcal{V}. \end{cases}$$

6.2.2 Kinetic definition

We define the kinetics of the reaction system by taking into account that a minimal concentration of ribosomal proteins $r_{\mathcal{V},\min} \in (0, \gamma/\beta)$ is required for protein synthesis to take place. In other words, a part of the bacterial volume is occupied by ribosomal proteins which do not directly contribute to growth [23]. Such behavior can be modeled as

$$v_R(p_\mathcal{V}, r_\mathcal{V}) \doteq w_R(p_\mathcal{V}) \, (r_\mathcal{V} - r_{\mathcal{V},\min})^+, \quad \text{with } (r_\mathcal{V} - r_{\min})^+ = \begin{cases} r_\mathcal{V} - r_{\mathcal{V},\min} & \text{if } r_\mathcal{V} \geq r_{\mathcal{V},\min}, \\ 0 & \text{if } r_\mathcal{V} < r_{\mathcal{V},\min}. \end{cases}$$

Later on, we will see that there is no need to define $v_R(p_\mathcal{V}, r_\mathcal{V})$ for $r_\mathcal{V} < r_{\mathcal{V},\min}$ if the initial conditions lie in a particular region of the state space. The rate of nutrient uptake is defined as

$$v_M(s, m_\mathcal{V}) \doteq w_M(s) \, m_\mathcal{V}.$$

We will make the following assumption for functions $w_R(p_\mathcal{V})$ and $w_M(s)$.

Hypothesis 6.2.1. *Function $w_i(x) : \mathbb{R}_+ \to \mathbb{R}_+$ is*

- *Continuously differentiable w.r.t. x,*

- *Null at the origin:* $w_i(0) = 0$,
- *Strictly increasing:* $w_i'(x) > 0, \forall x \geq 0$,
- *Strictly concave:* $w_i''(x) < 0, \forall x \geq 0$,
- *Upper bounded:* $\lim_{x\to\infty} w_i(x) = k_i > 0$.

The classical Michaelis-Menten kinetics satisfies Hypothesis 6.2.1. While most of the mathematical results are based on this general definition, for the calibration of the model and numerical simulations, we will resort to the particular case where the functions are defined as

$$w_R(p_\mathcal{V}) \doteq k_R \frac{p_\mathcal{V}}{K_R + p_\mathcal{V}}, \quad w_M(s) \doteq k_M \frac{s}{K_S + s}, \tag{6.4}$$

where k_R and k_M are the maximal reaction rates in h^{-1}, and K_M and K_R are the half-saturation constants of the synthesis rates in g L^{-1}. For the general case introduced in Hypothesis 6.2.1 we will define

$$k_R \doteq \lim_{p_\mathcal{V}\to\infty} w_R(p_\mathcal{V}).$$

6.2.3 Constant environmental conditions

We assume that the availability of the substrate in the medium is constant over the time-window analyzed. The latter can be modeled by setting s constant, and thus removing the dynamics of s from the system.

Hypothesis 6.2.2. *The flow of substrate can be expressed as* $w_M(s) = e_M$ *with* $e_M > 0$ *constant.*

Using this assumption, the dynamical equation of $p_\mathcal{V}$ becomes

$$\dot{p}_\mathcal{V} = e_M m_\mathcal{V} - (1 + \beta p_\mathcal{V}) w_R(p_\mathcal{V})(r_\mathcal{V} - r_{\mathcal{V},\min})^+.$$

The constant e_M models the substrate availability of the medium, but it is also related to the quality of the nutrient and the efficiency of the macroreaction that produces precursor metabolites.

6.2.4 Mass fraction formulation and non-dimensionalization

We define mass fractions of the total bacterial mass as

$$p \doteq \beta p_{\mathcal{V}}, \quad r \doteq \beta r_{\mathcal{V}}, \quad r_{\min} \doteq \beta r_{\mathcal{V},\min}, \quad m \doteq \beta m_{\mathcal{V}}, \quad q \doteq \beta q_{\mathcal{V}},$$

which, replacing in (6.3), yields the relation

$$r + m + q = 1. \tag{6.5}$$

We also define the non-dimensional time variable $\hat{t} \doteq k_R t$, and the non-dimensional growth rate

$$\hat{\mu}(p, r) \doteq \frac{\mu(p_{\mathcal{V}}, r_{\mathcal{V}})}{k_R} = \hat{w}_R(p)(r - r_{\min}), \tag{6.6}$$

with $\hat{w}_R(p) : \mathbb{R}_+ \to [0, 1)$ defined as $\hat{w}_R(p) \doteq w_R(p_{\mathcal{V}})/k_R$, and $E_M \doteq e_M/k_R$. For the sake of simplicity, let us drop all hats from the current notation. Then, the model becomes

$$\begin{cases} \dot{p} = E_M\, m - (p+1) w_R(p)(r - r_{\min})^+, \\ \dot{r} = (\gamma\alpha - r) w_R(p)(r - r_{\min})^+, \\ \dot{m} = (\gamma(1-\alpha) - m) w_R(p)(r - r_{\min})^+, \\ \dot{\mathcal{V}} = w_R(p)(r - r_{\min})^+ \mathcal{V}, \\ m + r \leq 1, \end{cases} \tag{S}$$

where q has been removed since it can be expressed in terms of the other concentrations through (6.5); and the constraint $m + r \leq 1$ is required to comply with $q \geq 0$. The model differs from that of Giordano *et al.* by the addition of the category of housekeeping proteins (q) and a minimum concentration of ribosomes for protein synthesis ($r_{\min}$). In what follows, we will systematically investigate how these differences affect the asymptotic behavior and optimal resource allocation strategies.

6.3 Asymptotic behavior

In the present section, we will study the asymptotic behavior of the reduced system representing the intracellular dynamics

$$\begin{cases} \dot{p} = E_M\, m - (p+1) w_R(p)(r - r_{\min})^+, \\ \dot{r} = (\gamma\alpha - r) w_R(p)(r - r_{\min})^+, \\ \dot{m} = (\gamma(1-\alpha) - m) w_R(p)(r - r_{\min})^+, \\ m + r \leq 1, \end{cases} \tag{6.7}$$

where $\mathcal{V}$ has been removed since none of the remaining states explicitly depends on it, and it only reaches a steady state when there is no bacterial growth (otherwise, $\dot{\mathcal{V}} > 0$). We will start by stating the invariant set of interest.

Lemma 6.3.1. *The set*

$$\Gamma = \left\{(p, r, m) \in \mathbb{R}^3 \,:\, p \geq 0,\, \gamma \geq r \geq r_{\min},\, \gamma \geq m \geq 0,\, m + r \leq 1\right\}$$

is positively invariant by (6.7).

Proof. This can be easily verified by evaluating the differential equations of system (6.7) over the boundaries of Γ. As for the condition $m + r \leq 1$, we can define a variable $z \doteq m + r$ that obeys the dynamics

$$\dot{z} = (\gamma - z) w_R(p)(r - r_{\min})^+$$

which, when evaluated at $z = 1$ yields $\dot{z} \leq 0$, as $r_{\max} < 1$, which proves its invariance. □

This Lemma states that $\gamma \geq r \geq r_{\min}$ for any trajectory with initial conditions in Γ. As a consequence, there is no need to define the flow $v_R(p, r)$ for values of r under $r_{\min}$. The same thing can be said for the constraint $m + r \leq 1$, which is valid for every trajectory starting in Γ. Additionally, since γ represents the maximal ribosomal mass fraction, we will define the following parameter.

Definition 6.3.2. *The maximal ribosomal mass fraction is* $r_{\max} \doteq \gamma$.

Then, we will reduce the study of the system to this set and so, using Definition 6.3.2, we redefine (6.7) as

$$\begin{cases} \dot{p} = E_M\, m - (p+1) w_R(p)(r - r_{\min}) \\ \dot{r} = (r_{\max}\alpha - r) w_R(p)(r - r_{\min}) \\ \dot{m} = (r_{\max}(1-\alpha) - m) w_R(p)(r - r_{\min}) \end{cases} \tag{S'}$$

where $(r - r_{\min})^+$ has been replaced by $r - r_{\min}$, and the constraint $m + r \leq 1$ has been removed. Furthermore, we will define the minimum constant allocation parameter $\alpha^*_{\min}$ necessary to allow steady-state self-replication, given by

$$\alpha^*_{\min} \doteq \frac{r_{\min}}{r_{\max}}.$$

Its importance will be analyzed throughout the current section.

6.3.1 Local stability

Theorem 6.3.1. *System* (S') *has the equilibria*

- $E_1 \doteq (p^*, r^*, m^*)$*, locally stable if* $\alpha^* > \alpha^*_{\min}$.
- $E_2 \doteq (p, r_{\min}, 0)$*, locally unstable if* $\alpha^* > \alpha^*_{\min}$.
- $E_3 \doteq (0, r, 0)$*, locally unstable if* $r \neq r_{\min}$.

with

$$\begin{aligned} p^* &\doteq \left\{ p \in \mathbb{R}_+ \,:\, (p+1) w_R(p) = \frac{E_M m^*}{r^* - r_{\min}} \right\}, \\ r^* &\doteq r_{\max}\alpha^*, \\ m^* &\doteq r_{\max}(1 - \alpha^*). \end{aligned} \tag{6.8}$$

Proof. The general Jacobian matrix of the system (S') is

$$\begin{bmatrix} -\Big(w_R(p)+(p+1)w_R'(p)\Big)(r-r_{\min}) & -(p+1)w_R(p) & E_M \\ (r_{\max}\alpha - r)w_R'(p)(r-r_{\min}) & (r_{\max}\alpha - 2r + r_{\min})w_R(p) & 0 \\ \Big(r_{\max}(1-\alpha)-m\Big)w_R'(p)(r-r_{\min}) & \Big(r_{\max}(1-\alpha)-m\Big)w_R(p) & -w_R(p)(r-r_{\min}) \end{bmatrix}. \tag{6.9}$$

We first see that, if $\alpha^* > \alpha^*_{\min}$, the value p^* is unique since $(p+1)w_R(p)$ is a monotone increasing function satisfying $w_R(0) = 0$ and $\lim_{p\to\infty}(p+1)w_R(p) = \infty$ (as stated in Hypothesis 6.2.1), and $E_M m^*/(r^* - r_{\min}) > 0$, so the set (6.8) yields a unique solution. For $\alpha^* < \alpha^*_{\min}$, the equation for p^* in (6.8) has no valid solution as $E_M m^*/(r^* - r_{\min})$ becomes negative, and therefore the equilibrium does not exist. The Jacobian (6.9) for E_1 becomes

$$J_1 = \begin{bmatrix} -\Big(w_R(p^*)+(p^*+1)w_R'(p^*)\Big)(r^*-r_{\min}) & -(p^*+1)w_R(p^*) & E_M \\ 0 & -(r^*-r_{\min})w_R(p^*) & 0 \\ 0 & 0 & -w_R(p^*)(r^*-r_{\min}) \end{bmatrix}$$

and so the local stability of the equilibrium is given by the signs of the roots of the characteristic polynomial, which are $\lambda = -\Big(w_R(p^*)+(p^*+1)w_R'(p^*)\Big)(r^*-r_{\min})$, $\lambda = -(p^*+1)w_R(p^*)$, and $\lambda = -w_R(p^*)(r^*-r_{\min})$. As the three roots are negative, we conclude that, if the equilibrium exists, it is locally stable. For the second equilibrium E_2, the Jacobian is

$$J_2 = \begin{bmatrix} 0 & -w_R(p) & E_M \\ 0 & (r^*-r_{\min})w_R(p) & 0 \\ 0 & r_{\max}(1-\alpha^*)w_R(p) & 0 \end{bmatrix}$$

with characteristic polynomial

$$P_2(\lambda) = \lambda^2\Big(\lambda - (r^*-r_{\min})w_R(p)\Big).$$

If $\alpha^* > \alpha^*_{\min}$, then J_2 has one positive eigenvalue and E_2 becomes locally unstable.

As for E_3, the Jacobian is

$$J_3 = \begin{bmatrix} -w_R'(0)(r - r_{\min}) & 0 & E_M \\ (r_{\max}\alpha - r)w_R'(0)(r - r_{\min}) & 0 & 0 \\ r_{\max}(1-\alpha)w_R'(0)(r - r_{\min}) & 0 & 0 \end{bmatrix}$$

with characteristic polynomial

$$P_3(\lambda) = \lambda^2\Big(\lambda + w_R'(0)(r - r_{\min})\Big) - E_M r_{\max}(1-\alpha)w_R'(0)(r - r_{\min})\lambda.$$

One root is $\lambda = 0$, and the two remaining roots can be found by solving the equation

$$\lambda^2 + \lambda w_R'(0)(r - r_{\min}) - E_M r_{\max}(1-\alpha)w_R'(0)(r - r_{\min}) = 0.$$

By the Routh-Hurwitz criterion, the two remaining roots are in the open left half plane if and only if $w_R'(0)(r - r_{\min}) > 0$ and $E_M r_{\max}(1-\alpha)w_R'(0)(r - r_{\min}) < 0$, which is never true. Consequently, for $r \neq r_{\min}$, there is at least one positive root, and so the equilibrium is unstable. □

6.3.2 Global behavior

We will study the global behavior of system (S') for the initial conditions

$$p(0) > 0, \quad r(0) \in (r_{\min}, r_{\max}), \quad m(0) \in (0, r_{\max}), \quad r(0) + m(0) \leq 1. \qquad \text{(IC)}$$

and for a given constant allocation parameter

$$\alpha(t) = \alpha^* \in (\alpha^*_{\min}, 1)\,.$$

Under this constraint, we see that the dynamics of r and m become

$$\dot{r} = (r^* - r)w_R(p)(r - r_{\min}), \quad \dot{m} = (m^* - m)w_R(p)(r - r_{\min}),$$

which means that, if $p > 0$ and $r > r_{\min}$, the signs of $\dot{r}$ and $\dot{m}$ are given by the signs of $r^* - r$ and $m^* - m$, respectively (and both $\dot{r}$ and $\dot{m}$ are zero if $p = 0$ or $r = r_{\min}$).

Then, let us divide Γ into the subsets

$$\mathcal{R}^- \doteq \{(p,r,m) \in \Gamma \,:\, r \in (r_{\min}, r^*)\}, \quad \mathcal{M}^- \doteq \{(p,r,m) \in \Gamma \,:\, m \in (0, m^*)\},$$
$$\mathcal{R}^+ \doteq \{(p,r,m) \in \Gamma \,:\, r \in (r^*, r_{\max})\}, \quad \mathcal{M}^+ \doteq \{(p,r,m) \in \Gamma \,:\, m \in (m^*, r_{\max})\},$$

such that $\Gamma = \overline{\mathcal{R}}^- \cup \overline{\mathcal{R}}^+ = \overline{\mathcal{M}}^- \cup \overline{\mathcal{M}}^+$. In these sets, the following holds.

Lemma 6.3.3. *For $\alpha(t) = \alpha^* \in (\alpha^*_{\min}, 1)$, the closed sets $\overline{\mathcal{R}}^-$, $\overline{\mathcal{R}}^+$, $\overline{\mathcal{M}}^-$ and $\overline{\mathcal{M}}^+$ are invariant by* (S'), *and*

$$\begin{cases} \dot{r} \geq 0 & \text{if } (p,r,m) \in \mathcal{R}^-, \\ \dot{r} \leq 0 & \text{if } (p,r,m) \in \mathcal{R}^+, \end{cases} \qquad \begin{cases} \dot{m} \geq 0 & \text{if } (p,r,m) \in \mathcal{M}^-, \\ \dot{m} \leq 0 & \text{if } (p,r,m) \in \mathcal{M}^+. \end{cases}$$

Again, the invariance of the sets can be checked by evaluating the vector field over the boundaries of the sets. Now we state a first result.

Proposition 6.3.4. *For $\alpha(t) = \alpha^* \in (\alpha^*_{\min}, 1)$ and initial conditions* (IC), *system* (S') *has a lower bound*

$$(p,r,m) \geq (p_{\text{low}}, r_{\text{low}}, m_{\text{low}}) \text{ for all } t \geq 0,$$

with

$$\begin{aligned} r_{\text{low}} &\doteq \min(r(0), r^*), \quad m_{\text{low}} \doteq \min(m(0), m^*), \\ p_{\text{low}} &\doteq \left\{ p \in \mathbb{R}_+ \,:\, (p+1) w_R(p) = \frac{E_M m_{\text{low}}}{r_{\max} - r_{\min}} \right\}. \end{aligned} \tag{6.10}$$

Proof. For a trajectory emanating from $\mathcal{R}^-$ (respectively, $\mathcal{R}^+$), it follows that $\dot{r} \geq 0$ (respectively, $\dot{r} \leq 0$) for all t (according to Lemma 6.3.3), and so $r \geq r(0)$ (respectively, $r \geq r^*$) for all t. This proves that $r \geq \min(r(0), r^*) > r_{\min}$ for all t (depending on whether the trajectory starts in $\mathcal{R}^-$ or $\mathcal{R}^+$). Similarly, a trajectory starting in $\mathcal{M}^-$ (respectively, $\mathcal{M}^+$) meets $\dot{m} \geq 0$ (respectively, $\dot{m} \leq 0$) for all t, and so $m \geq m(0)$ (respectively, $m \geq m^*$) for all t. Then, it follows that $m \geq \min(m(0), m^*)$ for all $t \geq 0$. The equation for p can thus be lower-bounded to

$$\dot{p} \geq E_M m_{\text{low}} - (p+1) w_R(p)(r_{\max} - r_{\min}),$$

which means $p \geq p_{\text{low}}$ for all $t \geq 0$, with p_{low} the solution of (6.10), which is unique

by the same arguments as those used in Theorem 6.3.1. □

A lower bound on system (S') is a stronger condition than the classical persistence for biological populations, as the bound is imposed not only for $t \to \infty$ but for the whole trajectory. As a consequence, the growth rate never vanishes, as it meets $\mu(p, r) \geq w_R(p_{\text{low}})(r_{\text{low}} - r_{\min}) > 0$ for all $t \geq 0$. Then, the global stability of the system is straightforward.

Theorem 6.3.2. *For $\alpha(t) = \alpha^* \in (\alpha^*_{\min}, 1)$ and initial conditions* (IC)*, every solution of* (S') *converges to the equilibrium E_1.*

Proof. Since $p \geq p_{\text{low}} > 0$ and $r \geq r_{\text{low}} > r_{\min}$ for all $t \geq 0$, we have that $\text{sign}(\dot{r}) = \text{sign}(r^* - r)$ and $\text{sign}(\dot{m}) = \text{sign}(m^* - m)$, showing that r and m converge asymptotically to r^* and m^*, respectively. Consequently, the dynamical equation of p becomes $\dot{p} = E_M m^* - (p+1)w_R(p)(r^* - r_{\min})$ and so $\text{sign}(\dot{p}) = \text{sign}(p^* - p)$, which means that p converges asymptotically to the steady-state value p^*. □

Remark 6.3.5. *For the case over the invariant plane given by $r(0) = r_{\min}$ and $m(0) > 0$, concentrations m and r are constant along the whole trajectory, and p increases linearly with time (as $\dot{p} = E_M m(0)$). This is a degenerate case that contradicts the assumption $p \ll 1$, and lacks biological relevance.*

6.3.3 Maximum steady-state growth rate

A classical hypothesis in the literature is to suppose bacterial populations in steady-state regimes maximize their growth rate ([30] and references therein). We are interested in finding the static allocation strategy α^* that produces this situation. Since the only equilibrium that admits bacterial growth is E_1, we will express the static optimization problem as

$$\max_{\alpha^* \in [\alpha^*_{\min}, 1]} \mu(p^*, r^*),$$

which can be rewritten as $\mu(p^*, r^*) = w_R(p^*)(r^* - r_{\min})$. It is possible to express α^* in terms of p^* through the relation

$$\alpha^*(p^*) = \frac{E_M + (p^* + 1)w_R(p^*)\alpha^*_{\min}}{E_M + (p^* + 1)w_R(p^*)}. \tag{6.11}$$

Moreover, since the above function $\alpha^*(p^*) : \mathbb{R}_+ \to (\alpha^*_{\min}, 1]$ is monotone decreasing, it is possible to write the optimization problem in terms of p^* instead of α^*. The growth rate in terms of p^* writes

$$w_R(p^*)(r^* - r_{\min}) = (r_{\max} - r_{\min}) \left(\frac{E_M w_R(p^*)}{E_M + (p^* + 1) w_R(p^*)} \right). \tag{6.12}$$

We differentiate w.r.t. p^* and we get the relation $w_R(p^*)^2 = E_M w'_R(p^*)$, which has a unique solution since, according to Hypothesis 6.2.1, $w_R(p)^2$ is a monotone increasing function satisfying $w_R^2(0) = 0$ and $\lim_{p\to\infty} w_R^2(p) = 1$, and $w'_R(p)$ is a monotone decreasing function satisfying $w'_R(0) > 0$ and $\lim_{p\to\infty} w'_R(p) = 0$ (as $w_R(p)$ is a strictly increasing upper-bounded function). Then, the condition for optimality can be expressed as

$$\frac{w_R(p^*_{\text{opt}})^2}{E_M w'_R(p^*_{\text{opt}})} = 1. \tag{6.13}$$

Thus, the optimal allocation parameter α^* is obtained by replacing p^*_{opt} in (6.11), and the maximal static growth rate can be calculated using (6.12). From (6.13), it can be seen that p^*_{opt} depends neither on $r_{\min}$ nor on $r_{\max}$, suggesting that the steady-state precursor concentration is independent of the housekeeping protein fraction q and of the growth rate-independent ribosomal fraction. Conversely, the precursor concentration is rather determined by the environmental conditions and by the nature of the function $w_R(p)$. It can be proven that the latter result is not a consequence of assumption (6.1): when considering a definition of the bacterial volume as $\beta(P + R + M + Q)$, which takes into account the mass P, the optimal precursor concentration amounts to $p^*_{\text{opt}}/(1 + p^*_{\text{opt}})$.

In addition, from $\dot{p} = 0$ in (S'), we get:

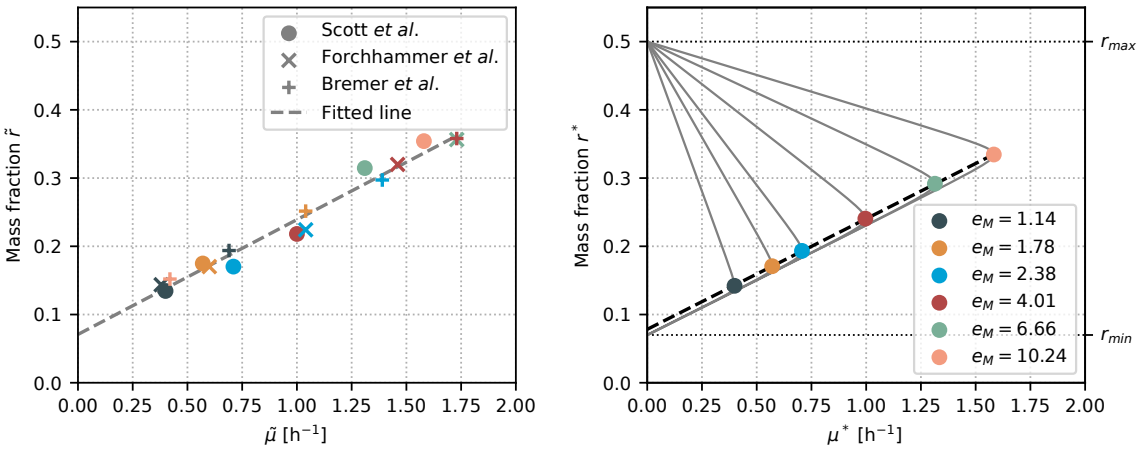
$$\frac{r^* - r_{\min}}{m^*} = \frac{E_M}{(p^* + 1) w_R(p^*)}.$$

This shows that, for the optimal steady state, the concentration ratio of the active gene expression machinery over the metabolic machinery does not depend on $r_{\max}$ either. Thus, a cellular strategy regulating the precursor concentration and the balance between gene expression and metabolism could lead to the optimal equilibrium, regardless of the demand for Q.

