



**HAL**  
open science

# Dynamical analysis and optimization of a generalized resource allocation model of microbial growth

Agustín Gabriel Yabo, Jean-Baptiste Caillau, Jean-Luc Gouzé, Hidde de Jong, Francis Mairet

► **To cite this version:**

Agustín Gabriel Yabo, Jean-Baptiste Caillau, Jean-Luc Gouzé, Hidde de Jong, Francis Mairet. Dynamical analysis and optimization of a generalized resource allocation model of microbial growth. SIAM Journal on Applied Dynamical Systems, 2022, 21 (1), pp.137-165. 10.1137/21M141097X . hal-03251044v2

**HAL Id: hal-03251044**

<https://inria.hal.science/hal-03251044v2>

Submitted on 9 Jan 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Dynamical analysis and optimization of a generalized resource allocation model of microbial growth\*

Agustín Gabriel Yabo<sup>†</sup>, Jean-Baptiste Caillau<sup>‡</sup>, Jean-Luc Gouzé<sup>†</sup>, Hidde de Jong<sup>§</sup>, and Francis Mairet<sup>¶</sup>

**Abstract.** Gaining a better comprehension of the growth of microorganisms is a major scientific challenge, which has often been approached from a resource allocation perspective. Simple mathematical self-replicator models based on resource allocation principles have been surprisingly effective in accounting for experimental observations of the growth of microorganisms. Previous work, using a three-variable resource allocation model, predicted an optimal resource allocation scheme for the adaptation of microbial cells to a sudden nutrient change in the environment. We here propose an extended version of this model considering also proteins responsible for basic housekeeping functions, and we study their impact on predicted optimal strategies for resource allocation following changes in the environment. 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. We then explore the optimal allocation strategies using optimization and optimal control theory. We show that the solutions to this dynamical problem have a complicated structure that includes a second-order singular arc given in feedback form, and characterized by i) Fuller's phenomenon, and ii) the turnpike effect, producing a very particular asymptotic behaviour towards the solution of the static optimization problem. Our work thus provides a generalized perspective on the analysis of microbial growth by means of simple self-replicator models.

**Key words.** systems biology, bacterial growth laws, resource allocation, nutritional shifts, optimal control, turnpike

**AMS subject classifications.** 37N25, 49K15, 92C42

**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 [1]. 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 [2], 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

---

\*Submitted to the editors 8th April 2021.

**Funding:** This work was partially supported by ANR project Maximic (ANR-17-CE40-0024-01), Inria IPL Cosy and Labex SIGNALIFE (ANR-11-LABX-0028-01). We acknowledge the support of the FMJH Program PGMO and the support to this program from EDF-THALES-ORANGE.

<sup>†</sup>Université Côte d'Azur, Inria, INRAE, CNRS, Sorbonne Université, Biocore Team, Sophia Antipolis, France ([agustin.yabo@inria.fr](mailto:agustin.yabo@inria.fr), [www.agustinyabo.com.ar](http://www.agustinyabo.com.ar))

<sup>‡</sup>Université Côte d'Azur, CNRS, Inria, LJAD, France ([jean-baptiste.caillau@univ-cotedazur.fr](mailto:jean-baptiste.caillau@univ-cotedazur.fr))

<sup>§</sup>Université Grenoble Alpes, Inria, 38000 Grenoble, France ([hidde.de-jong@inria.fr](mailto:hidde.de-jong@inria.fr))

<sup>¶</sup>Ifremer, Physiology and Biotechnology of Algae laboratory, rue de l'Île d'Yeu, 44311 Nantes, France ([francis.mairet@ifremer.fr](mailto:francis.mairet@ifremer.fr))

34 production [3, 4, 5].