6.4 Model calibration

Whereas the parameter values do not affect the results above and the optimal control analysis in the next section, they are nevertheless important for simulations illustrating the dynamics and optimal allocation strategies of system (S'). Below, we derive such parameters for the model bacterium *Escherichia coli*, using published sources. The β constant used in the definition of the bacterial volume (6.1) corresponds to the inverse of the protein density, which is set to 0.003 [L g^{-1}] based on [30]. According to [2], the ribosomal fraction of the proteome[1] can vary between 6% and 55%. In more recent studies [98], this sector is divided into growth-rate dependent and independent fractions. The maximal growth-rate dependent ribosomal fraction of the proteome is estimated to be 41%, and the growth rate-independent fraction is 9%. Based on these experimental estimations, we set $r_{\max} = 0.41 + 0.09 = 0.5$. We performed further calibrations using data sets from [5, 6, 2, 98] containing measurements of various strains of *E. coli* growing in different media. The data sets are composed essentially of data points (*growth rate*, *RNA/protein mass ratio*) measured at steady state. Most RNA is ribosomal RNA found overwhelmingly in ribosomes, the main constituent of the gene expression machinery. In order to adjust the measurements to model (S'), the observed RNA/protein ratios can be converted to mass fractions r through multiplication with a conversion factor $\rho = 0.76$ μg of protein/μg of RNA [2]. As a result, we have n measurements of form $(\widetilde{\mu}_k, \widetilde{r}_k)$ which are assumed to follow a linear relation [2], as seen in Figure 6.2a. From the vertical intercept of the linear regression performed using the data points, we obtain $r_{\min} = 0.07$, in agreement with previous studies [2, 23, 98]. Each data point, composed of an observed growth rate and its associated ribosomal mass fraction, can be related to an optimal steady state of system (S') for a certain environmental condition e_M. Thus, each kth pair $(\widetilde{\mu}_k, \widetilde{r}_k)$ of the n measurements should yield a constant environmental condition $e_{M,k}$, and all pairs should simultaneously adjust the rate constant k_R. Such fitting can be done by resorting to the Michaelis-Menten kinetic form introduced in (6.4). Based on [30], we fix the half-saturation constant of protein synthesis $K_R = 1$ g L^{-1}. We then define the parameter vector $\theta = (k_R, e_{M,1}, \ldots, e_{M,n})$ which is computed by solving a least-squares regression problem. Using the relation (6.6), the

[1] The proteome is the total amount of protein in the cell



(a) Data set (b) Steady-state growth rate

Figure 6.2: Experimental data from [2, 5, 6] plotted in (a) shows a linearity of $r^2 = 0.9739$ (dashed line, fitted to data) with a vertical intersect $r_{\min} = 0.07$ and slope $k_R = 6.23$ h^{-1}. In (b), steady-state growth rate curves μ^* are shown in terms of the mass fraction $r^* \in (r_{\min}, r_{\max})$ for different fitted values of e_M. Each optimal pair $(\mu^*_{\text{opt}}, r^*_{\text{opt}})$ marked with color circles corresponds to a sample from the data set of Scott *et al.* denoted in (a) with circles of matching colors.

cost function to minimize is

$$\min_{\theta \in \mathbb{R}_+^{n+1}} \sum_{k=1}^{n} (\widetilde{\mu}_k - \mu^*_{\text{opt}}(k_R, e_{M,k}))^2 + (\widetilde{r}_k - r^*_{\text{opt}}(k_R, e_{M,k}))^2,$$

where the non-dimensional growth rate μ^*_{opt} is calculated using (6.12), and the optimal steady state $(r^*_{\text{opt}}, p^*_{\text{opt}}, m^*_{\text{opt}})$ is expressed in terms of α^*_{opt} (using Theorem 6.3.1) which is, at the same time, a function of k_R and $e_{M,k}$. The numerical solution yields $k_R = 6.23$ h^{-1}, and different values of e_M matching different nutrients from the dataset (see Figure 6.2b). We can validate these results by computing the maximal growth rate $k_R(r_{\max} - r_{\min}) = 2.68$ h^{-1} based on the adjusted parameters, which is a value that corresponds well with literature values of the maximal growth rate of *E. coli* in rich media [6].

6.5 Optimal resource allocation

6.5.1 Problem definition

In this section we formulate the dynamic optimization problem under the hypothesis that microbial populations have evolved resource allocation strategies enabling them to maximize their biomass [13, 100]. This is represented by an OCP (Optimal Control Problem) where the objective is to maximize the final volume at time T given by $\mathcal{V}(T)$. For the sake of convenience, we propose to maximize the quantity $\log \mathcal{V}(T)$ (since log is an increasing function) given by

$$\log \mathcal{V}(T) = \int_0^T \mu(p, r)\, dt + \log \mathcal{V}(0).$$

As the initial condition $\mathcal{V}(0)$ is fixed, we define the cost function

$$J(u) \doteq \int_0^T \mu(p, r)\, dt = \int_0^T w_R(p)(r - r_{\min})\, dt.$$

Since $\mathcal{V}$ appears neither in the dynamics nor in the cost function, the optimal problem will be written considering the reduced system introduced in (S') with initial conditions given by (IC). We write the optimal control problem

$$\begin{cases} \textit{maximize} & J(u) = \displaystyle\int_0^T w_R(p)(r - r_{\min})\, dt \\ \textit{subject to} & \textit{dynamics} \text{ (S')}, \\ & \textit{initial conditions} \text{ (IC)}, \\ & \alpha(\cdot) \in \mathcal{U}, \end{cases} \tag{OCP}$$

with $\mathcal{U}$ the set of admissible controllers, which are Lebesgue measurable real-valued functions defined on the interval $[0, T]$ and satisfying the constraint $\alpha(t) \in [0, 1]$. Problem (OCP) has neither final state constraints nor path constraints. In the context of dynamical optimization, the use of path constraints can be useful to restrict the solutions to those meeting certain physical and biological limitations, especially when dealing with more complex models. While enforcing additional constraints on the OCP increases the dimension of the problem, standard optimal control solvers are able to handle such formulations. In this work, imposing initial

conditions (IC) guarantees that every trajectory of the system stays within the set Γ defined in Lemma 6.3.1, which ensures that the solutions are consistent with the biological assumptions. In principle, this formulation of the problem resembles the optimal control problem proposed in [30]: the objective is to maximize the accumulation of a certain quantity within the system during a fixed time interval $[0, T]$. The main difference lies in the dynamics of the system, as the introduction of the protein Q increases the system dimension by one, which yields a more relevant (and more complex) associated OCP. We will see in following sections that the problem raised in this work can be solved by a generalization of Giordano *et al.*'s approach.

6.5.2 Pontrjagin Maximum Principle

Existence of a solution for this class of OCPs is rather trivial. Given that there are no terminal constraints, there is no controllability issue. Moreover, the dynamics is affine in the control with the latter included in a compact and convex set (a closed interval), and one can easily check that every finite-time trajectory remains bounded. So existence is guaranteed by Filippov's theorem [53]. Then, for an optimal control problem (OCP) with state $\varphi \in \mathbb{R}^n$, Pontrjagin maximum principle (PMP) ensures that there exist $\lambda^0 \leq 0$ and a piecewise absolutely continuous mapping $\lambda(\cdot) : [0, T] \to \mathbb{R}^n$, with $(\lambda(\cdot), \lambda^0) \neq (0, 0)$, such that the extremal $(\varphi, \lambda, \lambda^0, \alpha)$ satisfies the generalized Hamiltonian system

$$\begin{cases} \dot{\varphi} = \frac{\partial}{\partial \lambda} H(\varphi, \lambda, \lambda^0, \alpha), \\ \dot{\lambda} = -\frac{\partial}{\partial \varphi} H(\varphi, \lambda, \lambda^0, \alpha), \\ H(\varphi, \lambda, \lambda^0, \alpha) = \max_{\alpha \in [0,1]} H(\varphi, \lambda, \lambda^0, \alpha), \end{cases} \tag{PMP}$$

for almost every $t \in [0, T]$. For our particular case, we have the state vector $\varphi \doteq (p, r, m)$ and adjoint vector $\lambda \doteq (\lambda_p, \lambda_r, \lambda_m)$ and the Hamiltonian given by

$$H(\varphi, \lambda, \lambda^0, \alpha) = \lambda^0 w_R(p)(r - r_{\min}) + \langle \lambda, F(\varphi, u) \rangle, \tag{6.14}$$

where F represents the right-hand side of system (S'). Given that in (OCP) there is no terminal condition on the state $\varphi(T)$, the transversality condition for the adjoint

state is $\lambda(T) = 0$, and we can discard abnormal extremals from the analysis. In other words, any extremal $(\varphi, \lambda, \lambda^0, \alpha)$ satisfying PMP is normal, so $\lambda^0 \neq 0$. Developing (6.14) yields the Hamiltonian

$$\begin{aligned} H =& \Big(E_M\, m - (p+1)w_R(p)(r - r_{\min})\Big)\lambda_p + (r_{\max}\alpha - r)w_R(p)(r - r_{\min})\lambda_r \\ &+ (r_{\max}(1-\alpha) - m)w_R(p)(r - r_{\min})\lambda_m - \lambda^0 w_R(p)(r - r_{\min}), \end{aligned}$$

and the adjoint system is

$$\left\{ \begin{aligned} \dot\lambda_p =\ & w_R(p)(r - r_{\min})\lambda_p + (p+1)w_R'(p)(r - r_{\min})\lambda_p - (r_{\max}\alpha - r)w_R'(p)(r - r_{\min})\lambda_r \\ & -(r_{\max}(1-\alpha) - m)w_R'(p)(r - r_{\min})\lambda_m + \lambda^0 w_R'(p)(r - r_{\min}), \\ \dot\lambda_r =\ & (p+1)w_R(p)\lambda_p + w_R(p)(r - r_{\min})\lambda_r - (r_{\max}\alpha - r)w_R(p)\lambda_r \\ & -(r_{\max}(1-\alpha) - m)w_R(p)\lambda_m + \lambda^0 w_R(p), \\ \dot\lambda_m =\ & -E_M\lambda_p + w_R(p)(r - r_{\min})\lambda_m. \end{aligned} \right. \tag{6.15}$$

Since the Hamiltonian is linear in the control α, we rewrite it in the input-affine form $H = H_0 + \alpha H_1$ with

$$\begin{aligned} H_0 &= \Big(E_M\, m - (p+1)w_R(p)(r - r_{\min})\Big)\lambda_p - r w_R(p)(r - r_{\min})\lambda_r \\ &\quad + \Big(r_{\max} - m\Big)w_R(p)(r - r_{\min})\lambda_m - \lambda^0 w_R(p)(r - r_{\min}), \\ H_1 &= r_{\max} w_R(p)(r - r_{\min})(\lambda_r - \lambda_m). \end{aligned} \tag{6.16}$$

The constrained optimal control α should maximize the Hamiltonian, so the solution is

$$\alpha(t) = \left\{ \begin{array}{cc} 0 & \textit{if } H_1 < 0, \\ 1 & \textit{if } H_1 > 0, \\ \alpha_{\text{sing}}(t) & \textit{if } H_1 = 0, \end{array} \right.$$

where $\alpha_{\text{sing}}(t)$ is called a singular control, showing that any optimal control is a concatenation of bangs ($\alpha = \pm 1$) and singular arcs, depending on the sign of the switching function H_1. As obtained in [AYT5, AYT2], a bang arc $\alpha = 0$ (respec-

tively $\alpha = 1$) corresponds to a pure allocation strategy where the production of R (respectively M) is completely switched off. While a full description of the optimal control is often difficult to obtain through PMP, there are certain analyses that can be performed to help understand its structure. We will first see that the final bang of the optimal control is an upper bang $\alpha = 1$.

Lemma 6.5.1. *There exists ϵ such that the optimal control solution of* (OCP) *is $\alpha(t) = 1$ for the interval of time $[T - \epsilon, T]$.*

Proof. We define $\lambda_z = \lambda_r - \lambda_m$, where its dynamics can be obtained from (6.15). It can be seen that, when evaluating its dynamics at final time, we get

$$\dot{\lambda}_z(T) = \lambda^0 w_R(p(T)) < 0,$$

due to the whole adjoint state being null at final time except for λ^0. As $\lambda_z(T)$ also vanishes due to the transversality conditions, we have $\lambda_z(T - \epsilon) > 0$ for a certain ϵ. Then, $H_1 > 0$ for the interval $[T - \epsilon, T]$, which corresponds to a bang arc $\alpha = 1$. □

A control $\alpha = 1$ implies a strategy in which all resources are allocated to ribosome synthesis, thus favoring the synthesis of proteins. An intuitive interpretation of Lemma 6.5.1 is that, when approaching the final time T, the most efficient strategy is to exploit as much as possible the available precursors. This is achieved by maximizing the proteins catalyzing v_R, at the expense of arresting the uptake of nutrients v_M from the environment. In order to further describe the optimal control, we can analyze the singular extremals. A singular arc occurs when the switching function H_1 vanishes over a subinterval of time. A detailed description of the singular arcs can be done by differentiating succesively the switching function H_1 until the singular control α_{sing} can be obtained as a function of the state φ and the adjoint state λ.

6.5.3 Study of the singular arcs

Introduction

We assume H_1 vanishes on a whole sub-interval $[t_1, t_2] \subset [0, T]$, so the extremal belongs to the singular surface

$$\Sigma \doteq \left\{(\varphi, \lambda) \in \mathbb{R}^6 \,:\, H_1(\varphi, \lambda) = 0\right\}.$$

Since H_1 vanishes identically, so does its derivative with respect to time. Differentiating along an extremal (φ, λ) amounts to taking a Poisson bracket[2] with the Hamiltonian H [53]. Indeed, along the singular arc,

$$0 = \dot{H}_1 = \frac{\partial H_1}{\partial \varphi}\dot{\varphi} + \frac{\partial H_1}{\partial \lambda}\dot{\lambda} = \sum_{i=1}^{n}\left(\frac{\partial H}{\partial \lambda_i}\frac{\partial H_1}{\partial \varphi_i} - \frac{\partial H}{\partial \varphi_i}\frac{\partial H_1}{\partial \lambda_i}\right) = \{H, H_1\} = \{H_0, H_1\}.$$

The first derivative $\dot{H}_1 = H_{01} \doteq \{H_0, H_1\}$ is equal to $\langle \lambda, F_{01} \rangle$, where F_{01} corresponds to the Lie bracket of the vector fields F_0 and F_1. Differentiating again we obtain

$$0 = \dot{H}_{01} = H_{001} + \alpha H_{101}.$$

Again, $H_{001} \doteq \langle \lambda, F_{001} \rangle$ where, with the same notation as before, F_{001} is the Lie bracket of F_0 with F_{01}. If, on the set

$$\Sigma' \doteq \left\{(\varphi, \lambda) \in \mathbb{R}^6 \,:\, H_1(\varphi, \lambda) = H_{01}(\varphi, \lambda) = 0\right\},$$

the bracket H_{101} is also zero, the control disappears from the previous equality, and one has to differentiate at least two more times to retrieve the control: H_{0001} is also zero, and

$$0 = H_{00001} + \alpha H_{10001}. \tag{6.17}$$

If the length-five bracket H_{10001} is not zero, the singular arc is of *order two*. When H_{101} vanishes not only on Σ' but on all $\mathbb{R}^6$, the order is said to be *intrinsic* and connections between bang and singular arcs can only occur through an infinite number of switchings [69], the so-called Fuller phenomenon. Otherwise, the order is termed *local*, and Fuller phenomenon may or may not occur. Using (6.17), the singular control u_s is obtained as a function of both the state φ and the adjoint state λ as

$$\alpha_s(\varphi, \lambda) \doteq -\frac{H_{00001}}{H_{10001}}.$$

In our low-dimensional situation, there exists the possibility that the singular control

[2]The Poisson bracket $\{f, g\}$ of two functions f and g along an extremal (φ, λ) is defined as

$$\{f, g\} = \sum_{i=1}^{n}\left(\frac{\partial f}{\partial \lambda_i}\frac{\partial g}{\partial \varphi_i} - \frac{\partial f}{\partial \varphi_i}\frac{\partial g}{\partial \lambda_i}\right).$$

is in feedback form, that is, as a function of the state only. The latter can be verified by rewriting the system in dimension four (Mayer optimal control formulation where the final volume is maximized), in terms of $\widetilde{\varphi} \doteq (p, r, m, \mathcal{V})$ and its adjoint $\widetilde{\lambda} \doteq (\lambda_p, \lambda_r, \lambda_m, \lambda_{\mathcal{V}})$. The dynamics is affine in the control,

$$\dot{\widetilde{\varphi}} = \widetilde{F}_0(\widetilde{\varphi}) + \alpha \widetilde{F}_1(\widetilde{\varphi}),$$

and so is the Hamiltonian:

$$\widetilde{H}(\widetilde{\varphi}, \widetilde{\lambda}, \alpha) = \widetilde{H}_0 + \alpha \widetilde{H}_1,$$

with $\widetilde{H}_i = \langle \widetilde{\lambda}, \widetilde{F}_i \rangle$, $i = 0, 1$. The same computation as before leads to the following relations along a singular arc of order two:

$$0 = \dot{\widetilde{H}}_1 = \dot{\widetilde{H}}_{01} = \dot{\widetilde{H}}_{001} = \dot{\widetilde{H}}_{0001},$$

and

$$0 = \widetilde{H}_{00001} + \alpha \widetilde{H}_{10001}.$$

Proposition 6.5.2. *Assume that, for all φ, $\widetilde{F}_1$, $\widetilde{F}_{01}$, and $\widetilde{F}_{001}$ are independent. Then, an order two singular control depends only on the state $\widetilde{\varphi}$, and can be expressed as*

$$\alpha_s(\widetilde{\varphi}) = -\frac{\det\left(\widetilde{F}_1, \widetilde{F}_{01}, \widetilde{F}_{001}, \widetilde{F}_{00001}\right)}{\det\left(\widetilde{F}_1, \widetilde{F}_{01}, \widetilde{F}_{001}, \widetilde{F}_{10001}\right)}.$$

Proof. The previous relations imply that $\widetilde{\lambda}$ is orthogonal to $\widetilde{F}_1$, $\widetilde{F}_{01}$, $\widetilde{F}_{001}$, and also to $\widetilde{F}_{00001} + \alpha \widetilde{F}_{10001}$. If these four vector fields were independent at some point along the singular arc, $\widetilde{\lambda} \in \mathbb{R}^4$ would vanish: for a problem in Mayer form, this would contradict the maximum principle. So their determinant must vanish everywhere along the arc and

$$\det\left(\widetilde{F}_1, \widetilde{F}_{01}, \widetilde{F}_{001}, \widetilde{F}_{00001}\right) + \alpha \det\left(\widetilde{F}_1, \widetilde{F}_{01}, \widetilde{F}_{001}, \widetilde{F}_{10001}\right) = 0.$$

If the second determinant was zero, given the rank assumption on the first three vector fields, F_{10001} would belong to their span; but this is impossible since it would imply $H_{10001} = 0$, contradicting the fact that the singular is of order two. □

Going back to the three-dimensional formulation, one can explicit the computations by successively differentiating the expression (6.16).

Singular arc in feedback form

The condition $H_1 = 0$ could be a consequence of the growth rate $w_R(p)(r - r_{\min})$ vanishing over the whole interval $[t_1, t_2]$. We will see this is not possible given the dynamics of the system.

Proposition 6.5.3. *The growth rate $\mu(p, r) = w_R(p)(r - r_{\min})$ cannot vanish along the optimal solution of* (OCP).

Proof. For any trajectory of (S') with initial conditions (IC), control $\alpha(\cdot) \in \mathcal{U}$ and $t \in [0, T]$, we have $\dot{p} \leq E_M r_{\max}$, which means $p \leq p^T_{\max} \doteq E_M r_{\max} T + p(0)$. Then, $\dot{r} \geq -r_{\max} w_R(p^T_{\max})(r - r_{\min})$. Additionally, since $w_R(p)$ is continuously differentiable, there exists c such that $cp \geq w_R(p)$, which means that $\dot{p} \geq -cp(p^T_{\max}+1)(r_{\max}-r_{\min})$. Then, at worst, the state p (respectively, r) decays exponentially towards the value 0 (respectively, $r_{\min}$), which cannot be attained in finite time. □

As a consequence of Proposition 6.5.3, the condition $H_1 = 0$ becomes

$$\lambda_r - \lambda_m = 0. \qquad \text{(Condition 1)}$$

We define the quantity $\phi(\varphi, \lambda) \doteq (r_{\max} - m - r)\lambda_r - (p+1)\lambda_p - \lambda^0$, so that the time derivative of (Condition 1) is

$$\phi(\varphi, \lambda) w_R(p) - E_M \lambda_p = 0. \qquad \text{(Condition 2)}$$

Along a singular arc, the Hamiltonian can be rewritten as

$$H = E_M m \lambda_p + \phi(\varphi, \lambda) w_R(p)(r - r_{\min}), \tag{6.18}$$

and, using (Condition 1) and (Condition 2), the adjoint system becomes

$$\begin{cases} \dfrac{d\lambda_p}{dt} = w_R(p)(r - r_{\min})\lambda_p - \phi(\varphi, \lambda) w'_R(p)(r - r_{\min}), \\ \dfrac{d\lambda_r}{dt} = w_R(p)(r - r_{\min})\lambda_r - \phi(\varphi, \lambda) w_R(p). \end{cases}$$

Proposition 6.5.4. *Neither $\phi(\varphi,\lambda)$ nor λ_p can vanish along a singular arc.*

Proof. According to (Condition 2), if either $\phi(\varphi,\lambda)$ or λ_p are null, then both of them are null. Then, if $\phi(\varphi,\lambda)=\lambda_p=0$, equation (6.18) would imply that the Hamiltonian vanishes in Σ, and therefore it would vanish for the whole interval $[0,T]$ (as it is constant along the solution). However, one can see in (6.14) that the Hamiltonian evaluated at final time is $-\lambda^0 w_R(p(T))(r(T)-r_{\min})$ which cannot be 0 due to Proposition 6.5.3 and $\lambda^0\neq 0$. □

We differentiate (Condition 2) w.r.t. time and we get $\dot{\phi}(\varphi,\lambda)w_R(p)+\phi(\varphi,\lambda)w_R'(p)\dot{p}-E_M\dot{\lambda}_p=0$. Replacing the latter and using Proposition 6.5.4 allows us to reduce the expression to

$$-(r_{\max}-r_{\min})w_R(p)^2+E_M(m+r-r_{\min})w_R'(p)=0, \qquad \text{(Condition 3)}$$

which allows us to express $m+r$ in terms of p.

Lemma 6.5.5. *Along a singular arc over the interval $[t_1,t_2]$,*

$$m+r=x(p)$$

with $x(p):\mathbb{R}_+\to[r_{\min},\infty)$ defined as

$$x(p)\doteq(r_{\max}-r_{\min})\frac{w_R(p)^2}{E_M w_R'(p)}+r_{\min},$$

which, using (6.13)*, yields $x(p^*_{\text{opt}})=r_{\max}$.*

The fact that the control does not show up in (Condition 3)—which is obtained by differentiating (Condition 1) twice—means that the singular arc is *at least* of order two. We differentiate (Condition 3) and we get

$$\Big(r_{\max}-x(p)+(p+1)x'(p)\Big)w_R(p)(r-r_{\min})-E_M m x'(p)=0. \qquad \text{(Condition 4)}$$

We define the function

$$y(p)\doteq w_R(p)\Big(r_{\max}-x(p)+(p+1)x'(p)\Big). \tag{6.19}$$

Using (Condition 3) and (6.19) in (Condition 4) yields

$$(x(p) - r_{\min})y(p) - \Big(E_M x'(p) + y(p)\Big)m = 0,$$

which means we can express m and r in terms of p along the singular arc.

Lemma 6.5.6. *Along a singular arc over the interval* $[t_1, t_2]$,

$$m = (x(p) - r_{\min})\frac{y(p)}{E_M x'(p) + y(p)}, \tag{6.20}$$

$$r = x(p) - (x(p) - r_{\min})\frac{y(p)}{E_M x'(p) + y(p)}. \tag{6.21}$$

We differentiate (Condition 4) and we get

$$\begin{aligned} &-(r_{\max}(1-\alpha) - m)w_R(p)(r - r_{\min}) + x'(p)\frac{y(p)}{E_M x'(p) + y(p)}\dot{p} \\ &+(x(p) - r_{\min})\left(\frac{y'(p)}{E_M x'(p) + y(p)} - \frac{y(p)}{(E_M x'(p) + y(p))^2}(E_M x''(p) + y'(p))\right)\dot{p} = 0, \end{aligned} \tag{6.22}$$

meaning that we can express

$$\alpha_{\text{sing}}(p) = 1 - \frac{m}{r_{\max}}\left(\left(\frac{x'(p)}{x(p) - r_{\min}} + \frac{y'(p)}{y'(p)} - \frac{E_M x''(p) + y'(p)}{E_M x'(p) + y(p)}\right)\frac{\dot{p}}{w_R(p)(r - r_{\min})} + 1\right).$$

While (Condition 3) showed that the order of the singular arc is *at least* two, the latter relation proves that it is *exactly* two. Indeed, the coefficient before α in (6.22) is $-r_{\max}w_R(p)(r - r_{\min})$, which cannot vanish as proven in Proposition 6.5.3. The singular arc is said to be *locally of order two*, as the coefficient of α in (Condition 3) is zero along the singular arc, but not everywhere on the cotangent bundle [69]. In this case, the presence of the Fuller phenomenon (i.e., the junctions between bang and singular arcs constituting an infinite number of switchings) is not guaranteed. However, this turns out to be the case as it will be shown in the numerical computations. Besides, in accordance with Proposition 6.5.2, the order two singular control can be expressed in feedback form, i.e., as a function of the state only. We performed a numerical rank test using Singular Value Decomposition, which confirmed that the rank condition is fulfilled. More precisely, the actual computation proves that the singular control can be expressed as a function of p only (Lemma 6.5.6 entails that

r, m and therefore $\dot{p}$ can be expressed in terms of p), which allows to retrieve the turnpike behaviour as described in the following section.

The turnpike phenomenon

Using (6.20) and (6.21), we see that, along a singular arc, the dynamical equation of p becomes

$$\dot{p} = E_M w_R(p) \frac{x(p) - r_{\min}}{E_M x'(p) + y(p)} \Big(r_{\max} - x(p) \Big),$$

which is only equal to 0 when $r_{\max} = x(p)$. This is only true at $p = p^*_{\text{opt}}$, and so

$$\text{sign}\,(\dot{p}) = \text{sign}\,\left(p^*_{\text{opt}} - p\right),$$

meaning that, in a singular arc over the interval $[t_1, t_2]$, the concentration p converges asymptotically to the optimal value p^*_{opt}. This means that m and r would also converge to the optimal values m^*_{opt} and r^*_{opt}, respectively, and the singular control α_{sing} to α^*_{opt}. We formalize this in the following theorem.

Theorem 6.5.1. *On a singular arc, the system states and singular control tend asymptotically to*

$$\begin{aligned}(p, r, m) &= (p^*_{\text{opt}}, r^*_{\text{opt}}, m^*_{\text{opt}}),\\ \alpha_{\text{sing}}(t) &= \alpha^*_{\text{opt}}.\end{aligned}$$

The above theorem is an explicit proof of the presence of the turnpike property: an optimal control characterized by a singular arc that stays exponentially close to the steady-state solution of the static optimal control problem [71]. This phenomenon has been considerably studied in econometry [101], and more recently in biology [102, 30, 3]. It has been shown that, for large final times, the trajectory of the system spends most of the time near the optimal steady state, and that in infinite horizon problems, it converges to this state.

6.5.4 Numerical results

The computations of the optimal trajectories were performed with Bocop [4], which solves the optimal control problem through a direct method. An online version of

the numerical computations can be visualized and executed on the gallery of the `ct` (Control Toolbox) project[3]. The time discretization algorithm used is Lobato IIIC (implicit, 4-stage, order 6) with 2000 time steps. Figure 6.3 shows an optimal trajectory with $r(0) + m(0) < r_{\max}$, where most of the bacterial mass corresponds to class Q proteins. The obtained optimal control confirms the conclusions of the latter section: a large part of the time, the optimal control remains near the optimal steady-state allocation α^*_{opt}, according to the turnpike theory (Theorem 6.5.1). The solution presents chattering after and before the singular arc, as expected in the presence of Fuller's phenomenon (even if only a finite number of bangs is computed by the numerical method), and the final bang corresponds to $\alpha = 1$ (Lemma 6.5.1). In order to verify the optimality of the singular arc, we performed a numerical

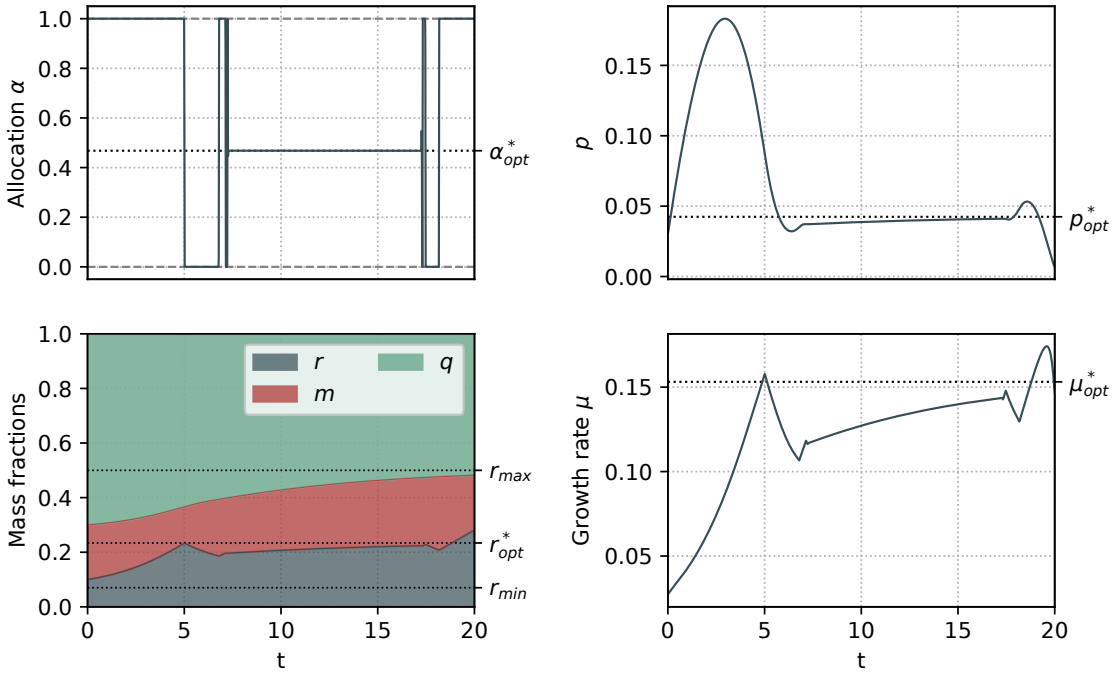


Figure 6.3: Numerical simulation of (OCP) obtained with Bocop, for the parameter values derived in Section 4. Initial state is $p(0) = 0.03$, $r(0) = 0.1$, $m(0) = 0.2$ with $E_M = 0.6$. As predicted, the optimal control α involves chattering after and before the singular arc. The mass fraction q converges to $1 - r_{\max}$ and $m + r$ to $r_{\max}$. Moreover, along the singular arc, the states (p^*, r^*, m^*) converge asymptotically to $(p^*_{\text{opt}}, r^*_{\text{opt}}, m^*_{\text{opt}})$.

[3] `https://ct.gitlabpages.inria.fr/gallery/bacteria/bacteria.html`

computation of the derivatives of H_1, which is shown in Figure 6.4. The fact that the factor of α in the fourth derivative is different from 0 confirms that the singular arc is of order 2. Moreover, its negativity complies with the *generalized Legendre-Clebsch* condition given by

$$(-1)^k \frac{\partial}{\partial \alpha} \left(\frac{d^{2k}}{dt^{2k}} H_1 \right) < 0, \tag{6.23}$$

along the singular arc, which is a necessary condition for optimality. As we state in [AYT2], even if there exist no available sufficient condition to verify local optimality of extremals with Fuller arcs, a check of the Legendre-Clebsch condition along the singular arc can ensure that the extremal obtained is not a too crude local minimizer. For the second-order singular arc case, the condition corresponds to the case $k = 2$. The initial conditions used in Figure 6.3 were only chosen to confirm the theoretical results found throughout this section, by emphasizing the main features of the solution. However, from a biological perspective, a situation where $r + m$ is significantly different from its steady-state value $r_{\max}$ is not to be expected: a common assumption in these classes of coarse-grained models is that the transcription of Q proteins is autoregulated around stable levels [103], which translates into a constant $q = 1 - r_{\max}$ (and therefore $m + r = r_{max}$) for the whole interval $[0, T]$. We will see in next section that this hypothesis produces a very particular structure of the optimal control solution.

6.6 Biologically relevant scenarios

Despite their simplicity, self-replicator models have been capable of accounting for a number of observable phenomena during steady-state microbial growth, under the assumption that bacteria allocate their resources in such a way as to maximize growth. Here, we apply the general optimal allocation strategy derived in the previous section to predict the bacterial response to certain environmental changes. We consider two situations that commonly affect bacteria: changes in the nutrient concentration in the medium, and changes in the environment submitting the cell to a particular stress.

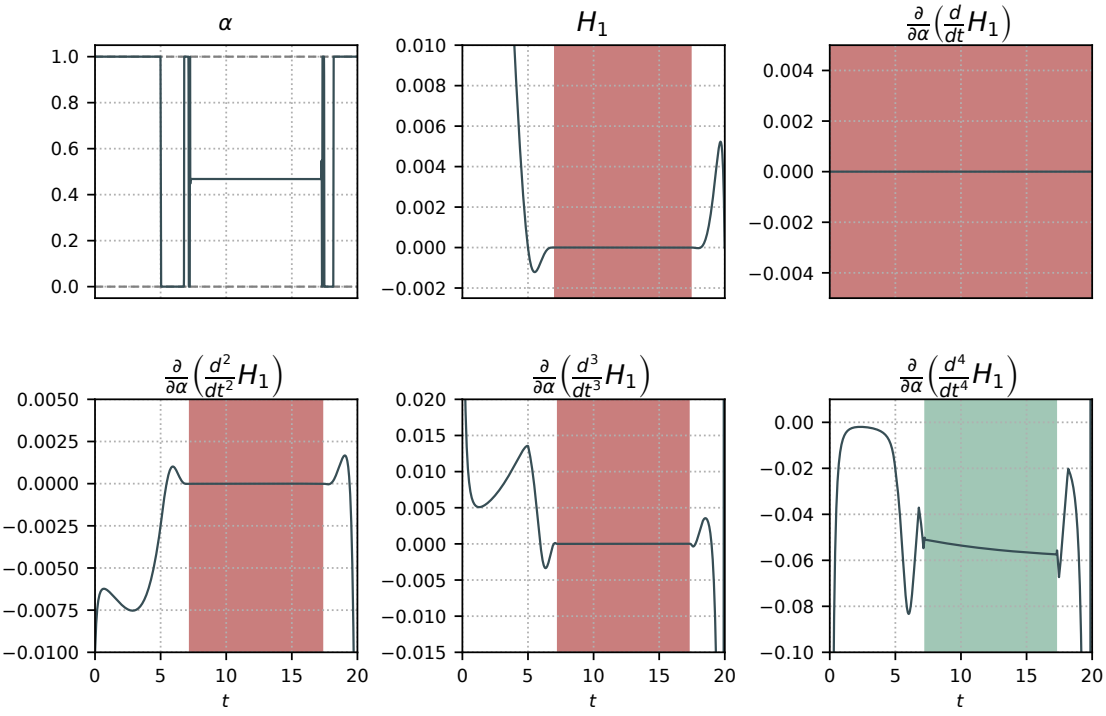


Figure 6.4: Factors of α in the derivatives of H_1 evaluated over the trajectory plotted in Figure 6.3. The intervals where the functions vanish are marked in red. As expected, all functions vanish along the singular arc except for the factor in the fourth derivative (highlighted in green) which is negative according to the Legendre-Clebsch condition (6.23).

6.6.1 Nutrient shift

Bacteria are known to traverse different habitats throughout their lifetime, experiencing fluctuating nutrient concentrations in the medium. In [30], we explored how bacteria dynamically adjust their allocation strategy when facing a nutrient upshift. In this work, we show that considering a class of growth rate-independent proteins in the model refines these previous results. We consider the optimal control problem with the initial state being the optimal steady state for a low value of E_M, and we set a higher E_M for the time interval $[0, T]$, representing a richer medium. Setting initial conditions at steady state has an impact on the singular arc of the optimal control: it holds that $m + r = r_{\max}$ and $q = 1 - r_{\max}$ for the whole trajectory, which yields a constant singular arc.

Theorem 6.6.1. *If $r(0) + m(0) = r_{\max}$ (i.e., q starts from a steady-state value),*

then any singular arc over the interval $[t_1, t_2]$ *of the optimal control corresponds to the optimal steady state.*

Proof. The dynamical equation for q is $\dot{q} = ((1 - r_{\max}) - q)w_R(p)(r - r_{\min})$, where it can be seen that the set $q = 1 - r_{\max}$ is invariant. This means that, for any trajectory emanating from a steady state, q remains constant even under changes of the nutrient quality E_M. Then, by using the relation (6.5), we obtain

$$m + r = r_{\max}. \tag{6.24}$$

Along the singular arc, it holds that $m + r = x(p)$, which, using (6.24), implies that $p = p^*_{\text{opt}}$, meaning that the precursor concentration along the singular arc is constant and optimal. Then, $\alpha_{\text{sing}} = \alpha^*_{\text{opt}}$, $m = m^*_{\text{opt}}$ and $r = r^*_{\text{opt}}$ for the whole singular arc. □

A numerical simulation of this scenario is shown in Figure 6.5. As expected, the increase in E_M produces a higher ribosomal mass fraction r, which translates into an increase of the growth rate, stabilizing at the maximal steady-state growth rate μ^*_{opt} through an oscillatory phase. It is noteworthy that, in comparison to Giordano *et al.*'s model, the relative changes in mass fractions r and m are much lower, which corresponds well with the relative changes observed in [2]. Additionally, while the presence of $r_{\min}$ does not noticeably affect the solution of the optimal control problem, it contributes to a model that more accurately reproduces the experimental data (Figure 6.2a), representing a significant improvement from the modeling point of view.

6.6.2 Bacterial response to stress

The other scenario of interest is an environmental change imposing a certain stress on the microbial population, which is counteracted through the synthesis of a stress response protein W. This protein is also growth rate-independent like Q, and its production can be triggered by many different situations. For instance, when subject to extreme temperatures, the production of so-called molecular chaperones helps bacteria counter the effect of protein unfolding [104, 90]. Likewise, the production of other proteins is known to protect bacteria like *E. coli* against acid stress [105]. Another possible scenario is the response to metabolic load imposed by the induced

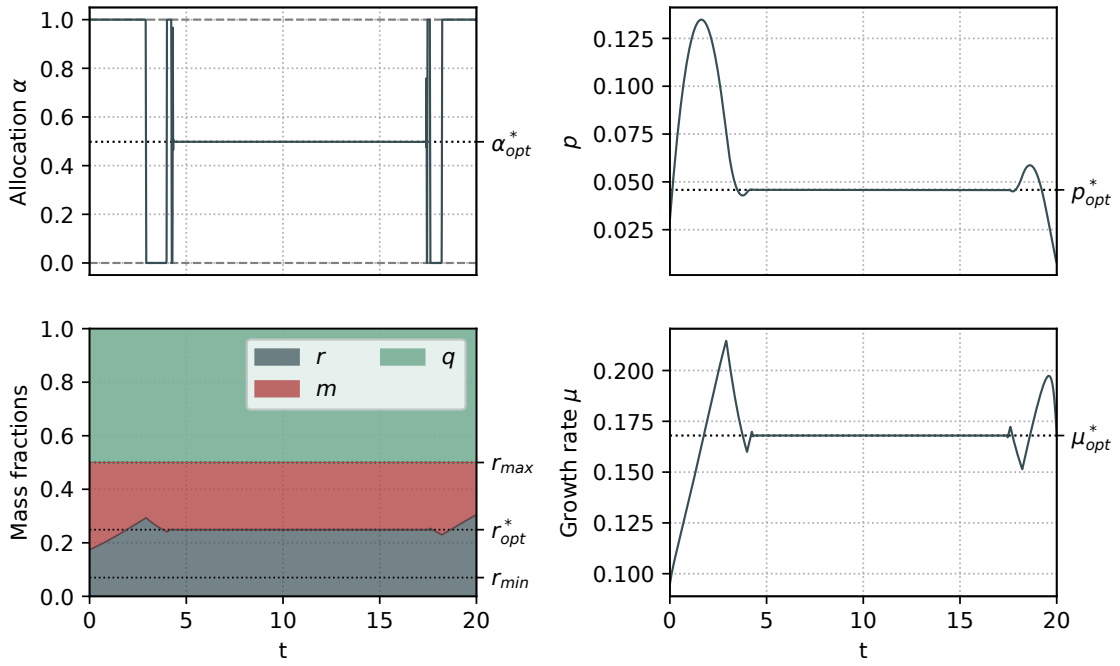


Figure 6.5: Numerical simulation of the optimal control problem starting from a steady state. The initial state corresponds to the optimal steady state for $E_M = 0.3$ (poor medium), and the new environmental constant is fixed to $E_M = 0.7$ (rich medium). As predicted, $m + r$ $(= 1 - q)$ remains constant, even if they vary individually, in opposition to the previous case. Naturally, an increase in the nutrient quality produces a higher steady-state ribosomal mass fraction r^*, which yields an increased steady-state growth rate μ^*_{opt} with respect to the growth rate before the upshift.

overexpression of a heterologous protein [106]. All of these situations are known to reduce the resources available for growth-associated proteins (Figure 6.6), consequently decreasing the maximal growth rate attainable. Here, we model a general stress response through the production of the W protein that takes up a fraction w of the proteome, thus reducing $r_{\max}$ to a certain $r^w_{\max} < r_{\max}$.

As before, we assume q takes up a constant fraction $1 - r_{\max}$ of the proteome, but the proportions of resources allocated to M and R are now $r^w_{\max}\alpha$ and $r^w_{\max}(1-\alpha)$ respectively. By construction, we have $w = r_{\max} - m - r$, which means we can express

$$\dot{w} = (r_{\max} - r^w_{\max} - w)w_R(p)(r - r_{\min}),$$

Figure 6.6: Left: original case. Right: new proposed case, where q remains unchanged, but the maximal allocation $m + r$ is restricted to a $r^w_{\max} < r_{\max}$.

showing that the mass fraction w converges asymptotically to the difference $r_{\max} - r^w_{\max}$. The remaining mass fractions p, r and m obey the dynamics of system (S'), so the application of the optimal solution found in last section is straightforward. An example is shown in Figure 6.7. As predicted, $m + r$ converges to the reduced $r^w_{\max}$, q remains constant at $1 - r_{\max}$ and w converges to $r^w_{\max} - r_{\max}$. The reduction of resources available for growth-associated proteins (M and R) causes the growth rate to drop, as was shown experimentally [2].

6.7 Conclusion

In this work, we proposed a dynamical self-replicator model of bacterial growth based on the work of [30], which introduces a growth rate-independent class of protein. As a consequence, the proteome of the bacterial cell can be divided into the metabolic machinery M, the gene expression machinery R, and the housekeeping machinery Q. While Q is growth rate-independent, this is also the case for a fraction of R required for cell replication to occur. As a consequence of this hypothesis, a maximum ribosomal concentration $r_{\max}$ appears in the model kinetics, limiting the allocation of resources to M and R. We studied the asymptotic behavior of the system, showing that, under certain conditions, all solutions converge towards the only globally attractive equilibrium. We then explored the optimal dynamic allocation strategies that consider maximizing the bacterial population volume in terms of the resource allocation parameter α. This involved a study of the static and dynamic aspects of optimal strategies. For the first one, we showed there is a unique optimal steady state, which corresponds to experimental observations of growing cultures of *E. coli* [5, 6, 2, 98]. The dynamic problem is approached through optimal control theory, by

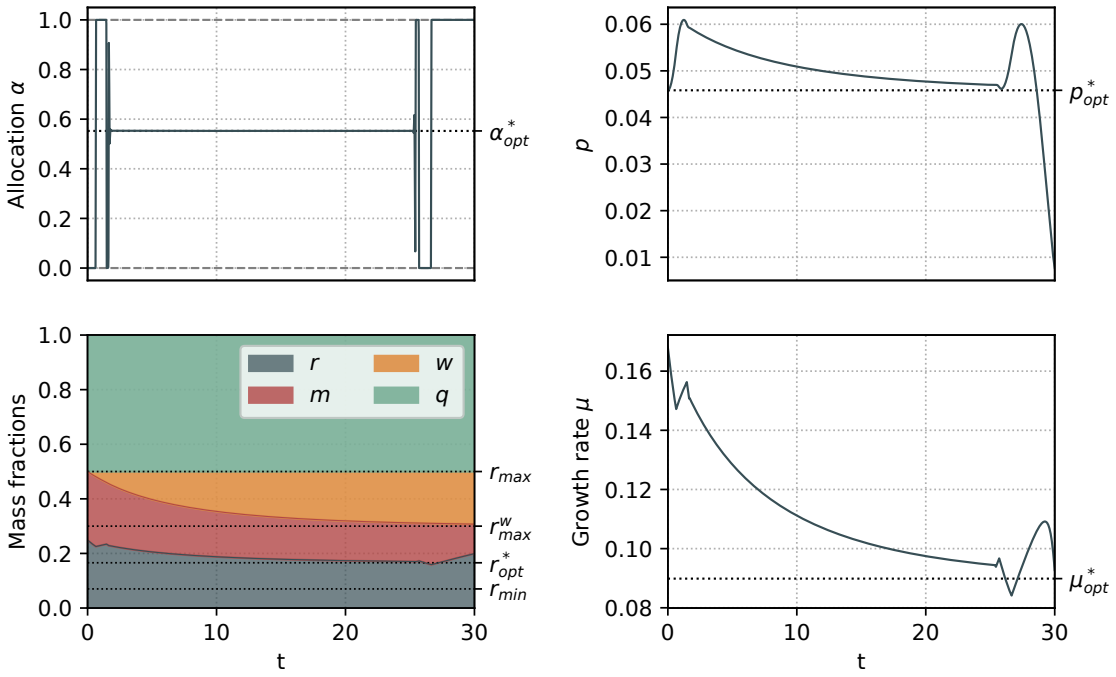


Figure 6.7: Numerical simulation of an optimal trajectory where the initial conditions are the optimal steady state for $E_M = 0.7$ and $r_{\max} = 0.5$. A certain stress is induced at $t = 0$, which triggers the synthesis of the growth rate-independent protein w, reducing the fraction $r_{\max}$ to $r^w_{\max} = 0.3$. As a result, the steady-state growth rate is significantly reduced.

application of the Pontrjagin's Maximum Principle. The obtained optimal control has a Fuller-singular-Fuller structure with a non-constant singular arc, in contrast to the constant singular arc obtained in Giordano *et al.*'s approach. We performed a detailed analysis of the OCP in both analytic and numerical ways. In particular, the singular arc of the optimal solution is characterized by i) its feedback form (i.e., being expressed as a function of the state only), ii) being exactly of order 2, and iii) the turnpike phenomenon (where the state trajectory and optimal control converge asymptotically towards the optimal steady state and control). Moreover, we showed that, when the mass fraction of class Q proteins is at steady state, the singular arc of the optimal solution corresponds to the optimal steady state. Additionally, we showed that the dynamical approach can be used to predict the behavior of the system when subject to stress. The latter is modeled through a reduction of the fraction of growth rate-dependent protein synthesis as the production of a w protein

that reduces $r_{\max}$.

While the main features of Giordano *et al.*'s work are present in this approach, our generalization shows a better agreement with the experimental data given by the introduction of the parameters $r_{\max}$ and $r_{\min}$ in the model. Additionally, the proposed partitioning of the proteome in a dynamic setting can account for certain natural phenomena known to reduce the fraction of growth rate-dependent proteins in the cell. These modifications yield interesting optimal control problems, which could potentially help understand the internal decision-making mechanisms evolved by bacteria.

Our approach was built on the joint exploitation of theoretical and numerical results. When tackling more complex problems as proposed, e.g., in Tsiantis & Banga [107], a PMP perspective tends to yield very complicated mathematical formulations. Using direct methods has the advantage of avoiding these issues, but it often requires some knowledge to initialize the optimization algorithm or to check the validity of the solutions. In order to investigate complex biological systems, we advocate the development and theoretical analysis of simple models, in line with the question to be investigated, coupled with numerical exploration of optimal solutions (using larger models if necessary).

Acknowledgments

We would like to acknowledge the help of Sacha Psalmon and Baptiste Schall from Polytech Nice Sophia for the numerical simulations and the production of the online example in the control toolbox gallery.

Chapter 7

Conclusion

In this manuscript, we explored a series of problems related to the allocation of resources in mathematical models of bacterial growth, and its applications to the production of compounds of interest.

Chapter 2, 3 and 4 have been dedicated to the production of metabolites in the most relevant cost-effective operating modes used in industry for commercial production. The main difference in these approaches lies in how the rich medium is supplied to the bioreactor and how the end product is extracted, which greatly affects the way the bacterial culture interacts with the environment. In every case, we studied the dynamical behavior of a self-replicator system representing the resource allocation problem in bacteria. An ideal case with constant substrate is analyzed in Chapter 2, representing an environment where bacteria sense no changes in the medium. In Chapter 3, the nutrient is assumed to be consumed by the culture until there is no available substrate in the bioreactor, which represents a batch process. Finally, the continuous bioreactor case is analyzed in Chapter 4, where there are both an inflow and an outflow producing—under certain conditions—a steady-state of the system, which is a well-studied behavior in this framework. The main trade-off intervening in these production schemes arise from the competition between the natural objective of the bacteria that aims to maximize the growth rate, and the biotechnological challenge of maximizing the synthesis of the compound of interest.

The optimal control problems arising from these works are tackled analytically and numerically, and can be considerably instructive in how to cope with such trade-offs in order to accomplish the proposed biosynthetic objectives. However, they frequently yield open-loop allocation strategies, which are often difficult to

implement. This issue is dealt with in Chapter 5, where a model-predictive control loop is proposed. The closed-loop implementation is based on the optimal open-loop strategies found in previous works, and offers a potential alternative to be explored.

Finally, a resource allocation model that considers growth rate-independent proteins is presented in Chapter 6. The latter represents a generalization of the model introduced in [30], and aims to study how bacteria can adapt to changing environments (such as nutrient upshifts and stress) from a dynamical perspective.

The models introduced here are simplified representations of very complex biological systems, which take into account two main cellular processes: metabolism and gene expression. The simplicity of the models allows a qualitative analysis of specific biological and biotechnological problems, and of the trade-offs arising from each objective. By design, several phenomena are not considered in the models, such as protein degradation and cell division. Additionally, the self-replicator systems are intended to represent an average composition of the cells forming the culture, but, in practice, resource distribution in individual cells is not necessarily homogeneous across the population. A major advantage of this methodology is that it allows to obtain analytical results to the proposed mathematical problems, in contrast to other more comprehensive approaches that aim to integrate numerous cellular mechanisms in a single system, yielding rather complex models.