35 A fruitful perspective on microbial growth is to view it as a resource allocation problem  
36 [6]. Microorganisms must assign their available resources to different cellular functions, in-  
37 cluding the uptake and conversion of nutrients into molecular building blocks of proteins and  
38 other macromolecules (metabolism), the synthesis of proteins and other macromolecules from  
39 these building blocks (gene expression), and the detection of changes in the environment and  
40 the preparation of adequate responses (signalling and regulation). It is often assumed that  
41 microorganisms have evolved resource allocation strategies so as to maximize their growth  
42 rate, as this would allow them to outgrow competing species.

43 Simple mathematical models based on resource allocation principles have been surpris-  
44 ingly effective in accounting for experimental observations of the growth and physiology of  
45 microorganisms [6, 7, 8, 9, 10, 11, 12]. Instead of providing a detailed description of the en-  
46 tire biochemical reaction network, these models include a limited number of macroreactions  
47 responsible for the main growth-related functions of the cell. The models usually take the  
48 form of nonlinear ODE systems, typically 3-10 equations with parameters obtained from the  
49 experimental literature or estimated from published data. The models have been instrumental  
50 in explaining a number of steady-state relations between the growth rate and the cellular com-  
51 position, in particular the concentration of ribosomes, protein complexes that are responsible  
52 for the synthesis of new proteins [8, 6, 13, 10, 14]. Moreover, they have brought out a trade-off  
53 between the rate and yield of alternative metabolic pathways that produce energy-carrying  
54 molecules, necessary for driving forward many cellular reactions, such as those involved in the  
55 synthesis of proteins and other macromolecules [8, 15, 16].

56 In previous work, using a three-variable resource allocation model, it was possible to  
57 predict an optimal resource allocation scheme for the response of microbial cells to a sudden  
58 nutrient change in the environment [10]. The prediction was based on the Infinite Horizon  
59 Maximum Principle, a generalization of the well-known PMP (Pontrjagin Maximum Principle)  
60 [17, 18]. A feedback control strategy inspired by a known regulatory mechanism for growth  
61 control in the bacterial cell was shown to give a quasi-optimal approximation of the optimal  
62 solution. Strategies for optimal control were also explored for an extension of the model,  
63 inspired by recent experimental work [19], which comprises a pathway for the production  
64 of a metabolite of biotechnical interest as well as an external signal allowing growth to be  
65 switched off [20, 21, 22, 23]. We showed by a combination of analytical and computational  
66 means that the optimal solution for the targeted metabolite production problem consists of  
67 a phase of growth maximization followed by a phase of product maximization, in agreement  
68 with strategies proposed in metabolic engineering. Optimal control approaches have also been  
69 used for studying other dynamic optimization problems in biology (see [24] for a review). A  
70 classical example is the determination of optimal activation patterns of metabolic pathways,  
71 such as to minimize the transition time of metabolites or minimize enzyme costs [25, 26].

72 The resource allocation model that lies at the basis of the above-mentioned work [10] has  
73 a number of limitations. First, the biomass of the cell was assumed to consist of two classes of  
74 proteins, enzymes catalyzing metabolic reactions and ribosomes responsible for protein syn-  
75 thesis, whose relative proportions vary with the growth rate. However, experimental data  
76 show that a large fraction of the total protein contents of the cell is growth rate-independent  
77 [27]. This suggests the introduction of a third protein category, dedicated mainly to ba-

78 sic housekeeping functions of the cell. The proportion of these proteins is independent of  
79 the growth rate and thus constrains the variations in the other two, growth rate-dependent  
80 categories [6, 13]. Second, the concentration of ribosomes and enzymes, the two protein  
81 categories included in the original model, have both a growth rate-dependent and a growth  
82 rate-independent component [6, 27]. This implies that the protein synthesis rate, and thus the  
83 growth rate, does not depend on the total ribosome concentration, as in the original model,  
84 but only on its growth rate-dependent fraction [13].