There are numerous aspects to be explored in the future. All of the models are calibrated using data of previous steady-state experiments (or parameters in the literature obtained under the same principle). However, there is a lack of experimental results linking the trajectories of the dynamical models to the real-life bacterial behaviors. As explained in the introduction of this manuscript, a series of experiments with the *E. coli* bacteria were conducted throughout my PhD, with the objective of validating the theoretical results. The experiments were carried out in the context of the ANR Maximic project[1] with the Ibis team (Inria Grenoble - Rhône-Alpes) with H. de Jong, J. Geiselmann and E. Cinquemani. However, experimentation in the microbiological framework requires multiple iterations in order to consider the obtained data as representative of the studied phenomenon, which was not possible to accomplish due to lack of time. Yet, it remains an important follow-up of this project.

Another potential extension of these works is the inclusion of additional bio-

[1] `https://project.inria.fr/maximic/`

logical mechanisms and constraints into the bacterial models, and the study of the impact of such modifications in the resulting optimal allocation strategies. This was the objective of the work described in Chapter 6, where the presence of growth rate-independent proteins (representing, for example, stress response proteins) imposes an upper bound on the ribosomal mass fraction of the cell. An important mechanism to contemplate in such models could be the energy metabolism, known to play an important limiting role in protein synthesis, which could refine the results previously obtained. Similarly, changing the optimization criteria can yield very different results in the distribution of resources to different cellular functions. The hypothesis of growth-gate maximization has been validated in many experiments of *E. coli*, and proven to be a selective advantage in certain environments. However, such assumptions are context-dependent, and could loose validity under different environmental conditions.

A modelling assumption that can be reviewed is the action of the external control I, modelled in (2.1) (and in the rest of the manuscript) through a multiplicative term affecting directly the natural allocation process. The idea behind this modelling choice is that the external control can arrest the production of RNA polymerase. However, an equally valid hypothesis is that the control can stall general protein synthesis, including enzymes, which would drastically change the control problem. In fact, the latter is part of an ongoing project extending the work on MPC loops presented in this manuscript. The approach is also complemented by a stage of state and parameter estimation, towards an implementation of a real-time bacterial resource allocation controller.

Finally, we recall that the work presented in Chapter 4 relies on a simplifying assumption linking metabolite production and protein synthesis through a certain function of the ribosomal concentration. This hypothesis provides an additional mass conservation law allowing to study the original dynamical system through a lower dimensional limiting system. However, the general case remains unstudied, as even the analysis of the local equilibria through the Routh-Hurwitz stability criterion represents a challenge for such system. This motivated an alternative approach to the problem, in which the stability of the local equilibria is examined by means of real algebraic tools, in collaboration with M. Safey El Din (PolSys team, LIP6, Sorbonne Université).

Manipulation of bacterial cultures for the production of added-value compounds

remains a complex endeavour. Hopefully, theoretical research on the field, boosted by our growing understanding of microorganisms, will progressively improve the existing biotechnological techniques towards more efficient and sustainable industrial schemes.

Appendix A

Time-optimal control of piecewise affine bistable gene-regulatory networks

This chapter reproduces [AYO1] and [AYO2], written in collaboration with Nicolas Augier, and accepted for publication in the International Journal of Robust and Nonlinear Control and the 7th IFAC Conference on Analysis and Design of Hybrid Systems respectively. While these works are not in line with the topic of resource allocation in bacterial growth, they have been included in the Appendix as they were produced during my PhD as a result of a joint work. In accordance with the usages in the mathematics field, the authors of this publications are in alphabetical order, regardless of the fact that we equally contributed to the works.

A.1 Introduction

Understanding complex biological phenomena has become of great interest in the last decades for the scientific community. In the context of synthetic biology, numerous fields of study have been employed to better comprehend and re-engineer the interactions within biological systems [108]. Such is the case of control theory, widely used to explain regulatory mechanisms in nature [30, 109], but also to artificially act upon them for biotechnological purposes [AYT5, AYT2]. A classical example is the metabolism of cells, described by multiple regulatory mechanisms forming complex networks. In this framework, the interaction between genes is a

crucial subject of study [110], whose typical behaviors can be described by positive and negative feedback loops [111, 112, 113, 114, 115]. The dynamics of these loops have been extensively analyzed, both from experimental and theoretical perspectives, and are known to present either multistability or oscillatory behaviors. From a mathematical modeling perspective, such a systems can be modeled through dynamical systems of several variables, where the positivity or negativity is given by the parity of negative interactions forming the loop.

Among all existing patterns, the simplest positive feedback loop is the two-dimensional bistable system, which is commonly used to represent the so-called *genetic toggle switch*. The latter is a synthetic flip-flop device first implemented experimentally in *E. coli* through the genes *lacI* and *tetR* mutually repressing each other [116]. The state of the device is determined by the concentration of the genes in the boolean form $(\text{low}, \text{high})$ and $(\text{high}, \text{low})$. This allows genetic toggle switches to act as biological memory units capable of storing 1 bit of information, by sustaining one of the two possible states through time [117], which offers a biosynthetic alternative to the classical electronic flip-flop. Since its creation, understanding how to regulate bistable systems in a reliable manner (e.g. by suppressing undesirable oscillations [118] or achieving transitions between states [AYO2] [119, 120]) has become highly relevant for their vast implications in biotechnology and biocomputing.

In practice, the state of a genetic toggle switch can be controlled by externally catalyzing or inhibiting the synthesis rate of the genes. This is done by introduction of a plasmid, which are essentially small circular DNA molecules that can be constructed to include an inducible promoter of the studied gene, thus affecting the synthesis rate of messenger RNA. Thus, the transcription rate can be directly modified by aggregation of an inducer. In *E. coli*, this is done by externally adding the diffusible molecules IPTG[1] and aTc[2], which are known to repress the *lacI* and *tetR* genes, respectively [116].

Motivated by this experimental scheme, some authors proposed exact control strategies based on a piecewise affine model of the bistable system [120]. As discussed in the work, the importance of studying the phenomenon through qualitative models arises from the constraints related to the experimental setup, both in measuring the state and in acting on the system. The proposed model is characterized by the

[1] isopropyl-β-D-thiogalactopyranoside
[2] anhydrotetracycline

existence of an "undifferentiated state", where no gene is predominant, and from which the system can evolve towards one of the two attractors. Mathematically, the unstability of this state appears as a Filippov non-smooth "saddle" singularity [121]. From a biological point of view, such a state plays a key role in cell decision making and cell fate differentiation [122]. Its role in fate commitment has also motivated experimental studies aiming at stabilizing genetic toggle switches around this undifferentiated point [123, 124].

Whereas most of the theoretical work in the subject has been dedicated to externally producing state transfers [125], the time efficiency of state switches has received little or no attention from the community. Indeed, one of the key issues in these genetic devices is the time needed to induce a transfer between its two stable states, due to its importance when studying more complex networks of systems involving different time scales. In particular, the latter becomes a major constraint in the framework of biological signal processing [126]. Genetic toggle switches operate at the level of gene transcription and translation, whose duration and timescales are the main factor delaying the availability of the proteins when facing a switch between steady states. In this context, the minimization of a state switch, which is directly linked to the production of the non-expressed protein, becomes highly relevant. Recent works [127] showed the importance of accelerating transitions times (and minimizing inducer usage) in artificially engineered bistable systems in order to obtain less costly (and therefore, more sustainable) chemical production schemes. Thus, in this paper, we investigate the time-optimal control strategies for the aforementioned bistable system. Our aim is to induce transitions between the two stable steady states in minimal time. Many complex systems are known to involve bistable processes [128, 129]. Hence, the reduction of the time needed for such transfers could allow experimentalists to speed up certain chemical reactions or to artificially increase bacterial growth rate, thus improving yield in biotechnological processes. In a different setting [130, 131], the time efficiency of bistable systems switches in two-level quantum systems was studied, so as to induce efficient transitions between two quantum states.

In addition to the biological relevancy of the subject, the resulting OCP (Optimal Control Problem) yields very interesting results in the framework of Hybrid Optimal Control. The steady states of the piecewise linear system cannot be reached in finite time due to the lack of controllability in certain regions, and so one has to consider

a relaxed OCP with "partial targets", that is, driving a given protein to a certain fixed value larger than its corresponding threshold. In this regard, we show that time-optimal strategies for such a problem have a very specific geometric description. When the initial state is far enough from the target, that is below a curve called *separatrix*, we show by an adaptation of the HMP (Hybrid Maximum Principle) to our setting that the optimal control consists in a concatenation of two bang arcs, and the optimal trajectories follow:

- a first phase in which the system reaches the separatrix;
- a second phase where the system slides along this curve, until reaching the "undifferentiated" point of the biological system in finite time;
- a third phase, where the system leaves this curve, slides along a second fixed curve and reaches its target.

These two curves correspond to the stable and unstable manifolds of the undifferentiated saddle-type singularity, and the point where the dynamics achieves its transfer is nothing but the corresponding Filippov equilibrium. The latter behavior can be compared to the turnpike phenomenon [71], where the optimal trajectory for a given OCP for large final times is shown to remain close to a steady-state trajectory solution of the associated static OCP. Besides its specific interest, we expect our method to open new prospects in the study of optimal control of higher dimensional genetic regulatory networks. In particular, it often occurs that trajectories belonging to a given domain may bifurcate into different domains, similarly to what happens in the toggle switch case, and some similar turnpike-like properties may hold in this case.

The paper is organized as follows: in Section A.2, we present both the non-controlled system and the studied controlled system, and we provide some technical results. In Section A.3, we introduce the time-optimal control problem and we adapt the HMP to our setting. In Section A.4, we present the main results, that prove the qualitative features of the optimal trajectories mentioned above. In Section A.6, we give a lower bound for the minimal time, then characterizing the minimal transfer time between the two states of the toggle switch model. In Section A.7, we provide numerical results implemented with Bocop [4], an open-source toolbox for solving OCPs. Additionally, we perform a numerical comparison between the trajectories of the relaxed OCP and the original OCP, that suggests that the results also hold for the original one.

A.2 Bistable-switch model

A.2.1 Free dynamics

Consider two variables x_1 and x_2 which represent two genes mutually inhibiting each other. The individual dynamics, defined in Filippov sense, is the following

$$\begin{cases} \dot{x}_1 = -\gamma_1 x_1 + k_1 s^-(x_2, \theta_2), \\ \dot{x}_2 = -\gamma_2 x_2 + k_2 s^-(x_1, \theta_1), \end{cases} \tag{A.1}$$

where for $j \in \{1, 2\}$, $x_j \in \mathbb{R}_+$, and for $\theta \in \mathbb{R}$, $s^-(\cdot, \theta) : \mathbb{R} \to \mathbb{R}$ is such that

$$s^-(x, \theta) = \begin{cases} 1 & \text{if } x < \theta, \\ 0 & \text{if } x > \theta. \end{cases}$$

It is assumed that $s^-(x) \in [0, 1]$ for $x = \theta$. The positive constants $(\gamma_j)_{j\in\{1,2\}}$, $(k_j)_{j\in\{1,2\}}$ correspond, respectively, to the degradation and the production rates of each variable. It is classical [120] that the domain $K = [0, k_1/\gamma_1] \times [0, k_2/\gamma_2]$ is forward invariant by the dynamics of Equation (A.1). From now on, we consider only solutions evolving in K. Define the regular domains

$$\begin{aligned} B_{00} &= \left\{(x_1, x_2) \in \mathbb{R}^2 \mid 0 < x_1 < \theta_1,\ 0 < x_2 < \theta_2\right\}, \\ B_{01} &= \left\{(x_1, x_2) \in \mathbb{R}^2 \mid 0 < x_1 < \theta_1,\ \theta_2 < x_2 < \tfrac{k_2}{\gamma_2}\right\}, \\ B_{10} &= \left\{(x_1, x_2) \in \mathbb{R}^2 \mid \theta_1 < x_1 < \tfrac{k_1}{\gamma_1},\ 0 < x_2 < \theta_2\right\}, \\ B_{11} &= \left\{(x_1, x_2) \in \mathbb{R}^2 \mid \theta_1 < x_1 < \tfrac{k_1}{\gamma_1},\ \theta_2 < x_2 < \tfrac{k_2}{\gamma_2}\right\}, \end{aligned}$$

which are defined as open sets in accordance with the HMP approach to be applied in Section A.3.3. Equation (A.1) restricted to a regular domain B_{ij} is an affine dynamical system on $\mathbb{R}^2$ having an asymptotically stable equilibrium, called *focal point* for system (A.1). Each region B_{ij} for $i, j \in \{0, 1\}$ has a focal point

$$\phi_{ij} = (\bar{x}_i, \bar{x}_j)$$

corresponding to

$$\bar{x}_i = \frac{k_i}{\gamma_i} s^-(\bar{x}_j, \theta_j).$$

Thus, system (A.1) has two locally asymptotically stable steady states

$$\phi_{10} = \left(\frac{k_1}{\gamma_1}, 0\right) \in \bar{B_{10}},$$
$$\phi_{01} = \left(0, \frac{k_2}{\gamma_2}\right) \in \bar{B_{01}},$$

and an unstable Filippov equilibrium point at (θ_1, θ_2). Figure A.1 illustrates the dynamics of the system for a given set of parameters.

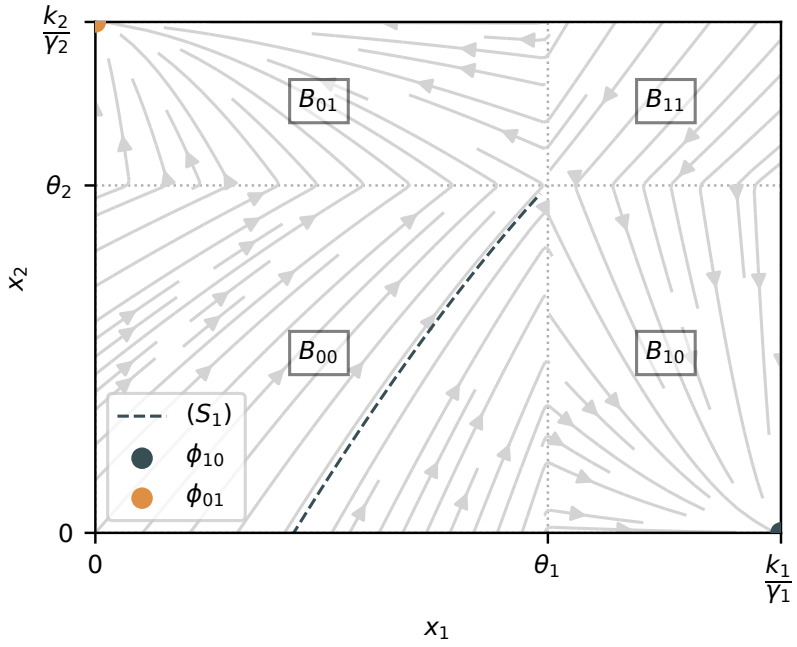


Figure A.1: Stream plot with free dynamics given by Equation (A.1). System parameters are $\gamma_1 = 1.1$, $\gamma_2 = 1.7$, $\theta_1 = 0.6$, $\theta_2 = 0.4$, $k_1 = k_2 = 1$.

A.2.2 Controlled dynamics and some related properties

We write the controlled dynamics assuming that the synthesis rates of each gene can be externally catalyzed or inhibited (e.g. through the introduction of inducible promoters of a given gene), as indicated in the previous section. Mathematically, this is represented by the control input u acting directly on the synthesis rate of each gene, in a multiplicative form. Then, the controlled system, defined in Filippov

sense, is

$$\begin{cases} \dot{x}_1 = -\gamma_1 x_1 + u(t) k_1 s^-(x_2, \theta_2), \\ \dot{x}_2 = -\gamma_2 x_2 + u(t) k_2 s^-(x_1, \theta_1), \end{cases} \tag{S}$$

where the control $u(\cdot) \in L^\infty([0, t_f], [u_{\min}, u_{\max}])$, with $0 < u_{\min} < 1 \leq u_{\max}$. This system is motivated by the one introduced in [120], which assumed the same control input u for the two variables. The latter aims to model a simple qualitative control, easier to implement in a molecular biology setting than the case with two distinct control variables. We make the following assumptions on the parameters of the system (for more details, see [120]).

Assumption A.2.1. *The parameters $(\gamma_j)_j$ and $(k_j)_j$ satisfy*

$$\theta_j < \frac{k_j}{\gamma_j},\ j \in \{1, 2\}; \qquad \frac{\theta_2}{\theta_1} > \frac{k_2}{k_1}\frac{\gamma_1}{\gamma_2}; \qquad \frac{\theta_2}{\theta_1} < \frac{k_2}{k_1}.$$

This assumption is based on intrinsic conditions of the parameters of the non-controlled system, and allows to find a control strategy driving the solution of Equation (S) from B_{10} to B_{01}, as well as from B_{01} to B_{10}. Note that it implies $\gamma_1 < \gamma_2$, and that the case where $\frac{\theta_2}{\theta_1} > \frac{k_2}{k_1}\frac{\gamma_1}{\gamma_2}$ and $\frac{\theta_2}{\theta_1} < \frac{k_2}{k_1}$ can be treated analogously by permutation of x_1 and x_2.

Separatrix

Now define the *separatrix*, which is a curve playing a fundamental role in in the global dynamics of both the open-loop system (A.1) and the controlled system (S). For a fixed value of $u(t) \equiv u \in [u_{\min}, u_{\max}]$, the separatrix (S_u) is defined as the stable manifold of the Filippov equilibrium (θ_1, θ_2) for Equation (S) restricted to $\overline{B_{00}}$. In the coordinates $(x_1, x_2) \in B_{00}$, for $u \geq 1$, the separatrix (S_u) can be written as the curve of equation

$$x_2 = \alpha(x_1, u) = \frac{k_2 u}{\gamma_2} - \left(\frac{k_2 u}{\gamma_2} - \theta_2\right)\left(\frac{\dfrac{k_1 u}{\gamma_1} - x_1}{\dfrac{k_1 u}{\gamma_1} - \theta_1}\right)^{\frac{\gamma_2}{\gamma_1}}.$$

Using the latter, we define the regions

$$(S_u)^+ = \left\{ (x_1, x_2) \in \mathbb{R}^2 \mid 0 < x_1 < \theta_1,\ \alpha(x_1, u) < x_2 < \frac{k_2}{\gamma_2} \right\},$$
$$(S_u)^- = \left\{ (x_1, x_2) \in \mathbb{R}^2 \mid 0 < x_2 < \theta_2,\ \alpha(x_1, u) > x_2,\ x_1 < \frac{k_1}{\gamma_1} \right\},$$

such that the domain $\overline{K}$ is divided into

$$\overline{K} = \overline{(S_u)^+} \cup \overline{(S_u)^-} \cup \overline{B_{11}}, \tag{A.2}$$

as shown in Figure A.2. The solutions of Equation (S) having initial conditions in

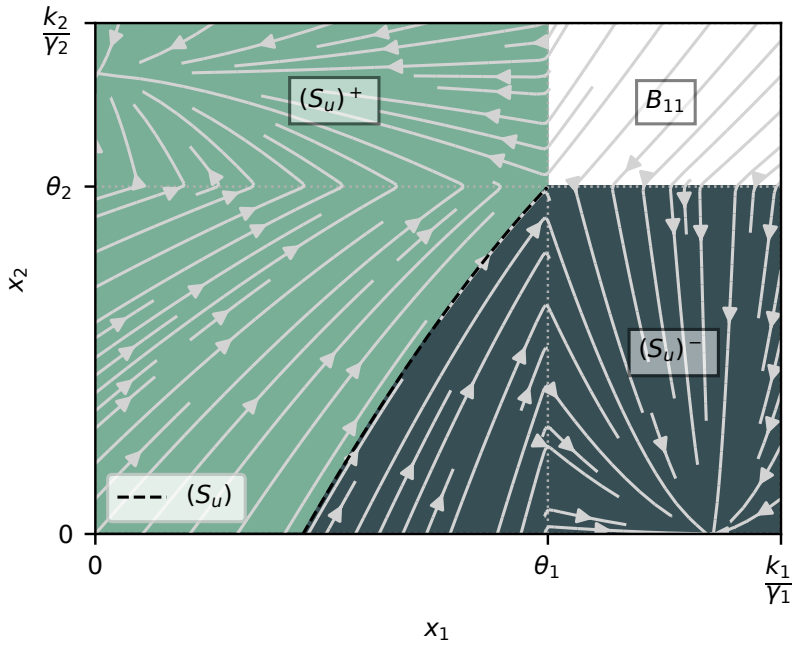


Figure A.2: Division of the domain $\overline{K}$ as defined in (A.2), with a vector field defined by a constant control $u < 1$.

$(S_u)^-$ (respectively, $(S_u)^+$) reach B_{10} (respectively, B_{01}) in finite time. Moreover, B_{10} (respectively, B_{10}) is included in the bassin of attraction of ϕ_{10} (respectively, ϕ_{01}). Notice that, for a fixed value of $u(t) \equiv u \in [u_{\min}, u_{\max}]$, the solutions of Equation (S) having initial conditions in (S_u) reach the Filippov point (θ_1, θ_2) in finite time. Once having reached this point, the solution of Equation (S) is then defined by differential inclusion in the Filippov sense. Roughly speaking, there exist

several solutions that will reach either B_{01} or B_{10} (see [120, Appendix] for more precise informations about Filippov solutions of such a system).

Lower separatrix

Now we define the *lower separatrix*, which will be useful in Section A.6.