85 In the present manuscript, we revise the above modeling assumptions and study their  
86 impact on predicted optimal strategies for resource allocation following changes in the envi-  
87 ronment of different nature (i.e., changes in the nutrient concentration or stress responses).  
88 This leads to a number of interesting problems in mathematical analysis and control, which  
89 are addressed using tools from dynamical systems analysis and optimal control theory. A full  
90 dynamical analysis of the system shows there is a single globally attractive equilibrium, which  
91 can be related to steady-state growth conditions of bacteria observed in experiments. In spite  
92 of the simplicity of the presented model, the solutions of the associated biomass maximization  
93 problems exhibit quite interesting features. Notably, the second-order singular arc is charac-  
94 terized by a) the Fuller’s phenomenon at its junctions, yielding an infinite set of switching  
95 points in a finite-time window, and b) the turnpike effect, which produces very particular as-  
96 ymptotic behaviors towards the solution of the static optimization problem. We provide a full  
97 description of the singular arc in terms of the state, as well as an explicit proof of the presence  
98 of the turnpike effect. While the predicted (optimal) control dynamics does not change much  
99 qualitatively in comparison with the previous model, the more realistic modeling assumptions  
100 offer a more general perspective of the biological problem. For example, in contrast with  
101 the previous model where the absence of growth-rate independent protein yields a constant  
102 singular arc equal to the solution of the static optimization problem, the singular arc of the  
103 new model is not constant, but governed by a turnpike phenomenon.

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

110 **2. Model definition.** We define a self-replicator system composed of the mass of precursor  
111 metabolites  $P$ , the gene expression machinery  $R$  (ribosomes, RNA polymerase, ...) and the  
112 metabolic machinery  $M$  (enzymes, transporters, ...), as shown in Figure 1. Essentially, the  
113 ribosomal proteins  $R$  are responsible for the fabrication of new proteins, and the metabolic pro-  
114 teins  $M$  are in charge of the uptake of nutrients for building precursor metabolites  $P$ . Following  
115 Scott *et al.* [6], we also introduce a class  $Q$  of proteins whose functions fall outside the range of  
116 tasks performed by  $M$  and  $R$ . This sector comprises mainly growth rate-independent proteins  
117 such as housekeeping proteins responsible for the maintenance of certain basic cellular func-  
118 tions. Needless to say, the synthesis of  $Q$  proteins draws resources away from the pathways to  
119  $M$  and  $R$ , and consequently imposes an upper bound on the fraction of resources dedicated to  
120 self-replication and nutrient uptake. This constraint appears in the model through a constant

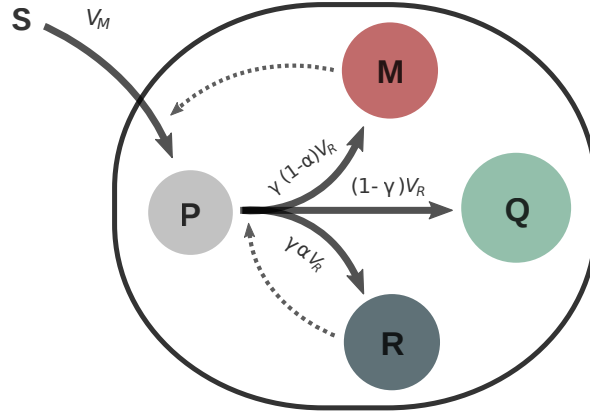
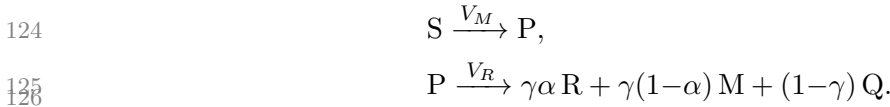


Figure 1: Coarse-grained self-replicator model. The external substrate  $S$  is consumed by bacteria and transformed into precursor metabolites  $P$  by the metabolic machinery  $M$ . The precursors are used to produce macromolecules of classes  $R$ ,  $M$  and  $Q$ , with proportions  $\gamma\alpha$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

121  $\gamma \in [0, 1]$ , and it indicates the maximum fraction of the protein synthesis rate available for  
 122 making ribosomes and metabolic enzymes. The overall allocation process can be represented  
 123 by the biochemical macroreactions



127 The first reaction describes the transformation of external substrate  $S$  into precursor me-  
 128 tabolites  $P$  at a rate  $V_M$ . The second reaction represents the conversion of precursors into  
 129 macromolecules  $R$ ,  $M$ , and  $Q$  at a rate  $V_R$ . The roles of the enzymes  $M$  in the uptake and  
 130 metabolization of nutrients and the ribosomal proteins  $R$  in the production of proteins are rep-  
 131 resented through catalytic effects, indicated with dotted arrows in Figure 1. In this context,  
 132 protein  $A$  *catalyzes* reaction  $B$  means that the rate of reaction  $B$  is proportional to the cellular  
 133 concentration of  $A$ , but the reaction itself does not consume  $A$ . The natural resource alloca-  
 134 tion strategy is modeled through the time-varying function  $\alpha(t) \in [0, 1]$ . Thus, the proportion  
 135 of the total synthesis rate of proteins dedicated to the gene expression machinery  $R$  is  $\gamma\alpha$ ,  
 136 while that of the metabolic machinery  $M$  is  $\gamma(1-\alpha)$ . In particular, the allocation parameter  
 137 does not influence the synthesis rate of  $Q$ , with constant proportion  $1-\gamma$ , as the synthesis of  
 138 proteins in this class is auto-regulated through mechanisms not relevant in this study. From  
 139 a biological perspective, the function  $\alpha(t)$  represents the naturally-evolved allocation strategy  
 140 of the cell which is, *a priori*, unknown. In the context of control theory, and throughout this  
 141 paper,  $\alpha$  is treated as the control input of the system.

142 **2.1. Self-replicator system.** Generalizing upon Giordano *et al.* [10], a mass balance  
 143 analysis yields the dynamical system

$$144 \quad \begin{cases} \dot{P} = V_M - V_R, \\ \dot{R} = \gamma\alpha V_R, \\ \dot{M} = \gamma(1 - \alpha)V_R, \\ \dot{Q} = (1 - \gamma)V_R. \end{cases}$$

145

146 where mass quantities  $P$ ,  $M$ ,  $R$  and  $Q$  are described in grams (g), the synthesis rates  $V_M$  and  
 147  $V_R$  in grams per hour, and  $\alpha$  is the dimensionless allocation parameter. In what follows, we  
 148 will assume that the proteins of classes  $R$ ,  $M$  and  $Q$  are responsible for most of the bacterial  
 149 mass [1], and so we define the bacterial volume  $\mathcal{V}$  measured in liter units (L) as

$$150 \quad (2.1) \quad \mathcal{V} = \beta(R + M + Q),$$

151

152 where  $\beta$  corresponds to a density constant relating mass and bacterial volume [28], such that  
 153 the total biomass in grams is given by  $\mathcal{V}/\beta$ . The above assumption implies that the mass of  
 154 precursor metabolites represents a negligible fraction of the total biomass, (in other words,  
 155  $P \ll \mathcal{V}/\beta$ ). We define the intracellular concentrations in grams per liter as

$$156 \quad (2.2) \quad p_{\mathcal{V}} \doteq \frac{P}{\mathcal{V}}, \quad r_{\mathcal{V}} \doteq \frac{R}{\mathcal{V}}, \quad m_{\mathcal{V}} \doteq \frac{M}{\mathcal{V}}, \quad q_{\mathcal{V}} \doteq \frac{Q}{\mathcal{V}}.$$

157

158 Using (2.1) and (2.2), we to obtain the relation

$$159 \quad (2.3) \quad r_{\mathcal{V}} + m_{\mathcal{V}} + q_{\mathcal{V}} = \frac{1}{\beta}.$$

160

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

$$163 \quad v_M(s, m_{\mathcal{V}}) \doteq \frac{V_M}{\mathcal{V}}, \quad v_R(p_{\mathcal{V}}, r_{\mathcal{V}}) \doteq \frac{V_R}{\mathcal{V}},$$