Definition A.2.2. *For $(\gamma_j)_{j\in\{1,2\}}$ and $(k_j)_{j\in\{1,2\}}$ satisfying Assumption A.2.1, define the* lower separatrix $(\widetilde{S_u})$ *as the straight line of equation*

$$x_2 = \beta(x_1, u) = \frac{k_2 u}{\gamma_1} - \left(\frac{k_2 u}{\gamma_1} - \theta_2\right)\left(\frac{\frac{k_1 u}{\gamma_1} - x_1}{\frac{k_1 u}{\gamma_1} - \theta_1}\right).$$

Lemma A.2.3. *Let $(\gamma_j)_{j\in\{1,2\}}$ and $(k_j)_{j\in\{1,2\}}$ satisfy Assumption A.2.1 and $u \geq 1$, and let $\bar{x}_1$ be the unique $x_1 \in [0, \theta_1]$ such that $\beta(x_1, u) = 0$. Then for every $x_1 \in [\bar{x}_1, \theta_1]$ and $u \geq 1$, we have $\beta(x_1, u) \leq \alpha(x_1, u)$, that is, $(\widetilde{S_u})$ is below (S_u) for every $x_1 \in [\bar{x}_1, \theta_1]$.*

Proof. For every $x_1 \in [0, \theta_1]$ and $u \geq 1$, define $X(x_1) = \frac{\frac{k_1 u}{\gamma_1} - x_1}{\frac{k_1 u}{\gamma_1} - \theta_1}$. We have easily that $\beta(x_1, u) = 0$ if and only if $x_1 = \bar{x}_1$, where $\bar{x}_1 \in [0, \theta_1]$ is such that $X(\bar{x}_1) = \frac{k_2 u/\gamma_1}{k_2 u/\gamma_1 - \theta_2}$. Evaluating in the expression of α at $x_1 = \bar{x}_1$, we have

$$\alpha(\bar{x}_1, u) = \frac{k_2 u}{\gamma_2} - \left(\frac{k_2 u}{\gamma_2} - \theta_2\right)\left(\frac{k_2 u/\gamma_1}{k_2 u/\gamma_1 - \theta_2}\right)^{\frac{\gamma_2}{\gamma_1}}$$

Using the fact that $\gamma_2 > \gamma_1$ under Assumption A.2.2, one can prove by a direct differentiation that the function

$$x \mapsto \frac{\left(\frac{k_2 u}{\gamma_1} - x\right)^{\frac{\gamma_2}{\gamma_1}}}{\frac{k_2 u}{\gamma_2} - x}$$

is non-decreasing on $[0, \theta_2]$, and that $x_1 \mapsto \alpha(x_1, u)$ is concave for $x_1 \in [0, \theta_1]$. Hence, we have

$$\left(\frac{\frac{k_2 u}{\gamma_1}}{\frac{k_2 u}{\gamma_1} - \theta_2}\right)^{\frac{\gamma_2}{\gamma_1}} \leq \frac{\frac{k_2 u}{\gamma_2}}{\frac{k_2 u}{\gamma_2} - \theta_2},$$

and we can deduce $\alpha(\bar{x}_1, u) \geq 0$. Provided that $\alpha(\theta_1, u) = \beta(\theta_1, u) = \theta_2$, we deduce that $\alpha(x_1, u) \geq \beta(x_1, u)$, for every $x_1 \in [\bar{x}_1, \theta_1]$. □

A.3 Time-optimal transfer

A.3.1 Problem formulation

The state of a genetic toggle switch is determined by gene expression in the boolean form $(\text{low}, \text{high})$ and $(\text{high}, \text{low})$, and so the objective in this work is to achieve a transition from one boolean state to the other in minimal time. In the mathematical context, the latter translates into finding trajectories that drive the solution $(x_1(t), x_2(t))$ of Equation (S) towards the steady states ϕ_{01} and ϕ_{10} of Equation (A.1) in minimum time (where these states correspond to the differentiated states aforementioned). However, due to the lack of controllability in direction x_1 (respectively, x_2) of Equation (S) restricted to B_{01} (respectively, B_{10}), one has to relax the problem. More precisely, the steady state ϕ_{01} (respectively, ϕ_{10}) cannot be reached in finite time, because u does not act on x_1 in the domain B_{01} (respectively, x_2 in the domain B_{10}). Thus, we will be first interested in driving $x_2(t)$ towards an arbitrary value $x_2(t_f) = x_2^f > \theta_2$ (for instance, the value $x_2^f = k_2/\gamma_2$ corresponding to the x_2-component of the steady state ϕ_{01}), with the constraint that at the final time, $x_1(t_f)$ belongs to the interval $[0, \theta_1)$. This target choice ensures that, at the final time, the gene x_2 is strongly expressed while the gene x_1 is weakly expressed. The symmetric problem, which is equivalent, consists in driving $x_1(t)$ towards an arbitrary value $x_1(t_f) = x_1^f > \theta_1$, with the constraint that at the final time, $x_2(t_f)$ belongs to the interval $[0, \theta_2)$. Fix $x_1^0 \geq \theta_1$, $x_2^0 \leq \theta_2$, $x_2^f \geq \theta_2$, and consider the minimization problem

$$\begin{cases} \textit{minimize } t_f \geq 0, \\ x(t) = (x_1(t), x_2(t)) \textit{ is subject to } (\text{S}), \\ x(0) = (x_1^0, x_2^0), \\ x_2(t_f) = x_2^f, \\ x_1(t_f) \in [0, \theta_1), \\ u(\cdot) \in [u_{\min}, u_{\max}]. \end{cases} \tag{OCP}$$

A.3.2 Reachability of the terminal state

A fundamental aspect of OCPs with fixed terminal state is the existence of a solution. Such a matter is directly linked to the reachability and controllability analysis of the dynamical system, which are often hard to conduct analytically. In this work, we provide sufficient conditions for the feasibility of the proposed trajectory, and we show that a simple piecewise constant control strategy [120] achieves the objective, serving as a candidate to (OCP). This strategy drives asymptotically the system from an initial state in B_{01} to ϕ_{10} (or B_{10} to ϕ_{01} for the symmetric problem). One can show that, under Assumption A.2.1, there exist $u_{\min} < \theta_1\gamma_1/k_1$ and $u_{\max} \geq 1$ such that

$$\Phi^*(u_{\min}) \in (S_{u_{\max}})^+,$$

with

$$\Phi^*(u_{\min}) \doteq \left(\frac{u_{\min}k_1}{\gamma_1}, \frac{u_{\min}k_2}{\gamma_2}\right). \tag{A.3}$$

In previous works [120, Section 3], authors proved the existence of $\bar{u}_{\min}, \bar{u}_{\max}$ such that for every $u_{\min}, u_{\max}$ such that $0 \leq u_{\min} \leq \bar{u}_{\min} \leq \bar{u}_{\max} \leq u_{\max}$, we have $\Phi^*(u_{\min}) \in (S_{u_{\max}})^+$. The latter condition is satisfied for $\bar{u}_{\min}, \bar{u}_{\max}$ when there exists $\delta < \frac{\theta_1 k_2}{\theta_2 k_1} - 1$ and $\epsilon > 0$ small enough such that

$$\bar{u}_{\max} > \max\left\{1, \frac{(1+\delta)\gamma_2\dfrac{\theta_2}{k_2} - \gamma_1\dfrac{\theta_1}{k_1}}{\delta}\right\},$$

and

$$0 < \bar{u}_{\min} < \min\left\{\frac{\gamma_1}{k_1}\theta_1, \frac{\gamma_2}{k_2}\theta_2, (1-\epsilon)x_1^*\frac{\gamma_1}{k_1}\right\},$$

where $x_1^\star \in (0, \theta_1)$ is the unique solution of $\alpha(x_1^\star, u_{\max}) = 0$. Concerning the symmetric problem, one can show the existence of another choice of $u_{\min} < \theta_1\gamma_1/k_1$ and $u_{\max} \geq 1$ such that $\Phi^*(u_{\min}) \in (S_{u_{\max}})^-$. Due to the symmetry of both problems, we will focus on the first case, and state the following assumption.

Assumption A.3.1. *Bounds u_{min} and u_{max} are chosen such that $0 \leq u_{min} \leq \bar{u}_{min} \leq$*

$\bar{u}_{max} \leq u_{max}$, *so that* $\Phi^*(u_{min}) \in (S_{u_{\max}})^+$.

Based on the solution developed in previous works [120], we first propose an input control constrained to two possible values $\{u_{min}, u_{max}\}$ corresponding to the low and high synthesis control. The control law is expressed in terms of the state and time as

$$u(x,t) = \begin{cases} u_{min} & \text{if } x \in B_{10}, \\ u_{min} & \text{if } t \in [0, t_s), x \in B_{00}, \\ u_{max} & \text{if } t \in [t_s, \infty), x \in B_{00}, \\ u_{max} & \text{if } x \in B_{01}. \end{cases} \tag{A.4}$$

for $t_s > 0$ sufficiently large. During the first phase with $u \equiv u_{min}$, every focal point of the system belongs to B_{00}, hence the solution $x(t)$ of Equation (S) converges towards the point $\Phi^*(u_{\min}) \in B_{00}$ when $t \to \infty$. During the second phase with $u \equiv u_{max}$, state $x(t)$ reaches B_{01} in finite time, and $x_2(t)$ converges towards x_2^f in finite time. From that point, an open-loop control $u \equiv 1$ drives $x(t)$ to ϕ_{01} when $t \to \infty$. An example illustrating this trajectory is shown in Figure A.3, where $x_2^f = k_2/\gamma_2$, matching the coordinate x_2 of the point ϕ_{01}. Indeed, under Assumptions A.2.1 and A.3.1, and by choosing t_s sufficiently large, the control strategy (A.4) ensures that any trajectory starting from (x_1^0, x_2^0) reaches a final point meeting $x_1 \in [0, \theta_1)$ and $x_2 = x_2^f$ in finite time, which shows that the set of admissible controllers for problem (OCP) is non empty.

Notice that, while the latter strategy serves as a candidate, the set of possible controllers is not limited to bang-bang solutions. One could consider, for instance, non bang-bang strategies based on A.4 with intermediate control values (e.g. replacing $u_{\max}$ by $\widetilde{u}_{\max} < u_{\max}$) which also achieve the transfer. This implies that there exist several trajectories reaching the target of Problem OCP, which motivates a study from a Pontryagin's Maximum Principle perspective.

A.3.3 Hybrid optimal control problem with a fixed domain sequence

Consider two compact subsets M_0 and M_1 of $\mathbb{R}^2$, and assume M_1 is reachable from M_0 for system (S), that is, such that there exists a time $t_f > 0$, a control $u(\cdot) \in L^\infty([0, t_f], \Omega)$ and $x_0 \in M_0$ such that the solution $x(t)$ of Equation (S), defined in

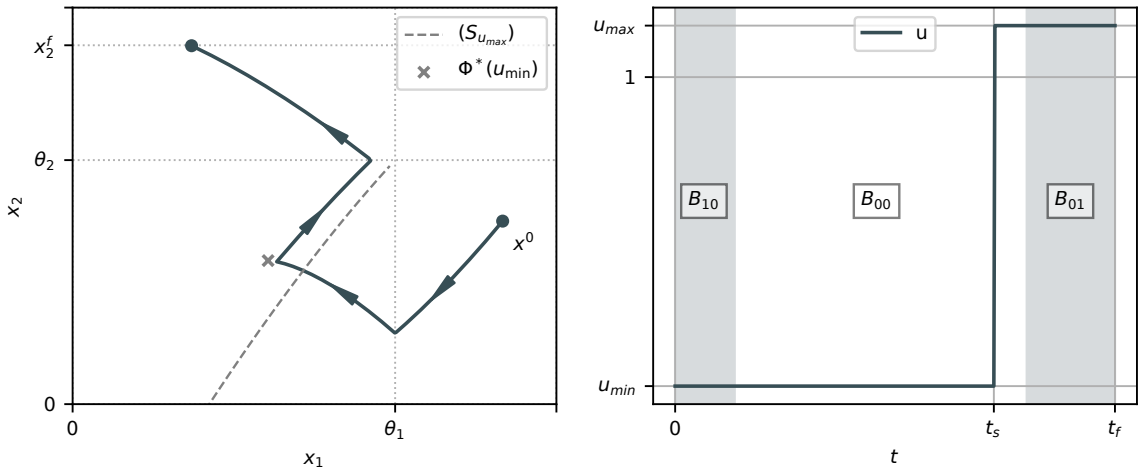


Figure A.3: Optimal trajectory with $x_1^0 = 0.8$, $x_2^0 = 0.3$ and $x_2^f = k_2/\gamma_2$. System parameters are $\gamma_1 = 1.1$, $\gamma_2 = 1.7$, $\theta_1 = 0.6$, $\theta_2 = 0.4$, and $k_1 = k_2 = 1$. Control bounds are set to $u_{min} = 0.4$ and $u_{max} = 1.1$. The control switches from $u \equiv u_{\min}$ to $u \equiv u_{\max}$ at time t_s ($= 3$ in this case), after the state $x(t)$ has crossed the separatrix $(S_{u_{\max}})$.

the Filippov sense with initial condition $x(0) = x_0$ satisfies $x(t_f) \in M_1$. Consider the problem of steering the system (S) from M_0 to M_1 in minimal time t_f. In order to properly define the problem, one has to choose a sequence B in the set $\{B_{00}, B_{01}, B_{10}, B_{11}\}$ of regular domains, and consider B-admissible trajectories of Equation (S), defined as follows.

Definition A.3.2. *Let $B = (B_j)_{j\in\{1,\dots,k\}}$ be a sequence of regular domains. We say that a solution $x(t)$ of Equation* (S) *is* B-admissible *if there exists a time $T > 0$, a control $u(\cdot) \in L^\infty([0, t_f], \Omega)$, and times $t_0 = 0 < t_1 < \cdots < t_k$ such that $x(t) \in B_j$ for every $t \in \Delta_j$, where $\Delta_j = (t_j, t_{j+1})$.*

In particular, the previous definition excludes sliding modes along the frontier between two successive regular domains. Additionally, we require two more assumptions related to the reachability of B-admissible solutions for the general case.

Assumption A.3.3. *M_0 (respectively, M_1) is included in the adherence $\bar{B}_{jk}$ (respectively, $\bar{B}_{qi}$) of a regular domain, for $j, k, q, i \in \{0, 1\}$.*

Assumption A.3.4. *Assume that there exists a time $T > 0$ and a B-admissible solution $(x(t), u(t))$ such that $x(0) \in M_0$ and $x(T) \in M_1$.*

Notice that for given sets M_0, M_1, the choice of the sequence B is not unique in general. Assume that M_0, M_1, B satisfy the assumptions A.3.3 and A.3.4. For a fixed sequence $B = (B_j)_{j\in\{1,\dots,k\}}$, we can consider Problem (A.4.1) restricted to B-admissible trajectories. Necessary conditions of optimality for this problem can be directly derived from the HMP, which we will state in Theorem A.3.1 of the following subsection.

A.3.4 Hybrid Maximum Principle for time optimal control

In this section, we provide an adaptation of the Hybrid Maximum Principle given by Dmitruk and Kaganovich [132] to the time optimal setting. Let $t_0 < t_1 < \cdots < t_\nu$ be real numbers. Denote by Δ_k the time interval $[t_{k-1}, t_k]$. For continuous functions $x^k : [t_0, t_\nu] \to \mathbb{R}^n$, $k \in \{1, \dots, \nu\}$, define the vector

$$p = \big(t_0, (t_1, x^1(t_0), x^1(t_1)), \dots, (t_\nu, x^\nu(t_{\nu-1}), x^\nu(t_\nu))\big) \in \mathbb{R}^d,$$

where $d = 1 + (2n+1)\nu$. Let $(f_k)_{k\in\{1,\dots,\nu\}}$ be smooth vector fields on $\mathbb{R}^n$, and $(\phi_i)_{i\in\{1,\dots,m\}}$, $(\eta_j)_{i\in\{1,\dots,q\}}$ be two families of smooth functions defined on $\mathbb{R}^{(\nu+1)(n+1)}$. For $t \in [t_0, t_\nu]$ and a collection $(U_k)_{k\in\{1,\dots,\nu\}}$ of subsets of $\mathbb{R}^q$, $q \geq 1$, consider the autonomous hybrid OCP

$$\begin{cases} \textit{minimize } t_\nu - t_0, \\ \dot{x^k}(t) = f_k(x^k(t), u^k(t)), \\ u^k(t) \in U_k, \\ t \in \Delta_k, \\ k \in \{1, \dots, \nu\}, \\ \eta_j(p) = 0,\ j = 1, \dots, q, \\ \phi_i(p) \leq 0,\ i = 1, \dots, m. \end{cases} \tag{HOCP}$$

Definition A.3.5. *For a tuple $w = (t_0; t_k, x^k(t), u^k(t), k = 1, \dots, \nu)$ which is extremal for Problem* (HOCP)*, define:*

- *the trajectory $(x(t))_{t\in[t_0,t_\nu]}$ which is equal to $x^k(t)$ for every $t \in \Delta_k \setminus \{t_k\}$ and $k \in \{1, \dots, \nu\}$;*

- *the adjoint trajectory* $(\lambda(t))_{t\in[t_0,t_\nu]}$ *which is equal to* $\lambda^k(t)$ *for every* $t \in \Delta_k \backslash \{t_k\}$ *and* $k \in \{1,\dots,\nu\}$*;*
- *the control* $(u(t))_{t\in[t_0,t_\nu]}$ *which is equal to* $u^k(t)$ *for every* $t \in \Delta_k \setminus \{t_{k-1}\}$ *and* $k \in \{1,\dots,\nu\}$.

Define, for every $k \in \{1,\dots,\nu\}$, $t \in \Delta_k$,

$$H^k(x^k,\lambda^k,\lambda_0,u^k) = \langle \lambda^k, f_k(x^k,u^k)\rangle - \lambda_0.$$

Theorem A.3.1. *Assume that* $(\widetilde{x}(\cdot),\widetilde{u}(\cdot),\widetilde{p})$ *is an optimal solution of Problem* (HOCP)*. Then there exists*

$$(\alpha,\beta,\lambda(\cdot),\lambda_0),$$

where $\alpha = (\alpha_1,\dots,\alpha_m) \in \mathbb{R}^m$*,* $\beta = (\beta_1,\dots,\beta_q) \in \mathbb{R}^q$*,* $\lambda = (\lambda^1,\dots,\lambda^\nu)$*, all* $\lambda^k : \Delta_k \to \mathbb{R}^n$ *for* $k \in \{1,\dots,\nu\}$ *being Lipschitz functions, and a constant* $\lambda_0 \geq 0$ *such that:*

- $(\lambda_0,\alpha,\beta) \neq 0$*;*
- *For every* $i \in \{1,\dots,m\}$*,* $\alpha_i \geq 0$*;*
- *For every* $i \in \{1,\dots,m\}$*,* $\alpha_i\phi_i(\widetilde{p}) = 0$*;*
- *For almost every* $t \in \Delta_k$*,*

$$\begin{aligned}
\dot{x}^k &= \frac{\partial H^k}{\partial \lambda}(x^k,\lambda^k,\lambda_0,\widetilde{u}),\\
\dot{\lambda}^k &= -\frac{\partial H^k}{\partial x^k}(x^k,\lambda^k,\lambda_0,\widetilde{u}),\\
H^k(x^k,\lambda^k,\lambda_0,\widetilde{u}) &= \max_{u\in\Omega} H^k(x^k,\lambda^k,\lambda_0,u) = 0.
\end{aligned} \tag{E}$$

Moreover, if we define $L(p) = \lambda_0(t_\nu - t_0) + \sum_{i=1}^m \alpha_i\phi_i(p) + \sum_{j=1}^q \beta_j\eta_j(p)$*, then we have the following transversality and discontinuity conditions at times* $t = t_0,\dots,t_\nu$*:*

- *At the initial and final times* t_0 *and* t_ν*, we have*

$$\begin{aligned}
\lambda^1(t_0) &= \frac{\partial L}{\partial x^1(t_0)}(\widetilde{p}),\\
\lambda^\nu(t_\nu) &= \frac{\partial L}{\partial x^\nu(t_\nu)}(\widetilde{p}).
\end{aligned}$$

- *At the crossing times* $(t_k)_{k\in\{1,\dots,\nu-1\}}$*, we have, for every* $k \in \{1,\dots,\nu-1\}$*,*

$$\lambda^k(t_{k-1}) = \frac{\partial L}{\partial x^k(t_{k-1})}(\widetilde{p}),$$
$$\lambda^k(t_k) = -\frac{\partial L}{\partial x^k(t_k)}(\widetilde{p}).$$

A.4 Main results

We are interested in solving (OCP) among continuous B-admissible trajectories, as defined in Section A.3.3. We first observe that the regular domain B_{11} is repulsive, and so any B-admissible trajectory with $x(0) \in B_{10}$ and $x(t_f) \in B_{01}$ should pass through B_{00}, as the point (θ_1, θ_2) cannot be reached from B_{10}. Thus, we fix the sequence of regular domains $B = (B_{10}, B_{00}, B_{01})$, with M_0 restricted to a point in K, and $M_1 = \{(x_1, x_2) \in K \mid x_1 \in [0, \theta_1),\ x_2 = x_2^f\}$ which has already been proven to be reachable in finite time, verifying assumptions A.3.3 and A.3.4. As previously said, the problem can be further analyzed by applying HMP. The Maximum Principle in the Hybrid framework requires to define functions $(\phi_i)_i$ and $(\eta_j)_j$ that guarantee the continuity of the trajectories and the changes of dynamics at the frontiers $x_1 = \theta_1$ and $x_2 = \theta_2$ [132]. Through its application, we obtain that (OCP) admits an optimal control which can be defined as a very simple feedback.

Theorem A.4.1. *The optimal strategy* $u(x)$ *solution of* (OCP) *for* B*-admissible trajectories is the feedback control*

$$u(x) = \begin{cases} u_{min} & \text{if } x \in (S_{u_{max}})^-, \\ u_{max} & \text{if } x \in (S_{u_{max}})^+ \cup (S_{u_{max}}). \end{cases}$$

Note that $u(x)$ is not defined in B_{11} due to the lack of control in the region. Figure A.4 illustrates the resulting vector field of (S) with the latter time-optimal control law. As a consequence of the latter theorem, the solutions of (OCP) for B-admissible trajectories are such that:

- the optimal control consists of two bang arcs $u \equiv u_{min}$ and $u \equiv u_{max}$, similar to the suboptimal control (A.4), with the switching between them occurring at the time when the trajectory reaches the separatrix $(S_{u_{\max}})$;

- the optimal trajectories passes by the unstable Filippov equilibrium (θ_1, θ_2),

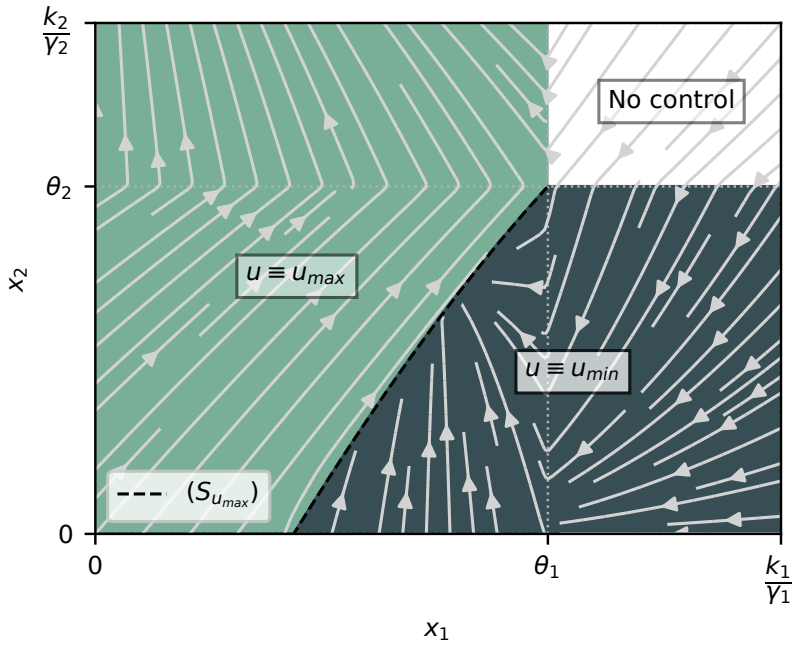


Figure A.4: Stream plot of the controlled dynamics (S) with the feedback control of Theorem A.4.1. System parameters are $\gamma_1 = 1.1$, $\gamma_2 = 1.7$, $\theta_1 = 0.6$, $\theta_2 = 0.4$, $k_1 = k_2 = 1$. Control bounds are set to $u_{min} = 0.5$ and $u_{max} = 1.5$.

> which is reached by its stable manifold corresponding to dynamics of Equation (S) with $u \equiv u_{\max}$. Then, the Filippov equilibrium is left by its unstable manifold corresponding to the dynamics of Equation (S) with $u \equiv u_{\max}$.

The proof of this result involves showing there are no singular arcs in the optimal control, and thus $u(t)$ can only be a concatenation of bang arcs. Additionally, because of the two-dimensional affine structure in each regular domain, the sign of the switching function in the Hamiltonian can switch at most once throughout the whole interval $[0, t_f]$. Consequently, the optimal control consists of at most two bang arcs (u_{min} or u_{max}), and the problem is reduced to finding the optimal switching time between the two arcs. An example of this trajectory and optimal control is shown in Figure A.5.

A.5 Proof of the main results

In this section, we provide the proof for Theorem A.4.1, which is organized as follows:

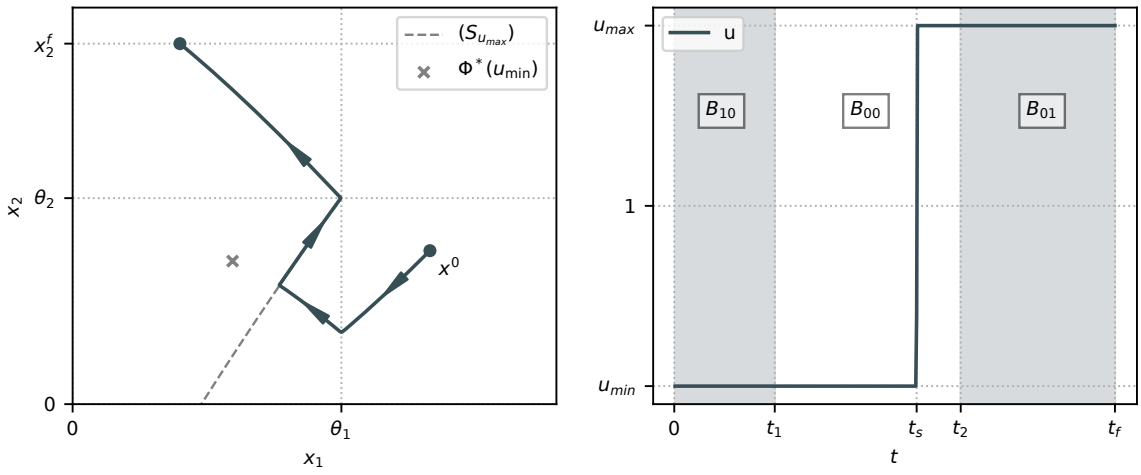


Figure A.5: Optimal trajectory with $x_1^0 = 0.8$, $x_2^0 = 0.3$ and $x_2^f = 0.7$. System parameters are $\gamma_1 = 1.2$, $\gamma_2 = 1.8$, $\theta_1 = 0.6$, $\theta_2 = 0.4$, and $k_1 = k_2 = 1$. Control bounds are set to $u_{min} = 0.5$ and $u_{max} = 1.5$. Times t_1 and t_2 are the transition times at which the state meets $x_1(t_1) = \theta_1$ and $x(t_2) = (\theta_1, \theta_2)$.

- In Section A.5.1, we reduce the problem to $B = (B_{00}, B_{01})$-admissible trajectories;
- In Section A.5.2, we prove that any optimal control admits no singular arcs in B_{00};
- In Section A.5.3, we show that the optimal control in B_{00} consists of two bang arcs with a switching time t_s such that $x(t_s) \in (S_{u_{max}})$, and we conclude the proof of Theorem A.4.1 showing that there are no singular arcs in B_{10}.

A.5.1 Reduction of the problem

Let $x(t)$ be the solution of Equation (S) such that $x(0) = x_0$ associated with an arbitrary control $u(t)$, and define the time at which the system crosses the frontier between B_{10} and B_{00} (respectively, between B_{00} and B_{01}) as t_1 (respectively, t_2). We notice that the time needed to achieve a transfer between the point $(x_1(t_2), x_2(t_2)) = (x_1', \theta_2)$ for $x_1' < \theta_1$, and the set $\{(x_1, x_2) \in K \mid x_2 = x_2^f,\ 0 \leq x_1 \leq \theta_1\}$ does not depend on $x_1' \leq \theta_1$, and by a direct property of Equation (S) restricted to B_{01}, we can easily prove that the optimal control strategy for Problem (OCP) is obtained when $u(t) = u_{\max}$ for $t \geq t_2$. Moreover, if $u(t) \in [u_{\min}, u_{\max}]$ is another optimal

control, then we obtain

$$\int_{t_2}^{t_f} e^{-\gamma_2 s}(u(s) - u_{\max})ds = 0,$$

hence $u(t) = u_{\max}$ for almost every $t \in [t_2, t_f]$. As a consequence, we can reduce the problem to solving (OCP) with $x_2^f = \theta_2$ among $B = (B_{10}, B_{00})$-admissible trajectories.

A.5.2 Absence of singular arcs in B_{00}

In order to apply Theorem A.3.1, given the choice $B = \{B_{10}, B_{00}\}$, we set $\nu = 2$, and we define the vector fields, for $x = (x_1, x_2) \in \mathbb{R}^2$, $u \in [u_{\min}, u_{\max}]$, by

$$f_1(x_1, x_2, u) = \begin{pmatrix} -\gamma_1 x_1 + uk_1 \\ -\gamma_2 x_2 \end{pmatrix}, \quad f_2(x_1, x_2, u) = \begin{pmatrix} -\gamma_1 x_1 + uk_1 \\ -\gamma_2 x_2 + uk_2 \end{pmatrix}.$$

The times where changes of regular domains occur for the dynamics are denoted by $t_0 = 0 < t_1$, and the final time is $t_2 = t_f$. Notice that $t_0 = 0$ is assumed to be fixed while t_1, t_f are not fixed quantities a priori. We introduce the following functions $(\eta_j)_{j\in\{1,\dots,7\}}$, which will guarantee the B-admissibility of the trajectories $x(t)$, which are solutions of Equation (S). In accordance with Definition A.3.5 of Section A.3.4, for a trajectory $x(t)$ which is solution of Equation (S), we define $p = (t_0, (t_1, x^1(t_0), x^1(t_1)), (t_f, x^2(t_1), x^2(t_f)))$. In order to guarantee the B-admissibility and the continuity of the trajectory $x(t)$ at $t = t_1$, we define the functions

$$\begin{cases} \eta_1(p) = t_0, \\ \eta_2(p) = x_1^1(t_0) - x_1^0, \\ \eta_3(p) = x_2^1(t_0) - x_2^0, \\ \eta_4(p) = x_1^1(t_1) - \theta_1, \\ \eta_5(p) = x_1^2(t_1) - \theta_1, \\ \eta_6(p) = x_2^1(t_1) - x_2^2(t_1), \\ \eta_7(p) = x_2^2(t_f) - \theta_2. \end{cases}$$

As in Theorem A.3.1, for $p = (t_0, (t_1, x^1(t_0), x^1(t_1)), (t_f, x^2(t_1), x^2(t_f)))$, $\alpha \in \mathbb{R}$, and $\beta = (\beta_1, \ldots, \beta_7) \in \mathbb{R}^7$, define the Lagrangian

$$L(p) = \alpha t_f + \sum_{j=1}^{7} \beta_j \eta_j(p).$$

For $k \in \{1, 2\}$, the Hamiltonian H^k defined in Theorem A.3.1 can be written as $H^k = H_0 + u^k H_1^k$, with $u^k \in [u_{\min}, u_{\max}]$ where, for every $x^k = (x_1^k, x_2^k) \in \mathbb{R}^2$ and $\lambda^k = (\lambda_1^k, \lambda_2^k)$,

$$\begin{aligned} H_0(x^k, \lambda^k, \lambda_0) &= -\gamma_1 x_1^k \lambda_1^k - \gamma_2 x_2^k \lambda_2^k - \lambda_0, \\ H_1^k(x^k, \lambda^k) &= \xi_1^k k_1 \lambda_1^k + \xi_2^k k_2 \lambda_2^k, \end{aligned}$$

with $\xi_1^1 = 0$, $\xi_1^2 = 1$, $\xi_2^1 = 1$, and $\xi_2^2 = 1$. In this setting, the Adjoint State Equation (E) of Theorem A.3.1 writes

$$\begin{cases} \dot{\lambda}_1^k = \gamma_1 \lambda_1^k, \\ \dot{\lambda}_2^k = \gamma_2 \lambda_2^k, \end{cases} \tag{AD}$$

which is independent of $k \in \{1, 2\}$. Then, we can derive conditions from Theorem A.3.1 concerning singular arcs of Equation (OCP) along B-admissible trajectories, as defined in Section A.3.3. For $k \in \{1, 2\}$, extremal singular arcs occur when the variables $(x^k(t), \lambda^k(t), \lambda_0, u^k(t))$ are extremal and satisfy

$$H_1^k(x^k(t), \lambda^k(t)) = 0, \tag{Sing}$$

for every $t \in [T_1, T_2]$, where $t_1 \leq T_1 < T_2 \leq t_f$. Along such trajectories, the vanishing condition of the k-th Hamiltonian H^k becomes

$$-\gamma_1 x_1^k(t) \lambda_1^k(t) - \gamma_2 x_2^k(t) \lambda_2^k(t) - \lambda_0 = 0, \tag{V}$$

for every $t \in [T_1, T_2]$. Define the B_{00} switching function as $\phi(t) = \text{sign}(k_1 \lambda_1^2(t) + k_2 \lambda_2^2(t))$ for $t \in [0, t_f]$. As a direct consequence of Theorem A.3.1, we have the following result.

Lemma A.5.1. *At times $t_0 = 0$, t_1 and t_f, we have the following transversality and*

discontinuity conditions:

$$\begin{cases} \beta_1 = -\alpha = \lambda_0, \\ \lambda_1^1(0) = \beta_2, \\ \lambda_2^1(0) = \beta_3, \\ \lambda_1^1(t_1) = -\beta_4, \\ \lambda_1^2(t_1) = \beta_5, \\ \lambda_2^1(t_1) = \lambda_2^2(t_1) = \beta_6, \\ \lambda_1^2(t_f) = 0, \\ \lambda_2^2(t_f) = -\beta_7. \end{cases} \tag{TD}$$

We can deduce the following property of extremal trajectories of Problem (OCP).

Lemma A.5.2. *Extremal trajectories of Problem (OCP) along B-admissible trajectories admit no singular arcs in B_{00}, that is, for $t \in [t_1, t_2]$.*

Proof. In this case, Condition (Sing) becomes

$$k_1\lambda_1^2(t) + k_2\lambda_2^2(t) = 0,$$

for $t \in [T_1, T_2]$. Differentiating this equality, we obtain

$$k_1\gamma_1\lambda_1^2(t) + k_2\gamma_2\lambda_2^2(t) = 0,$$

for $t \in [T_1, T_2]$. Then we get, for $t \in [T_1, T_2]$

$$\lambda_1^2(t)(\gamma_1 - \gamma_2) = 0.$$

Knowing that $\gamma_1 \neq \gamma_2$ by Assumption A.2.1, we obtain $\lambda_1^2(t) = 0$ for $t \in [T_1, T_2]$, and Condition (V) implies

$$\lambda_1^2(t) = \lambda_2^2(t) = \lambda_0 = 0.$$

Hence, by Equation (AD), we have $\lambda_1^2(t) = \lambda_2^2(t) = \lambda_0 = 0$ for every $t \in [t_1, t_f]$. Applying Theorem A.3.1 with the functions $(\eta_j)_{j\in\{1,\dots,7\}}$, and the Lagrangian L as defined as above, we see easily that the transversality and discontinuity conditions (TD) at times $t_0 = 0$, t_1 and t_f provide that $\alpha = \beta_j = 0$, for every $j \in \{1, \dots, 7\}$. Indeed, the condition $\lambda_1^2(t) = \lambda_2^2(t) = \lambda_0 = 0$ for every $t \in [t_1, t_f]$

implies that $\alpha = \beta_1 = \beta_5 = \beta_6 = \beta_7 = 0$. By Equation (AD), we get that $\lambda_2^1(t) = \lambda_2^2(t) = 0$ for every $t \in [0, t_f]$, so that we can deduce $\beta_3 = \beta_6 = 0$. The null Hamiltonian condition (V) then implies $\lambda_1^1(t) = 0$ for $t \in [0, t_1]$. It follows that $\beta_2 = \beta_4 = 0$, so that the nontriviality condition $(\alpha, \beta) \neq 0$ of Theorem A.3.1 is violated.

□

A.5.3 Optimality of the two bang arcs trajectory for Problem (OCP)

Because of the two-dimensional affine structure in each regular domain, the switching function ϕ can switch at most once throughout the whole interval $[0, t_f]$. By reachability considerations, we can deduce the following result.

Proposition A.5.3. *Extremal trajectories of Problem (OCP) along B-admissible trajectories are made of two bang arcs in the domain B_{00}, that is, there exists $t_s \geq t_1$ such that $u(t) = u_{min}$ for $t_1 \leq t \leq t_s$, and $u(t) = u_{max}$ for $t > t_s$.*

Proof. By Equation (AD) and Lemma A.5.2, the switching function ϕ switches at most once for $t \in [t_1, t_f]$. Moreover, in order to achieve a transfer between the lines $x_1^2(t_1) = \theta_1$ and $x_2^2(t_f) = \theta_2$, at least one switch is needed. Indeed, a constant control strategy $u \equiv u_{\min}$ is such that the associated solution $x(t)$ of Equation (S) converges towards $\Phi^*(u_{\min})$ when $t \to \infty$, where $\Phi^*(u_{\min})$ is defined as in Equation (A.3), so that $x(t) < \theta_2$ for every $t \geq t_1$. Moreover, a constant control strategy $u \equiv u_{\max}$ is such that $x(t) \in B_{10}$ for $t \geq t_1$, so that $x(t)$ is not B-admissible. □

Now we prove that that the switching time t_s defined in Proposition A.5.3 for optimal trajectories of Problem (OCP) along B-admissible trajectories is such that $(x_1^2(t_s), x_2^2(t_s)) \in (S_{u_{\max}})$. First notice that, by a direct study of Equation (S) restricted to the domain B_{00}, we can define $t^\star > 0$ as the unique time at which the solution $(y_1(t), y_2(t))$ of Equation (S) with $u \equiv u_{\min}$ and $y_1(0) = \theta_1$ and $y_2(0) = x_2^2(t_1)$ satisfies $(y_1(t^\star), y_2(t^\star)) \in (S_{u_{\max}})$. In order to guarantee the conditions $x_1^2(t_f) \in [0, \theta_1]$ and $x_2^2(t_f) = \theta_2$, the time t_s defined in Proposition A.5.3 has to satisfy $t_s \geq t_1 + t^\star$. We have the following result.

Lemma A.5.4. *Optimal trajectories of Problem (OCP) along B-admissible trajectories are such that $t_s = t_1 + t^\star$ and $x_1^2(t_f) = \theta_1$.*

Proof. We have for $t_1 \leq t \leq t_s$,

$$x_1^2(t) = (\theta_1 - \tfrac{k_1 u_{\min}}{\gamma_1})e^{-\gamma_1(t-t_1)} + \tfrac{k_1 u_{\min}}{\gamma_1},$$
$$x_2^2(t) = (x_2(t_1) - \tfrac{k_2 u_{\min}}{\gamma_2})e^{-\gamma_2(t-t_1)} + \tfrac{k_2 u_{\min}}{\gamma_2},$$

and for $t \geq t_s$ we have

$$x_1^2(t) = (x_1^2(t_s) - \tfrac{k_1 u_{\max}}{\gamma_1})e^{-\gamma_1(t-t_s)} + \tfrac{k_1 u_{\max}}{\gamma_1},$$
$$x_2^2(t) = (x_2^2(t_s) - \tfrac{k_2 u_{\max}}{\gamma_2})e^{-\gamma_2(t-t_s)} + \tfrac{k_2 u_{\max}}{\gamma_2}.$$

Notice that the condition $x_2^2(t_1) < \theta_2$ implies $x_2^2(t_s) < \theta_2$ and $x_1^2(t_s) < \theta_1$. By direct computations, a time $T \geq t_s$ satisfies $x_2^2(T) = \theta_2$ if and only if

$$T = T(t_s) \equiv t_s + \frac{1}{\gamma_2} \ln \left(\frac{-x_2^2(t_s) + k_2 u_{\max}/\gamma_2}{-\theta_2 + k_2 u_{\max}/\gamma_2} \right).$$

Notice that the condition $\theta_2 < \frac{u_{\max}}{\gamma_2}$ implies we can define a positive function $T : t_s \mapsto T(t_s)$. Moreover, one can prove that, for every $t_s > 0$,

$$T'(t_s) = \frac{k_2(u_{\max} - u_{\min})}{-\gamma_2 x_2^2(t_s) + k_2 u_{\max}}.$$

Using the fact that $(x_1^2(t_s), x_2^2(t_s))$ belongs to B_{00}, we obtain that T is increasing on $\mathbb{R}^+$, and reaches its minimum in the interval $[t_1 + t^\star, +\infty)$ at $t_s = t_1 + t^\star$. The result follows from the definitions of $t^\star$ and $(S_{u_{\max}})$ (see Section A.2.2). □

There remains to understand the structure of an optimal trajectory in the regular domain B_{10}, that is, when $t \leq t_1$. In the next proposition, we eliminate the possibility of having singular arcs in B_{10} by a direct study of the dynamics of Equation (S) associated with the application of Lemma A.5.4.

Proposition A.5.5. *Optimal trajectories of Problem (OCP) along B-admissible trajectories admit no singular arc in B_{10}.*

Proof. Consider the solution $\bar{x}(t) = (\bar{x}_1(t), \bar{x}_2(t))$ of Equation (S) such that $u \equiv u_{\min}$ while $\bar{x}(t) \in B_{10}$, $u \equiv u_{\min}$ while $\bar{x}(t) \in B_{00} \cap (S_{u_{\max}})^-$, $u \equiv u_{\max}$ while $\bar{x}(t) \in (S_{u_{\max}})$, and the solution $\widetilde{x}(t) = (\widetilde{x}_1(t), \widetilde{x}_2(t))$ of Equation (S) such that $u \equiv \widetilde{u}(t)$ while $\widetilde{x}(t) \in B_{10}$ for an arbitrary control $t \mapsto \widetilde{u}(t) \in [u_{\min}, u_{\max}]$, $u \equiv u_{\min}$ while $\widetilde{x}(t) \in B_{00} \cap (S_{u_{\max}})^-$, $u \equiv u_{\max}$ while $\widetilde{x}(t) \in (S_{u_{\max}})$, with same initial conditions. Hence

we can define the time $\bar{T} > 0$ (respectively, $\widetilde{T}$) at which we have $\bar{x}(\bar{T}) = (\theta_1, \theta_2)$ (respectively, $\widetilde{x}(\widetilde{T}) = (\theta_1, \theta_2)$).In order to prove that $\widetilde{T} \geq \bar{T}$, let us first consider the time $\widetilde{t}_1 > 0$ (respectively, $\bar{t}_1 > 0$) at which $\widetilde{x}(t)$ (respectively,$\bar{x}(t)$) reaches the frontier between B_{10} and B_{00}. By a direct property of Equation (S) restricted to the domain B_{10}, we have $\widetilde{x}_1(t) \geq \bar{x}_1(t)$ and $\widetilde{x}_2(t) = \bar{x}_2(t)$ for every $t \leq \min(\widetilde{t}_1, \bar{t}_1)$. It follows that $\bar{t}_1 \leq \widetilde{t}_1$ and $\widetilde{x}_2(\widetilde{t}_1) \leq \bar{x}_2(\bar{t}_1)$. Now consider a solution $\widetilde{y}(t) = (\widetilde{y}_1(t), \widetilde{y}_2(t))$ (respectively, $\bar{y}(t) = (\bar{y}_1(t), \bar{y}_2(t))$) of Equation (S) with $u \equiv u_{\min}$ and such that $\widetilde{y}_1(0) = \bar{y}_1(0) = \theta_1$, $\widetilde{y}_2(0) = \widetilde{x}_2(\widetilde{t}_1)$ and $\bar{y}_2(0) = \bar{x}_2(\bar{t}_1)$. By a direct property of Equation (S), the times $\widetilde{s}_1$ (respectively, $\bar{s}_1$) at which $\widetilde{y}(t)$ (respectively,$\bar{y}(t)$) reaches $(S_{u_{\max}})$ are such that $\widetilde{s}_1 \geq \bar{s}_1$. Hence we can deduce $\widetilde{T} = \widetilde{t}_1 + \widetilde{s}_1 \geq \bar{T} = \bar{t}_1 + \bar{s}_1$, and the structure of optimal trajectories in B_{00} given by Lemma A.5.4 allows to prove that an optimal trajectory for Problem (OCP) is such that $u(t) = u_{\min}$ for almost every $t \in [0, t_1]$. In particular, optimal trajectories for Problem (OCP) have no singular arcs in B_{10}. □

The latter proposition concludes the proof of Theorem A.4.1. Additionally, as a direct consequence of the previous results, we obtain that for $t \in [0, t_f]$, the optimal control is made of a first bang arc with $u \equiv u_{\min}$ towards $(S_{u_{\max}})$ for $t \in [0, t_s]$, then a second bang arc with $u \equiv u_{\max}$ for $t \in [t_s, t_f]$ so that the system follows $(S_{u_{\max}})$ until reaching (θ_1, θ_2).

Remark A.5.6. *By a direct analysis of the dynamics of Equation* (S) *in the regular domain B_{00}, one can show that the time $t^\star = t_s - t_1 > 0$ is the unique non-negative solution of the equation*

$$\left(\frac{k_2 u_{min}}{\gamma_2} - x_2^2(t_1)\right) e^{-\gamma_2 t} + k_2 \frac{u_{max} - u_{min}}{\gamma_2} \tag{EQ}$$
$$= \frac{\left(\frac{k_2 u_{max}}{\gamma_2} - \theta_2\right)}{\left(\frac{k_1 u_{max}}{\gamma_1} - \theta_1\right)^{\gamma_2/\gamma_1}} \left(\left(\frac{k_1 u_{min}}{\gamma_1} - \theta_1\right) e^{-\gamma_1 t} + k_1 \frac{u_{max} - u_{min}}{\gamma_1}\right)^{\gamma_2/\gamma_1}.$$

The latter can be obtained by solving $y_2(t^\star) = \alpha(y_1(t^\star), u_{max})$, with $y_1(t^\star) = x_1^2(t_1 + t^\star)$ and $y_2(t^\star) = x_2^2(t_1 + t^\star)$. Equation (EQ) *is hard to solve explicitly in the general case where Assumption A.2.1 is satisfied, especially because the latter assumption implies $\gamma_1 \neq \gamma_2$.*

A.6 Lower bound on the minimal time

The time required to perform a transition can be minimized to a certain extent, which is imposed by the dynamics of the system and the choice of control bounds, as shown in previous sections. In this section, we show there exist a lower bound to the minimal time. However, an explicit computation requires to solve Equation (EQ) analytically, which is a challenging task. In this section, we give a lower bound on the minimal time of Problem (OCP) in Proposition A.6.3, which is uniform w.r.t. $[u_{\min}, u_{\max}) \subset [0, +\infty)$ and is a function of the parameters $(\gamma_j)_{j\in\{1,2\}}$, $(k_j)_{j\in\{1,2\}}$, $(\theta_j)_{j\in\{1,2\}}$ satisfying Assumption A.2.1. In this purpose, we introduce an additional system which provides a lower bound for Problem (OCP). Let $[u_{\min}, u_{\max}) \subset [0, +\infty)$ be such that Assumption A.3.1 is satisfied. Then, for every $u_{\min}, u_{\max} \geq 0$ such that $[\bar{u}_{\min}, \bar{u}_{\max}] \subset [u_{\min}, u_{\max}]$, we have $\Phi^\star(u_{\min}) \in (S_{u_{\max}})^-$. Hence, the optimal control strategy for Problem (OCP) associated with such values of $u_{\min}, u_{\max}$ is given by Theorem A.4.1.

Definition A.6.1. *Define the* lower trajectories *as the solutions of*

$$\begin{aligned} \dot{z}_1 &= -\gamma_1 z_1 + k_1 u(t) s^-(z_2, \theta_2) \\ \dot{z}_2 &= -\gamma_1 z_2 + k_2 u(t) s^-(z_1, \theta_1), \end{aligned} \tag{$\widetilde{S}$}$$

with $u(t) \equiv u_{min}$ for $z(t) \in (\widetilde{S}_{u_{max}})^-$, and $u(t) \equiv u_{max}$ for $z(t) \in (\widetilde{S}_{u_{max}})^+ \cup \widetilde{S}_{u_{max}}$, where $(\widetilde{S}_{u_{max}})$ is defined as in Definition A.2.3.

A direct application of Lemma A.2.3 proves that if a lower trajectory $z(t)$ is such that $z(0) \in (\widetilde{S}_{u_{\max}})^-$, then $z(t)$ reaches $(\widetilde{S}_{u_{\max}})$ in finite time $T_{\text{low}}(u_{\min}, u_{\max})$. Moreover, as a consequence of the condition $\gamma_2 > \gamma_1$, we obtain the following lemma.

Lemma A.6.2. *Consider the solution $x(t)$ of Equation* (S) *such that $x(0) = x_0 \in (S_{u_{max}})^-$, where u is defined as in Theorem A.4.1, and the solution $z(t)$ of Equation $(\widetilde{S})$ such that $z(0) = x_0$. Then we have $x_1(t) \leq z_1(t)$ and $x_2(t) \leq z_2(t)$, for every $t \in [0, T_{low}(u_{min}, u_{max})]$.*

As a direct consequence, we get that the time t_s needed by $x(t)$ in order to reach $(S_{u_{\max}})$ (defined as in Proposition A.5.3) is such that $t_s \geq T_{\text{low}}(u_{\min}, u_{\max})$.

Hence, if we denote the minimal time for Problem (OCP) by $T_f(u_{\min}, u_{\max})$, then we have $T_f(u_{\min}, u_{\max}) \geq T_{\text{low}}(u_{\min}, u_{\max})$, for every $u_{\min}, u_{\max}$ be such that $0 \leq$

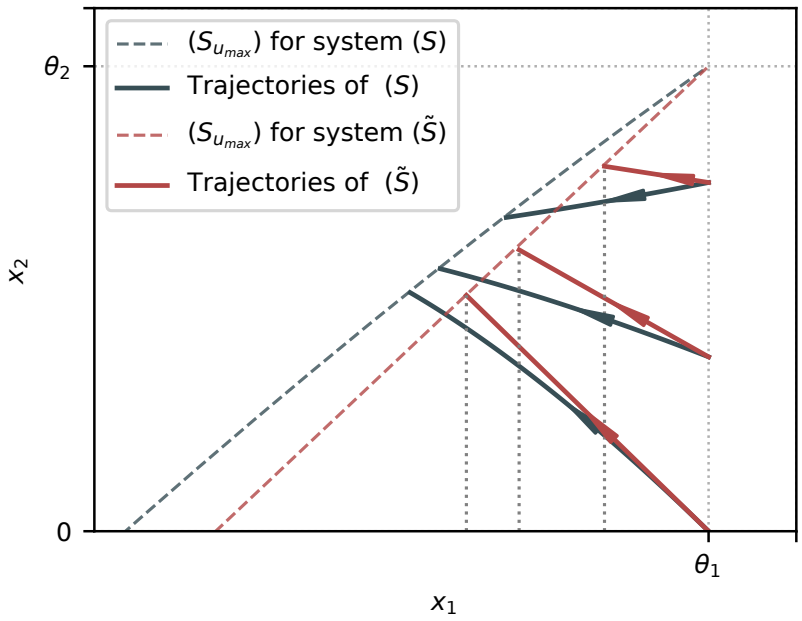


Figure A.6: Different trajectories starting from $x_1 = \theta_1$ with fixed control. System parameters are $\gamma_1 = 1.4$, $\gamma_2 = 2$, $\theta_1 = 0.6$, $\theta_2 = 0.4$, and $k_1 = k_2 = 1$. Control is set to $u \equiv u_{min}$ with $u_{min} = 0.5$. Trajectories of $(\widetilde{S})$ reach its associated separatrix at the lower-bound time T_{low}. Vertical lines at the interception indicate $x_1(T_{\text{low}})$.

$u_{\min} \leq u_{\max}$. Furthermore, by definition of Problem (OCP), we have $T_f(u_{\min}, u_{\max}) \geq T_f(\widetilde{u}_{\min}, \widetilde{u}_{\max})$, for every $u_{\min}, u_{\max}, \widetilde{u}_{\min}, \widetilde{u}_{\max}$ such that $0 \leq \widetilde{u}_{\min} \leq u_{\min} \leq u_{\max} \leq \widetilde{u}_{\max}$. It follows that for such a choice of $u_{\min}, u_{\max}, \widetilde{u}_{\min}, \widetilde{u}_{\max}$, we have

$$T_f(u_{\min}, u_{\max}) \geq T_{\text{low}}(\widetilde{u}_{\min}, \widetilde{u}_{\max}). \tag{A.5}$$

Proposition A.6.3. *Set $x_0 = (x_1^0, x_2^0) \in \bar{B}_{10}$ such that $x_2^0 < \theta_2$, and consider u_{min}, u_{max} such that $0 \leq u_{min} \leq u_{max}$. Let $x(t)$ be the solution of Equation* (S) *such that $x(0) = x_0$, where u is defined as in Theorem A.4.1. Then we have* $T_f(u_{min}, u_{max}) \geq -\frac{1}{\gamma_1} \ln\left(\frac{\theta_1 k_2 - \theta_2 k_1}{\theta_1 k_2 - x_2^0 k_1}\right) > 0$.