164

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

$$168 \quad \mu \doteq \frac{\dot{\mathcal{V}}}{\mathcal{V}} = \frac{\beta V_R}{\mathcal{V}} = \beta v_R(p_{\mathcal{V}}, r_{\mathcal{V}}).$$

169

170 We write the system in terms of the concentrations as

$$171 \quad \begin{cases} \dot{p}_{\mathcal{V}} = v_M(s, m_{\mathcal{V}}) - (1 + \beta p_{\mathcal{V}})v_R(p_{\mathcal{V}}, r_{\mathcal{V}}), \\ \dot{r}_{\mathcal{V}} = (\gamma\alpha - \beta r_{\mathcal{V}})v_R(p_{\mathcal{V}}, r_{\mathcal{V}}), \\ \dot{m}_{\mathcal{V}} = (\gamma(1 - \alpha) - \beta m_{\mathcal{V}})v_R(p_{\mathcal{V}}, r_{\mathcal{V}}), \\ \dot{q}_{\mathcal{V}} = ((1 - \gamma) - \beta q_{\mathcal{V}})v_R(p_{\mathcal{V}}, r_{\mathcal{V}}), \\ \dot{\mathcal{V}} = \beta v_R(p_{\mathcal{V}}, r_{\mathcal{V}})\mathcal{V}. \end{cases}$$

172

173 **2.2. Kinetic definition.** We define the kinetics of the reaction system by taking into  
174 account that a minimal concentration of ribosomal proteins  $r_{\mathcal{V},\min} \in (0, \gamma/\beta)$  is required for  
175 protein synthesis to take place. In other words, a part of the bacterial volume is occupied  
176 by ribosomal proteins which do not directly contribute to growth [13]. Such behavior can be  
177 modeled as

$$178 \quad v_R(p_{\mathcal{V}}, r_{\mathcal{V}}) \doteq w_R(p_{\mathcal{V}})(r_{\mathcal{V}} - r_{\mathcal{V},\min})^+, \quad \text{with } (r_{\mathcal{V}} - r_{\min})^+ = \begin{cases} r_{\mathcal{V}} - r_{\mathcal{V},\min} & \text{if } r_{\mathcal{V}} \geq r_{\mathcal{V},\min}, \\ 0 & \text{if } r_{\mathcal{V}} < r_{\mathcal{V},\min}. \end{cases}$$

179

180 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  
181 conditions lie in a particular region of the state space. The rate of nutrient uptake is defined  
182 as

$$183 \quad v_M(s, m_{\mathcal{V}}) \doteq w_M(s)m_{\mathcal{V}}.$$

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

186 *Hypothesis 2.1.* Function  $w_i(x) : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  is

- 187 • Continuously differentiable w.r.t.  $x$ ,
- 188 • Null at the origin:  $w_i(0) = 0$ ,
- 189 • Strictly increasing:  $w_i'(x) > 0, \forall x \geq 0$ ,
- 190 • Strictly concave:  $w_i''(x) < 0, \forall x \geq 0$ ,
- 191 • Upper bounded:  $\lim_{x \rightarrow \infty} w_i(x) = k_i > 0$ .

192 The classical Michaelis-Menten kinetics satisfies Hypothesis 2.1. While most of the math-  
193 ematical results are based on this general definition, for the calibration of the model and  
194 numerical simulations, we will resort to the particular case where the functions are defined as

$$195 \quad (2.4) \quad w_R(p_{\mathcal{V}}) \doteq k_R \frac{p_{\mathcal{V}}}{K_R + p_{\mathcal{V}}}, \quad w_M(s) \doteq k_M \frac{s}{K_S + s},$$

196

197 where  $k_R$  and  $k_M$  are the maximal reaction rates in  $\text{h}^{-1}$ , and  $K_M$  and  $K_R$  are the half-  
198 saturation constants of the synthesis rates in  $\text{g L}^{-1}$ . For the general case introduced in  
199 Hypothesis 2.1 we will define

$$200 \quad k_R \doteq \lim_{p_{\mathcal{V}} \rightarrow \infty} w_R(p_{\mathcal{V}}).$$

201

202 **2.3. Constant environmental conditions.** We assume that the availability of the sub-  
 203 strate in the medium is constant over the time-window analyzed. The latter can be modeled  
 204 by setting  $s$  constant, and thus removing the dynamics of  $s$  from the system.