Proof. First assume that $x_1^0 = \theta_1$. Then by an adaptation of the formula given in Remark A.5.6, replacing γ_2 by γ_1, the time $T_{\text{low}}(u_{\min}, u_{\max})$ needed by the lower trajectory $z(t)$ to reach $(\widetilde{S}_{u_{\max}})$ is

$$T_{\text{low}}(u_{\min}, u_{\max}) = -\frac{1}{\gamma_1} \ln\left(\frac{(A(u_{\max})k_1 - k_2)(u_{\max} - u_{\min})}{\gamma_1\left(\left(\frac{k_2 u_{\min}}{\gamma_1} - x_2^0\right) - A(u_{\max})\left(\frac{k_1 u_{\min}}{\gamma_1} - \theta_1\right)\right)}\right),$$

where $A(u_{\max}) = \frac{\frac{k_2 u_{\max}}{\gamma_1} - \theta_2}{\frac{k_1 u_{\max}}{\gamma_1} - \theta_1}$. For every $u_{\min}, u_{\max}, \widetilde{u}_{\max}$ such that $0 \leq u_{\min} \leq u_{\max} \leq \widetilde{u}_{\max}$, Inequality (A.5) provides

$$T_f(u_{\min}, u_{\max}) \geq T_{\text{low}}(0, \widetilde{u}_{\max}) = -\frac{1}{\gamma_1} \ln \left(\frac{(A(\widetilde{u}_{\max})k_1 - k_2)\widetilde{u}_{\max}}{\gamma_1 \left(-x_2^0 + A(\widetilde{u}_{\max})\theta_1\right)} \right).$$

Noticing that $T_{\text{low}}(0, \widetilde{u}_{\max}) \to -\frac{1}{\gamma_1} \ln \left(\frac{\theta_1 k_2 - \theta_2 k_1}{\theta_1 k_2 - x_2^0 k_1} \right)$ when $\widetilde{u}_{\max} \to +\infty$, we deduce that $T_f(u_{\min}, u_{\max}) \geq -\frac{1}{\gamma_1} \ln \left(\frac{\theta_1 k_2 - \theta_2 k_1}{\theta_1 k_2 - x_2^0 k_1} \right)$, for every $u_{\min}, u_{\max}$ such that $0 \leq u_{\min} \leq u_{\max}$. Moreover, Assumption A.2.1 and the condition $x_2^0 < \theta_2$ guarantee that $-\frac{1}{\gamma_1} \ln \left(\frac{\theta_1 k_2 - \theta_2 k_1}{\theta_1 k_2 - x_2^0 k_1} \right) > 0$. We deduce the general case $x_1^0 \geq \theta_1$ noticing that we have in this case $x_2(t_1) \leq x_2^0$, where $t_1 \geq 0$ is the time where $x(t)$ changes regular domain from B_{10} to B_{00}, then applying the case $x_1^0 = \theta_1$.

□

A.7 Numerical results

We illustrate our results with numerical simulations performed with Bocop [4], an open-source toolbox for solving OCPs. In order to guarantee the reproducibility of the numerical results, the computations can be executed through an online version of Bocop[3]. The original problem (OCP) is solved through a direct method, by approximating it by a finite dimensional optimization problem, using a Lobato time discretization method. As the algorithm requires s^- to be regularized to a smooth function, we define, for $x \in \mathbb{R}$ and $k \in \mathbb{N}$, the Hill function

$$\delta(x_i, \theta_i, k) = \frac{\theta_i^k}{x_i^k + \theta_i^k}, \tag{A.6}$$

which can approximate s^- for large values of k and, when $k \to \infty$, it verifies

$$\lim_{k \to \infty} \delta(x_i, \theta_i, k) = \begin{cases} 1 & x_i < \theta_i, \\ 0 & x_i > \theta_i, \\ 1/2 & x_i = \theta_i. \end{cases}$$

[3] `https://ct.gitlabpages.inria.fr/gallery/bistable/bistable.html`

Replacing s^- by Hill functions (A.6) in system (S) yields the non-hybrid system

$$\begin{cases} \dot{x}_1 = -\gamma_1 x_1 + u k_1 \delta(x_2, \theta_2, k), \\ \dot{x}_2 = -\gamma_2 x_2 + u k_2 \delta(x_1, \theta_1, k). \end{cases}$$

System parameters are fixed to $\gamma_1 = 1.2$, $\gamma_2 = 2$, $\theta_1 = 0.6$, $\theta_2 = 0.4$ and $k_1 = k_2 = 1$, which verify Assumption A.2.1; and control bounds are set to $u_{min} = 0.5$ and $u_{max} = 1.5$ satisfying Assumption A.3.1. The parameter k of the Hill function is set to $k = 500$, which proved an acceptable approximation of the s^- function. Figure A.5 shows an optimal trajectory representing the transition $(\text{high}, \text{low})$ to $(\text{low}, \text{high})$. In accorance with the analytical results, the optimal control is a bang-bang control: it consists of a first phase $[0, t_s]$ of low synthesis control u_{min} until x reaches the separatrix ($S_{u_{max}}$), followed by a phase $[t_2, t_f]$ of high synthesis control u_{max} until x_2 reaches x_2^f. As it is customary when solving OCPs with direct methods, the algorithm does not count on any *a priori* information of the structure of the optimal control. Yet, the obtained trajectory is in agreement with Theorem A.4.1, which confirms our theoretical results. Moreover, the solver is not restricted to consider only B-admissible trajectories, which suggests that the solution found in this work is optimal not only for Problem (OCP) along B-admissible trajectories but also for the general (OCP), without imposing the domain sequences. Figure A.7 shows different trajectories starting from $(S_{u_{max}})^+$ and $(S_{u_{max}})^-$. The streamplot represents the closed-loop dynamics for the optimal control defined in Theorem A.4.1. All trajectories starting in $(S_{u_{max}})^-$ approach asymptotically the point $\Phi^*(u_{\min})$ (denoted by a cross) until they reach the separatrix, point at which the state slides over it towards the Filippov equilibrium (θ_1, θ_2). The optimal control for trajectories starting in $(S_{u_{max}})^+$ consists in $u \equiv u_{max}$ for the whole interval $[0, t_f]$, and do not pass by the Filippov equilibrium.

Remark A.7.1. *As already mentioned in the introduction, the dynamics is not uniquely defined at the undifferentiated point (θ_1, θ_2), and the proposed solution is obtained by making a choice of dynamics at this point. Hence, concerning a biological implementation of our time-optimal strategy, it seems more reasonable to apply $u(t) \equiv u_{min}$ during a slightly longer time $\widetilde{t}_s = t_s + \epsilon$ with a small $\epsilon > 0$.*

In accordance with Remark A.7.1, one can be interested in comparing the sub-optimal control strategy given in Equation (A.4) with the optimal control given by

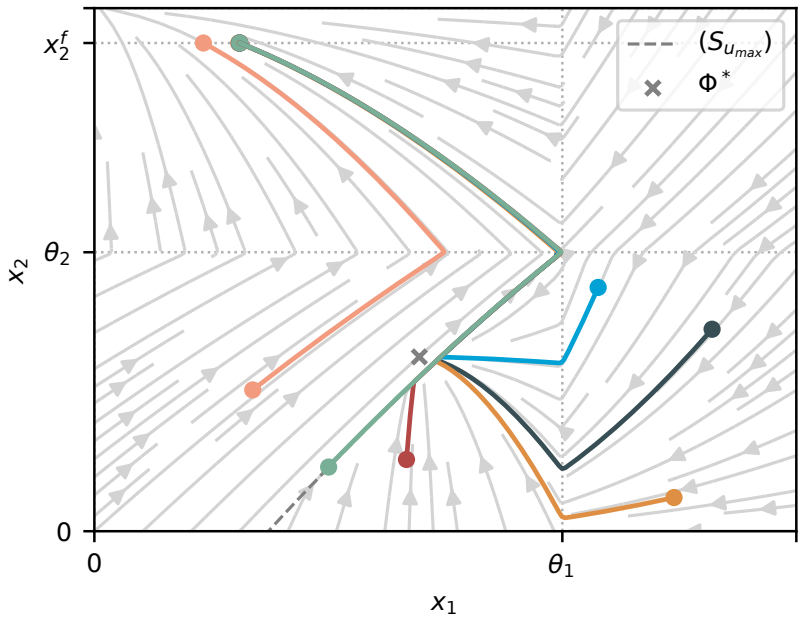


Figure A.7: Optimal trajectories starting from different initial points, with $x_2^f = 0.7$ and $k = 500$. The streamplot represents the vector field resulting from applying the optimal bang-bang strategy from Theorem A.4.1.

Theorem A.4.1. To this purpose, one can evaluate the time loss when delaying the switch by a time $\delta t > 0$, as $t_s = t_1 + t^\star + \delta t$, where t_1 and $t^\star$ are defined as in Section A.5.3 and depend on the parameters of the system. One can show by simple computations that the difference between the times needed to reach the target x_2^f for the modified trajectory w.r.t. the optimal trajectory is equal to

$$\frac{1}{\gamma_2} \ln \left(1 + e^{\gamma_2 (\widetilde{t} - t^*)} \left(e^{\gamma_2 \delta t} - 1 \right) \right),$$

where

$$\widetilde{t} = \frac{1}{\gamma_2} \ln \left(\frac{k_2 (u_{\min} - u_{\max})}{\gamma_2 \theta_2 - k_2 u_{\max}} \right).$$

A.7.1 Comparison with the smooth case

In order to explore the differences between the hybrid model studied in this paper and the smooth case, we obtained optimal trajectories for the continuous dynamical model given by (A.6) with lower Hill coefficients. Figure A.8 illustrates the impact of the Hill coefficient k in the functions $\delta(x, \theta, k)$, and how high values of k represent

a suitable approximation of the discrete case. Figure A.9 shows optimal transitions

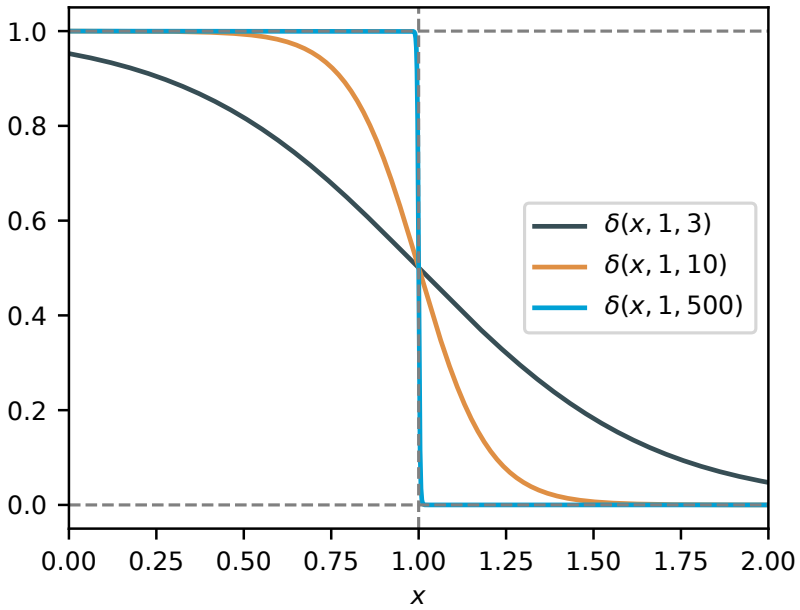


Figure A.8: Different Hill functions with different values of the Hill coefficient.

for $k = 3$ compared to the streamplot obtained from the optimal bang-bang strategy. We can observe that, even for lower values of k, the optimal control strategy remains bang-bang, with the switches being produced after the trajectories reach a certain region not necessarily delimited by $(S_{u_{max}})$. Finally, an important difference is that the bang-bang control does not yield trajectories passing by the unstable point (θ_1, θ_2) in the continuous case. In Figure A.10, three trajectories starting from the same initial conditions are compared for different values of k. We observe that, as k is reduced, the trajectory gets closer to the separatrix, and therefore, to the unstable point (θ_1, θ_2). Additionally, both the final time and the switching time are reduced as k increases towards the idealized hybrid case.

A.7.2 Supplementary condition $x_1(t_f) < x_1^{\mathbf{max}}$

In bistable systems, a binary switch implies taking the state towards the equilibria ϕ_{10} and ϕ_{01}. However, as stated in Section A.3, (OCP) represents a relaxed version of this problem where $x_1(t_f) > 0$, as it is not possible to control concentration x_1 in B_{01}. In order to compare the relaxed version with the original one, we investigate

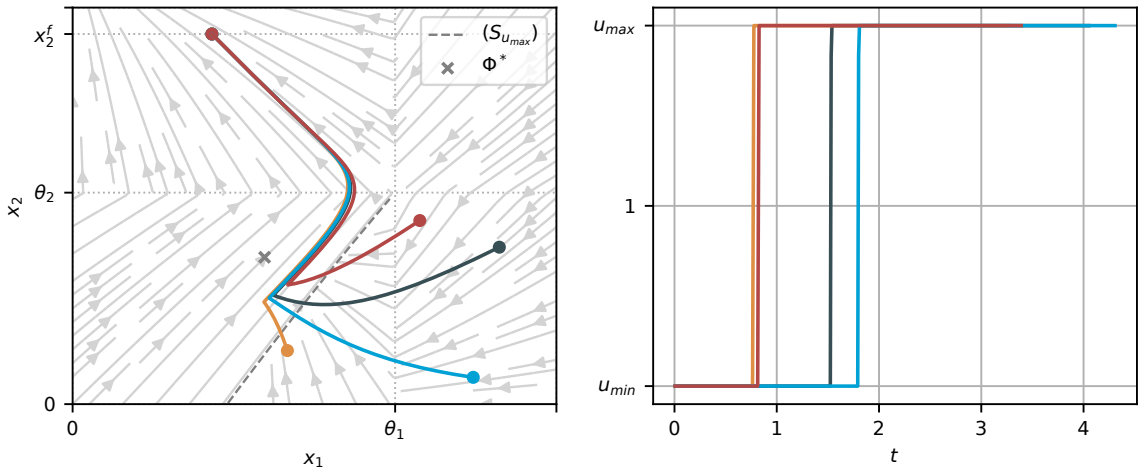


Figure A.9: Optimal trajectories starting from different initial points, with $x_2^f = 0.7$ and $k = 3$. The streamplot represents the vector field resulting from applying the optimal bang-bang strategy from Theorem A.4.1.

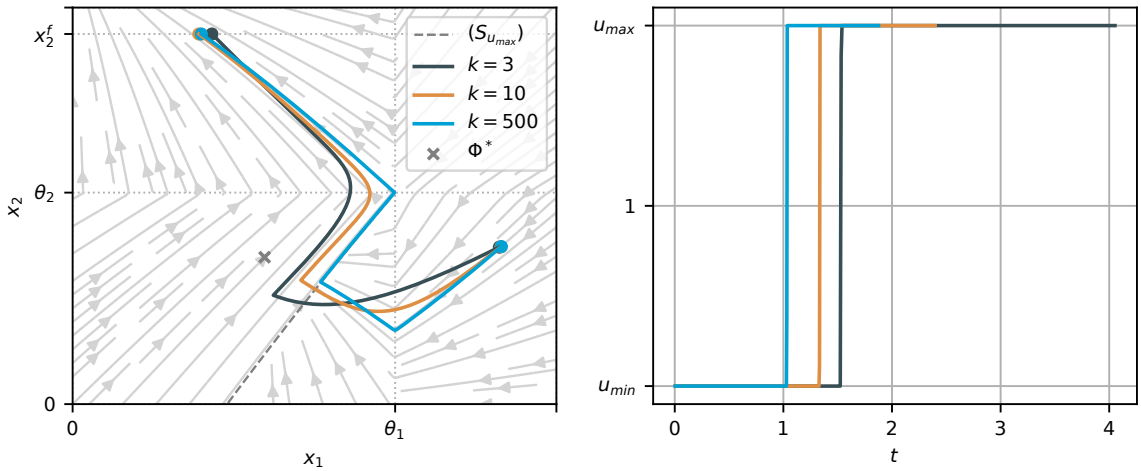


Figure A.10: Optimal trajectories starting from the initial point $(0.8, 0.3)$, with $x_2^f = 0.7$ and for different values of k. The streamplot represents the vector field resulting from applying the optimal bang-bang strategy from Theorem A.4.1.

numerically the following problem:

$$\begin{cases} \textit{minimize } t_f \geq 0 \\ x(t) = (x_1(t), x_2(t)) \textit{ is subject to } (\mathrm{S}), \\ x(0) = (x_1^0, x_2^0), \\ x_2(t_f) = x_2^f, \\ x_1(t_f) \in [0, x_1^{\max}), \\ u(\cdot) \in [u_{\min}, u_{\max}]. \end{cases} \tag{OCP_2}$$

In the particular case $x_2^f = k_2/\gamma_2$, solving Problem (OCP_2) allows to ensure that the state $x(t)$ is close enough to the steady state ϕ_{01} at time $t = t_f$. The main difference with (OCP) is that $x_1(t_f)$ is now constrained to the interval $[0, x_1^{\max}]$ with $x_1^{\max} < \theta_1$. For initial conditions in B_{01} given by $x_1^0 \in (x_1^{\max}, \theta_1]$ and $x_2^0 = \theta_2$, we notice that the time it takes for $x_1(t)$ to reach $x_1^{\max}$ does not depend on the control u (as there is no term depending on the control u in the dynamics of $x_1(t)$). Therefore, the final time t_f does not depend on the control, and so any control driving $x_2(t)$ from $x_2(0) = \theta_2$ to x_2^f in a time $t' \leq t_f$ is optimal for Problem (OCP_2). Thus, the problem has infinite solutions. Figure A.11 shows different trajectories for different values of $x_1^{\max}$. Among all infinite solutions, the ones found by Bocop depend on the initialization of the optimization algorithm, and have no particular meaning in the regular domain B_{01}. However, we verify that, as in (OCP), the switch in the control u occurs at the separatrix ($S_{u_{\max}}$), and then they follow the separatrix until the point (θ_1, θ_2). Thus, the simplest bang-bang strategy solution of (OCP_2) is

$$u_1(x) = \begin{cases} u_{min} & \text{if } x \in (S_{u_{max}})^-, \\ u_{max} & \text{if } x \in (S_{u_{max}})^+ \cup (S_{u_{max}}) \text{ and } x_2 < x_2^f, \\ \frac{\gamma_2}{k_2} x_2^f & \text{if } x_2 = x_2^f. \end{cases}$$

where the control $u \equiv x_2^f \gamma_2 / k_2$ is chosen so that $\dot{x}_2 = 0$ in the last phase. In the particular case where the final state is such that $x_2^f = k_2/\gamma_2$ (corresponding to the x_2-coordinate of the steady state ϕ_{01}), the optimal control in the last phase corresponds to the open loop system $u \equiv 1$.

A.8 Conclusion

This paper addressed the time-optimal control problem of a bistable gene-regulatory network. Through the application of HMP, we showed that any optimal control achieving state transition is a bang-bang control, where its value is a function of the state of the system (i.e. a feedback control). While in previous works [120], the bang-bang nature of the control is imposed as a constraint, we showed that such a characteristic is necessary to produce minimum-time transitions. Results also indicate that optimal trajectories should pass by the Filippov equilibrium (θ_1, θ_2), which represents the undifferentiated state, highly relevant from the biological point

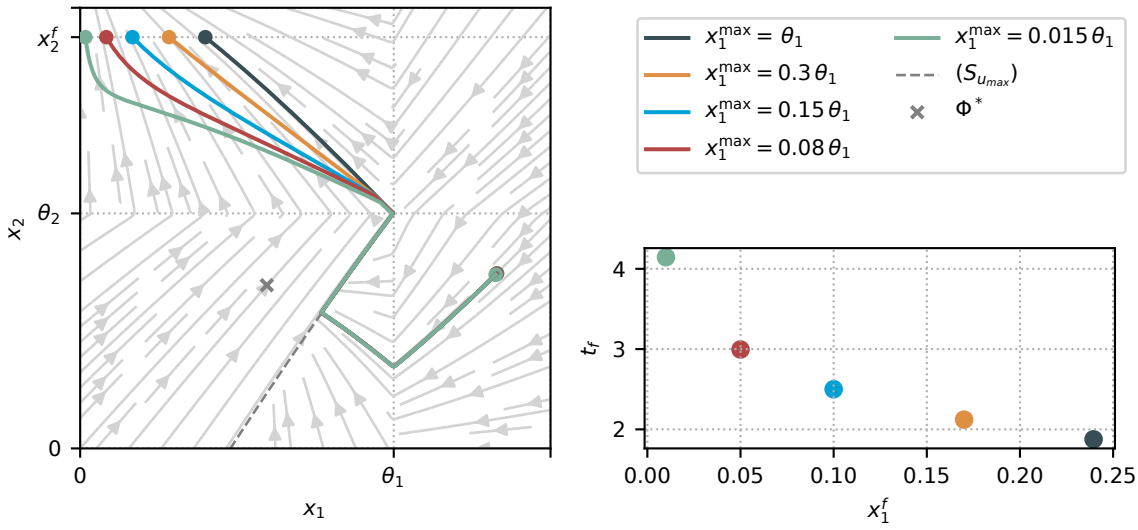


Figure A.11: Optimal trajectories obtained with Bocop starting from the same initial point $(0.8, 0.3)$, with $x_2^f = 0.7$ and for different values of $x_1^{\max}$. The streamplot represents the vector field resulting from applying the optimal bang-bang strategy from Theorem A.4.1. The first case (with $x_1^{\max} = \theta_1$) is the solution of (OCP).

of view. We showed the existence of a lower bound to the minimal time, by introducing the concept of *lower trajectories*. The numerical simulations obtained through direct methods confirm our analytical results, even when no prior knowledge of the structure of the optimal trajectories is specified. The latter are obtained by approximating the piecewise behavior of the systems with Hill functions, thus simulating a non-hybrid system. Additionally, the numerical results indicate that the trajectories found are optimal not only among B-admissible trajectories, but for all solutions of the hybrid system (S). Finally, we performed a numerical comparison of the trajectories obtained for the relaxed problem (i.e. with a constraint $x_1^f \leq \theta_1$) and those of the original one (i.e. with a constraint $x_1^f \leq x_1^{\max} < \theta_1$), which suggests that our results are also applicable to the original problem. Our work can be related to other results in the literature. For instance, in [127], an irreversible bistable switch in *E. coli* between the genes FadR and TetR is artificially engineered by replacing an endogenous negative autoregulation loop into a positive one. The control strategy is the feed-in of fatty acid, which is chosen to be bang-bang for simplicity. Our results supports such choice by proving it is not only simple, but also time optimal from a mathematical point of view. We expect that our result could be generalized to higher dimensional genetic regulatory networks, where it often occurs that trajec-

tories belonging to a given domain may bifurcate in different domains, similarly to what happens in the toggle switch case.

Bibliography

[1] Frederick C Neidhardt and Boris Magasanik. Studies on the role of ribonucleic acid in the growth of bacteria. *Biochimica et biophysica acta*, 42:99–116, 1960.

[2] Matthew Scott, Carl W Gunderson, Eduard M Mateescu, Zhongge Zhang, and Terence Hwa. Interdependence of cell growth and gene expression: origins and consequences. *Science*, 330(6007):1099–1102, 2010.

[3] Ivan Yegorov, Francis Mairet, Hidde De Jong, and Jean-Luc Gouzé. Optimal control of bacterial growth for the maximization of metabolite production. *Journal of mathematical biology*, 78(4):985–1032, 2019.

[4] Inria Saclay Team Commands. Bocop: an open source toolbox for optimal control. `http://bocop.org`, 2017.

[5] Jes Forchhammer and Lasse Lindahl. Growth rate of polypeptide chains as a function of the cell growth rate in a mutant of Escherichia coli 15. *Journal of molecular biology*, 55(3):563–568, 1971.

[6] HDPP Bremer, Patrick P Dennis, et al. Modulation of chemical composition and other parameters of the cell by growth rate. *Escherichia coli and Salmonella: cellular and molecular biology*, 2(2):1553–69, 1996.

[7] Liujie Huo, Joachim J Hug, Chengzhang Fu, Xiaoying Bian, Youming Zhang, and Rolf Müller. Heterologous expression of bacterial natural product biosynthetic pathways. *Natural Product Reports*, 36(10):1412–1436, 2019.

[8] Hidde De Jong, Stefano Casagranda, Nils Giordano, Eugenio Cinquemani, Delphine Ropers, Johannes Geiselmann, and Jean-Luc Gouzé. Mathematical modelling of microbes: metabolism, gene expression and growth. *Journal of the Royal Society Interface*, 14(136):20170502, 2017.

[9] Itzhak Fishov, Arieh Zaritsky, and NB Grover. On microbial states of growth. *Molecular microbiology*, 15(5):789–794, 1995.

[10] Jacques Monod. The growth of bacterial cultures. *Annual review of microbiology*, 3(1):371–394, 1949.

[11] Leonor Michaelis, Maud L Menten, et al. Die kinetik der invertinwirkung. *Biochem. z*, 49(333-369):352, 1913.

[12] Moselio Schaechter, Ole Maaløe, and Niels O Kjeldgaard. Dependency on medium and temperature of cell size and chemical composition during balanced growth of salmonella typhimurium. *Microbiology*, 19(3):592–606, 1958.

[13] Jeremy S Edwards, Rafael U Ibarra, and Bernhard O Palsson. In silico predictions of escherichia coli metabolic capabilities are consistent with experimental data. *Nature biotechnology*, 19(2):125–130, 2001.

[14] Rafael U Ibarra, Jeremy S Edwards, and Bernhard O Palsson. Escherichia coli K-12 undergoes adaptive evolution to achieve in silico predicted optimal growth. *Nature*, 420(6912):186, 2002.

[15] Erez Dekel and Uri Alon. Optimality and evolutionary tuning of the expression level of a protein. *Nature*, 436(7050):588–592, 2005.

[16] Andrea Y Weiße, Diego A Oyarzún, Vincent Danos, and Peter S Swain. Mechanistic links between cellular trade-offs, gene expression, and growth. *Proceedings of the National Academy of Sciences*, 112(9):E1038–E1047, 2015.

[17] Benjamin D Towbin, Yael Korem, Anat Bren, Shany Doron, Rotem Sorek, and Uri Alon. Optimality and sub-optimality in a bacterial growth law. *Nature communications*, 8(1):1–8, 2017.

[18] Julio R Banga. Optimization in computational systems biology. *BMC systems biology*, 2(1):1–7, 2008.

[19] Kenneth J Kauffman, Purusharth Prakash, and Jeremy S Edwards. Advances in flux balance analysis. *Current opinion in biotechnology*, 14(5):491–496, 2003.

[20] Anne Goelzer, Vincent Fromion, and Gérard Scorletti. Cell design in bacteria as a convex optimization problem. *Automatica*, 47(6):1210–1218, 2011.

[21] Evert Bosdriesz, Douwe Molenaar, Bas Teusink, and Frank J Bruggeman. How fast-growing bacteria robustly tune their ribosome concentration to approximate growth-rate maximization. *The FEBS journal*, 282(10):2029–2044, 2015.

[22] Oren Shoval, Hila Sheftel, Guy Shinar, Yuval Hart, Omer Ramote, Avi Mayo, Erez Dekel, Kathryn Kavanagh, and Uri Alon. Evolutionary trade-offs, pareto optimality, and the geometry of phenotype space. *Science*, 336(6085):1157–1160, 2012.

[23] M. Scott, S. Klumpp, E.M. Mateescu, and T. Hwa. Emergence of robust growth laws from optimal regulation of ribosome synthesis. *Molecular Systems Biology*, 10:747, 2014.

[24] Manlu Zhu and Xiongfeng Dai. Growth suppression by altered (p) ppgpp levels results from non-optimal resource allocation in escherichia coli. *Nucleic acids research*, 47(9):4684–4693, 2019.

[25] Moshe Sipper. Fifty years of research on self-replication: An overview. *Artificial life*, 4(3):237–257, 1998.

[26] Douwe Molenaar, Rogier Van Berlo, Dick De Ridder, and Bas Teusink. Shifts in growth strategies reflect tradeoffs in cellular economics. *Molecular systems biology*, 5(1):323, 2009.

[27] Tracey A Lincoln and Gerald F Joyce. Self-sustained replication of an rna enzyme. *Science*, 323(5918):1229–1232, 2009.

[28] Christoph Flamm, Lukas Endler, Stefan Müller, Stefanie Widder, and Peter Schuster. A minimal and self-consistent in silico cell model based on macromolecular interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1486):1831–1839, 2007.

[29] Nils Giordano, Francis Mairet, Jean-Luc Gouzé, Johannes Geiselmann, and Hidde De Jong. Dynamic optimisation of resource allocation in microorganisms. In *21st International Symposium on Mathematical Theory of Networks and Systems (MTNS 2014)*, 2014.

[30] Nils Giordano, Francis Mairet, Jean-Luc Gouzé, Johannes Geiselmann, and Hidde De Jong. Dynamical allocation of cellular resources as an optimal control

problem: novel insights into microbial growth strategies. *PLoS computational biology*, 12(3):e1004802, 2016.

[31] Lev Semenovich Pontryagin. *Mathematical theory of optimal processes*. Routledge, 2018.

[32] Steffen Waldherr and Henning Lindhorst. Optimality in cellular storage via the pontryagin maximum principle. *IFAC-PapersOnLine*, 50(1):9889–9895, 2017.

[33] David W Erickson, Severin J Schink, Vadim Patsalo, James R Williamson, Ulrich Gerland, and Terence Hwa. A global resource allocation strategy governs growth transition kinetics of Escherichia coli. *Nature*, 551(7678):119–123, 2017.

[34] Yael Korem Kohanim, Dikla Levi, Ghil Jona, Benjamin D Towbin, Anat Bren, and Uri Alon. A bacterial growth law out of steady state. *Cell reports*, 23(10):2891–2900, 2018.

[35] François Bertaux, Julius Von Kügelgen, Samuel Marguerat, and Vahid Shahrezaei. A bacterial size law revealed by a coarse-grained model of cell physiology. *PLoS computational biology*, 16(9):e1008245, 2020.

[36] Tae Seok Moon, Sang-Hwal Yoon, Amanda M Lanza, Joseph D Roy-Mayhew, and Kristala L Jones Prather. Production of glucaric acid from a synthetic pathway in recombinant escherichia coli. *Applied and environmental microbiology*, 75(3):589–595, 2009.

[37] Songyuan Li, Christian Bille Jendresen, and Alex Toftgaard Nielsen. Increasing production yield of tyrosine and mevalonate through inhibition of biomass formation. *Process Biochemistry*, 51(12):1992–2000, 2016.

[38] Tat-Ming Lo, Si Hui Chng, Wei Suong Teo, Han-Saem Cho, and Matthew Wook Chang. A two-layer gene circuit for decoupling cell growth from metabolite production. *Cell systems*, 3(2):133–143, 2016.

[39] Andreas Milias-Argeitis, Marc Rullan, Stephanie K Aoki, Peter Buchmann, and Mustafa Khammash. Automated optogenetic feedback control for precise and robust regulation of gene expression and cell growth. *Nature communications*, 7(1):1–11, 2016.

[40] Diego A Oyarzún and Guy-Bart V Stan. Synthetic gene circuits for metabolic control: design trade-offs and constraints. *Journal of the Royal Society Interface*, 10(78):20120671, 2013.

[41] Irene Otero-Muras, Ahmad A Mannan, Julio R Banga, and Diego A Oyarzún. Multiobjective optimization of gene circuits for metabolic engineering. *IFAC-PapersOnLine*, 52(26):13–16, 2019.

[42] Axel Nyström, Antonis Papachristodoulou, and Andrew Angel. A dynamic model of resource allocation in response to the presence of a synthetic construct. *ACS synthetic biology*, 7(5):1201–1210, 2018.

[43] Ivan Yegorov, Francis Mairet, Hidde De Jong, and Jean-Luc Gouzé. Optimal control of bacterial growth for the maximization of metabolite production. *Journal of mathematical biology*, pages 1–48, 2018.

[44] Jérôme Izard, Cindy DC Gomez Balderas, Delphine Ropers, Stephan Lacour, Xiaohu Song, Yifan Yang, Ariel B Lindner, Johannes Geiselmann, and Hidde De Jong. A synthetic growth switch based on controlled expression of RNA polymerase. *Molecular systems biology*, 11(11):840, 2015.

[45] Arthur L Koch. Why can't a cell grow infinitely fast? *Canadian journal of microbiology*, 34(4):421–426, 1988.

[46] Hal L Smith and Paul Waltman. *The Theory of the Chemostat: Dynamics of Microbial Competition*, volume 13. Cambridge university press, 1995.

[47] Mathukumalli Vidyasagar. Decomposition techniques for large-scale systems with nonadditive interactions: Stability and stabilizability. *IEEE transactions on automatic control*, 25(4):773–779, 1980.

[48] Joseph LaSalle. Some extensions of Liapunov's second method. *IRE Transactions on circuit theory*, 7(4):520–527, 1960.

[49] Bernard Bonnard and Monique Chyba. *Singular trajectories and their role in control theory*, volume 40. Springer Science & Business Media, 2003.

[50] Yacine Chitour, Frédéric Jean, and Emmanuel Trélat. Singular trajectories of control-affine systems. *SIAM Journal on Control and Optimization*, 47(2):1078–1095, 2008.

[51] VF Borisov. Fuller's phenomenon. *Journal of Mathematical Sciences*, 100(4):2311–2354, 2000.

[52] JF Bonnans, V Grelard, and P Martinon. Bocop, the optimal control solver, open source toolbox for optimal control problems. *URL http://bocop. org*, 2011.

[53] Andrei A Agrachev and Yuri Sachkov. *Control Theory from the Geometric Viewpoint*, volume 87. Springer Science & Business Media, 2013.

[54] Jan Dirk Van Elsas, Alexander V Semenov, Rodrigo Costa, and Jack T Trevors. Survival of escherichia coli in the environment: fundamental and public health aspects. *The ISME journal*, 5(2):173–183, 2011.

[55] J Shu and ML Shuler. A mathematical model for the growth of a single cell of e. coli on a glucose/glutamine/ammonium medium. *Biotechnology and bioengineering*, 33(9):1117–1126, 1989.

[56] David V Goeddel, Dennis G Kleid, Francisco Bolivar, Herbert L Heyneker, Daniel G Yansura, Roberto Crea, Tadaaki Hirose, Adam Kraszewski, Keiichi Itakura, and Arthur D Riggs. Expression in escherichia coli of chemically synthesized genes for human insulin. *Proceedings of the National Academy of Sciences*, 76(1):106–110, 1979.

[57] Jérôme Harmand, Claude Lobry, Alain Rapaport, and Tewfik Sari. *The chemostat: Mathematical theory of microorganism cultures*. John Wiley & Sons, 2017.

[58] Meike T Wortel, Evert Bosdriesz, Bas Teusink, and Frank J Bruggeman. Evolutionary pressures on microbial metabolic strategies in the chemostat. *Scientific reports*, 6:29503, 2016.

[59] David William Spitzer. Maximization of steady-state bacterial production in a chemostat with ph and substrate control. *Biotechnology and bioengineering*, 18(2):167–178, 1976.

[60] Rixa R Lichtl, Michael J Bazin, and David O Hall. The biotechnology of hydrogen production by nostoc flagelliforme grown under chemostat conditions. *Applied microbiology and biotechnology*, 47(6):701–707, 1997.

[61] Marc C D'anjou and Andrew J Daugulis. A rational approach to improving productivity in recombinant pichia pastoris fermentation. *Biotechnology and bioengineering*, 72(1):1–11, 2001.

[62] Térence Bayen and Francis Mairet. Optimization of the separation of two species in a chemostat. *Automatica*, 50(4):1243–1248, 2014.

[63] Carlos Martínez, Olivier Bernard, and Francis Mairet. Maximizing microalgae productivity in a light-limited chemostat. *IFAC-PapersOnLine*, 51(2):735–740, 2018.

[64] Antoine Haddon, Hector Ramirez, and Alain Rapaport. First results of optimal control of average biogas production for the chemostat over an infinite horizon. *IFAC-PapersOnLine*, 51(2):725–729, 2018.

[65] Horst R Thieme. Convergence results and a poincaré-bendixson trichotomy for asymptotically autonomous differential equations. *Journal of mathematical biology*, 30(7):755–763, 1992.

[66] Sylvain Goutelle, Michel Maurin, Florent Rougier, Xavier Barbaut, Laurent Bourguignon, Michel Ducher, and Pascal Maire. The hill equation: a review of its capabilities in pharmacological modelling. *Fundamental & clinical pharmacology*, 22(6):633–648, 2008.

[67] Kenneth A Johnson and Roger S Goody. The original Michaelis constant: translation of the 1913 Michaelis–Menten paper. *Biochemistry*, 50(39):8264–8269, 2011.

[68] Andrei A Agrachev and Yuri Sachkov. *Control theory from the geometric viewpoint*, volume 87. Springer Science & Business Media, 2013.

[69] Michail I Zelikin and Vladimir F Borisov. *Theory of Chattering Control: with Applications to Astronautics, Robotics, Economics, and Engineering*. Springer Science & Business Media, 2012.

[70] AJ Van der Schaft. Symmetries in optimal control. *SIAM journal on control and optimization*, 25(2):245–259, 1987.

[71] Emmanuel Trélat and Enrique Zuazua. The turnpike property in finite-dimensional nonlinear optimal control. *Journal of Differential Equations*, 258(1):81–114, 2015.

[72] Walid Djema, Laetitia Giraldi, Sofya Maslovskaya, and Olivier Bernard. Turnpike features in optimal selection of species represented by quota models. submitted.

[73] HM Robbins. A generalized legendre-clebsch condition for the singular cases of optimal control. *IBM Journal of Research and Development*, 11(4):361–372, 1967.

[74] E Cinquemani, F Mairet, I Yegorov, H de Jong, and J-L Gouzé. Optimal control of bacterial growth for metabolite production: The role of timing and costs of control. In *2019 18th European Control Conference (ECC)*, pages 2657–2662. IEEE, 2019.

[75] Ivan Yegorov, Francis Mairet, and Jean-Luc Gouzé. Optimal feedback strategies for bacterial growth with degradation, recycling, and effect of temperature. *Optimal Control Applications and Methods*, 39(2):1084–1109, 2018.

[76] S Joe Qin and Thomas A Badgwell. A survey of industrial model predictive control technology. *Control engineering practice*, 11(7):733–764, 2003.

[77] Erdal Aydin, Dominique Bonvin, and Kai Sundmacher. Nmpc using pontryagin's minimum principle-application to a two-phase semi-batch hydroformylation reactor under uncertainty. *Computers & Chemical Engineering*, 108:47–56, 2018.

[78] Abhishek S Soni and Robert S Parker. Closed-loop control of fed-batch bioreactors: A shrinking-horizon approach. *Industrial & engineering chemistry research*, 43(13):3381–3393, 2004.

[79] Erdal Aydin, Dominique Bonvin, and Kai Sundmacher. Computationally efficient nmpc for batch and semi-batch processes using parsimonious input parameterization. *Journal of Process Control*, 66:12–22, 2018.

[80] V. Raghuraman, V. Renganathan, T. H. Summers, and J. P. Koeln. Hierarchical mpc with coordinating terminal costs*. In *2020 American Control Conference (ACC)*, pages 4126–4133, 2020.

[81] R. Scattolini and P. Colaneri. Hierarchical model predictive control. In *2007 46th IEEE Conference on Decision and Control*, pages 4803–4808, 2007.

[82] A. L. Dontchev, I. V. Kolmanovsky, M. I. Krastanov, V. M. Veliov, and P. T. Vuong. Approximating optimal finite horizon feedback by model predictive control. *Systems Control Lett.*, 139:104666, 2020.

[83] Moselio Schaechter, John L Ingraham, and Frederick Carl Neidhardt. *Microbe*. ASM Press, Washington, DC, 2006.

[84] F.C. Neidhardt. Bacterial growth: constant obsession with dN/dt. *J. Bacteriol.*, 181(24):7405–7408, 1999.

[85] Stanley Brul, Suzanne Van Gerwen, and Marcel Zwietering. *Modelling Microorganisms in Food*. Elsevier, 2007.

[86] Asher Brauner, Ofer Fridman, Orit Gefen, and Nathalie Q Balaban. Distinguishing between resistance, tolerance and persistence to antibiotic treatment. *Nature Reviews Microbiology*, 14(5):320–330, 2016.

[87] James C Liao, Luo Mi, Sammy Pontrelli, and Shanshan Luo. Fuelling the future: microbial engineering for the production of sustainable biofuels. *Nature Reviews Microbiology*, 14(5):288–304, 2016.

[88] H.A. van den Berg, Y.N. Kiselev, and M.V. Orlov. Optimal allocation of building blocks between nutrient uptake systems in a microbe. *Journal of Mathematical Biology*, 44(3):276–296, 2002.

[89] Hugo Dourado and Martin J Lercher. An analytical theory of balanced cellular growth. *Nature communications*, 11(1):1–14, 2020.

[90] Francis Mairet, Jean-Luc Gouzé, and Hidde de Jong. Optimal proteome allocation and the temperature dependence of microbial growth laws. *npj Systems Biology and Applications*, 7(1):1–11, 2021.

[91] M. Basan, S. Hui, H. Okano, Z. Zhang, Y. Shen, J.R. Williamson, and T. Hwa. Overflow metabolism in *Escherichia coli* results from efficient proteome allocation. *Nature*, 528(7580):99–104, 2015.

[92] A. Maitra and K.A. Dill. Bacterial growth laws reflect the evolutionary importance of energy efficiency. *Proceedings of the National Academy of Sciences of the USA*, 112(2):406–411, 2015.

[93] Dean A. Carlson, Alain B. Haurie, and Arie Leizarowitz. *Infinite Horizon Optimal Control*. Springer, Berlin, Heidelberg, 1991.

[94] J. Izard, C. Gomez Balderas, D. Ropers, S. Lacour, X. Song, Y. Yang, A.B. Lindner, J. Geiselmann, and H. de Jong. A synthetic growth switch based on controlled expression of RNA polymerase. *Molecular Systems Biology*, 11(11):840, 2015.

[95] Jan Ewald, Martin Bartl, and Christoph Kaleta. Deciphering the regulation of metabolism with dynamic optimization: an overview of recent advances. *Biochemical Society Transactions*, 45(4):1035–1043, 2017.

[96] Edda Klipp, Reinhart Heinrich, and Hermann-Georg Holzhütter. Prediction of temporal gene expression: Metabolic optimization by re-distribution of enzyme activities. *European journal of biochemistry*, 269(22):5406–5413, 2002.

[97] Diego A Oyarzún, Brian P Ingalls, Richard H Middleton, and Dimitrios Kalamatianos. Sequential activation of metabolic pathways: a dynamic optimization approach. *Bulletin of mathematical biology*, 71(8):1851, 2009.

[98] Sheng Hui, Josh M Silverman, Stephen S Chen, David W Erickson, Markus Basan, Jilong Wang, Terence Hwa, and James R Williamson. Quantitative proteomic analysis reveals a simple strategy of global resource allocation in bacteria. *Molecular systems biology*, 11(2):784, 2015.

[99] Markus Basan, Manlu Zhu, Xiongfeng Dai, Mya Warren, Daniel Sévin, Yi-Ping Wang, and Terence Hwa. Inflating bacterial cells by increased protein synthesis. *Molecular systems biology*, 11(10):836, 2015.

[100] Nathan E Lewis, Kim K Hixson, Tom M Conrad, Joshua A Lerman, Pep Charusanti, Ashoka D Polpitiya, Joshua N Adkins, Gunnar Schramm, Samuel O Purvine, Daniel Lopez-Ferrer, et al. Omic data from evolved e. coli are consistent with computed optimal growth from genome-scale models. *Molecular systems biology*, 6(1):390, 2010.

[101] David Cass. Optimum growth in an aggregative model of capital accumulation. *The Review of economic studies*, 32(3):233–240, 1965.

[102] Jean-Michel Coron, Pierre Gabriel, and Peipei Shang. Optimization of an amplification protocol for misfolded proteins by using relaxed control. *Journal of mathematical biology*, 70(1-2):289–327, 2015.

[103] Stefan Klumpp, Zhongge Zhang, and Terence Hwa. Growth rate-dependent global effects on gene expression in bacteria. *Cell*, 139(7):1366–1375, 2009.

[104] Davide Roncarati and Vincenzo Scarlato. Regulation of heat-shock genes in bacteria: from signal sensing to gene expression output. *FEMS microbiology reviews*, 41(4):549–574, 2017.

[105] Ningzi Guan and Long Liu. Microbial response to acid stress: mechanisms and applications. *Applied microbiology and biotechnology*, 104(1):51–65, 2020.

[106] Hengjiang Dong, Lars Nilsson, and Charles G Kurland. Gratuitous overexpression of genes in Escherichia coli leads to growth inhibition and ribosome destruction. *Journal of bacteriology*, 177(6):1497–1504, 1995.

[107] Nikolaos Tsiantis and Julio R Banga. Using optimal control to understand complex metabolic pathways. *BMC bioinformatics*, 21(1):1–33, 2020.

[108] Han-Yu Chuang, Matan Hofree, and Trey Ideker. A decade of systems biology. *Annual review of cell and developmental biology*, 26:721–744, 2010.

[109] Carlo Cosentino and Declan Bates. *Feedback control in systems biology*. Crc Press, 2019.

[110] Thomas Schlitt and Alvis Brazma. Current approaches to gene regulatory network modelling. *BMC Bioinformatics*, 8(6):S9, 2007.

[111] Murat Arcak and Eduardo D Sontag. A passivity-based stability criterion for a class of biochemical reaction networks. *Mathematical biosciences & engineering*, 5(1):1, 2008.

[112] Etienne Farcot and Jean-Luc Gouzé. A mathematical framework for the control of piecewise-affine models of gene networks. *Automatica J. IFAC*, 44(9):2326–2332, 2008.

[113] Leon Glass and Joel S. Pasternack. Prediction of limit cycles in mathematical models of biological oscillations. *Bull. Math. Biology*, 40(1):27–44, 1978. Papers presented at the Society for Mathematical Biology Meeting (Univ. Pennsylvania, Philadelphia, Pa., 1976).

[114] John Mallet-Paret and Hal L Smith. The poincaré-bendixson theorem for monotone cyclic feedback systems. *Journal of Dynamics and Differential Equations*, 2(4):367–421, 1990.

[115] Eduardo D. Sontag. Passivity gains and the "secant condition" for stability. *Systems & Control Letters*, 55(3):177–183, 2006.

[116] Timothy Gardner, Charles Cantor, and James Collins. Construction of a genetic toggle switch in escherichia coli. *Nature*, 403:339–42, 02 2000.

[117] Mara C Inniss and Pamela A Silver. Building synthetic memory. *Current Biology*, 23(17):R812–R816, 2013.

[118] Lucie Chambon, Ismail Belgacem, and Jean-Luc Gouzé. Qualitative control of undesired oscillations in a genetic negative feedback loop with uncertain measurements. *Automatica*, 112:108642, 2020.

[119] Nicolas Augier, Madalena Chaves, and Jean-Luc Gouzé. Qualitative control strategies for synchronization of bistable gene regulatory networks. `https://hal.inria.fr/hal-02953502`, 2020.

[120] Madalena Chaves and Jean-Luc Gouzé. Exact control of genetic networks in a qualitative framework: the bistable switch example. *Automatica J. IFAC*, 47(6):1105–1112, 2011.

[121] Aleksei Fedorovich Filippov. Differential equations with discontinuous right-hand side. *Matematicheskii sbornik*, 93(1):99–128, 1960.

[122] Gábor Balázsi, Alexander van Oudenaarden, and James J Collins. Cellular decision making and biological noise: from microbes to mammals. *Cell*, 144(6):910–925, 2011.

[123] Jean-Baptiste Lugagne, Sebastián Sosa Carrillo, Melanie Kirch, Agnes Köhler, Gregory Batt, and Pascal Hersen. Balancing a genetic toggle switch by real-time feedback control and periodic forcing. *Nature communications*, 8(1):1–8, 2017.

[124] Agostino Guarino, Davide Fiore, and Mario Di Bernardo. In-silico feedback control of a mimo synthetic toggle switch via pulse-width modulation. pages 680–685. 2019 18th European Control Conference (ECC), 2019.

[125] Etienne Farcot and Jean-Luc Gouzé. A mathematical framework for the control of piecewise-affine models of gene networks. *Automatica*, 44(9):2326–2332, 2008.

[126] Patrick Hillenbrand, Georg Fritz, and Ulrich Gerland. Biological signal processing with a genetic toggle switch. *PloS one*, 8(7):e68345, 2013.

[127] Ahmad A. Mannan and Declan G. Bates. Designing an irreversible metabolic switch for scalable induction of microbial chemical production. *Nature Communications*, 12(1):3419, 2021.

[128] Jr Loomis, W F and B Magasanik. Glucose-lactose diauxie in escherichia coli. *Journal of bacteriology*, 93(4):1397–1401, 04 1967.

[129] Pierre Salvy and Vassily Hatzimanikatis. Emergence of diauxie as an optimal growth strategy under resource allocation constraints in cellular metabolism. *Proceedings of the National Academy of Sciences*, 118(8), 2021.

[130] Ugo Boscain and Paolo Mason. Time minimal trajectories for a spin 1/ 2 particle in a magnetic field. *Journal of Mathematical Physics*, 47(6):062101, 2006.

[131] Ugo Boscain, Fredrik Grönberg, Ruixing Long, and Herschel Rabitz. Minimal time trajectories for two-level quantum systems with two bounded controls. *Journal of Mathematical Physics*, 55(6):062106, 2014.

[132] Andrei V Dmitruk and Alexander M Kaganovich. The hybrid maximum principle is a consequence of pontryagin maximum principle. *Systems & Control Letters*, 57(11):964–970, 2008.