205 *Hypothesis 2.2.* The flow of substrate can be expressed as  $w_M(s) = e_M$  with  $e_M > 0$   
 206 constant.

207 Using this assumption, the dynamical equation of  $p_V$  becomes

$$208 \quad \dot{p}_V = e_M m_V - (1 + \beta p_V) w_R(p_V) (r_V - r_{V,\min})^+.$$

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

213 **2.4. Mass fraction formulation and non-dimensionalization.** We define mass fractions  
 214 of the total bacterial mass as

$$215 \quad p \doteq \beta p_V, \quad r \doteq \beta r_V, \quad r_{\min} \doteq \beta r_{V,\min}, \quad m \doteq \beta m_V, \quad q \doteq \beta q_V,$$

217 which, replacing in (2.3), yields the relation

$$218 \quad (2.5) \quad r + m + q = 1.$$

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

$$222 \quad (2.6) \quad \hat{\mu}(p, r) \doteq \frac{\mu(p_V, r_V)}{k_R} = \hat{w}_R(p) (r - r_{\min}),$$

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

$$226 \quad (S) \quad \begin{cases} \dot{p} = E_M m - (p + 1) w_R(p) (r - r_{\min})^+, \\ \dot{r} = (\gamma \alpha - r) w_R(p) (r - r_{\min})^+, \\ \dot{m} = (\gamma(1 - \alpha) - m) w_R(p) (r - r_{\min})^+, \\ \dot{\mathcal{V}} = w_R(p) (r - r_{\min})^+ \mathcal{V}, \\ m + r \leq 1, \end{cases}$$

227  
 228 where  $q$  has been removed since it can be expressed in terms of the other concentrations  
 229 through (2.5); and the constraint  $m + r \leq 1$  is required to comply with  $q \geq 0$ . The model  
 230 differs from that of Giordano *et al.* by the addition of the category of housekeeping proteins  
 231 ( $q$ ) and a minimum concentration of ribosomes for protein synthesis ( $r_{\min}$ ). In what follows,  
 232 we will systematically investigate how these differences affect the asymptotic behavior and  
 233 optimal resource allocation strategies.



234 **3. Asymptotic behavior.** In the present section, we will study the asymptotic behavior  
 235 of the reduced system representing the intracellular dynamics

$$236 \quad (3.1) \quad \begin{cases} \dot{p} = E_M m - (p+1)w_R(p)(r-r_{\min})^+, \\ \dot{r} = (\gamma\alpha - r)w_R(p)(r-r_{\min})^+, \\ \dot{m} = (\gamma(1-\alpha) - m)w_R(p)(r-r_{\min})^+, \\ m+r \leq 1, \end{cases}$$

237

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

241 **Lemma 3.1.** *The set*

$$242 \quad \Gamma = \{(p, r, m) \in \mathbb{R}^3 : p \geq 0, \gamma \geq r \geq r_{\min}, \gamma \geq m \geq 0, m+r \leq 1\}$$

243 *is positively invariant by (3.1).*

244 *Proof.* This can be easily verified by evaluating the differential equations of system (3.1)  
 245 over the boundaries of  $\Gamma$ . As for the condition  $m+r \leq 1$ , we can define a variable  $z \doteq m+r$   
 246 that obeys the dynamics

$$247 \quad \dot{z} = (\gamma - z)w_R(p)(r-r_{\min})^+$$

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

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

254 **Definition 3.2.** *The maximal ribosomal mass fraction is  $r_{\max} \doteq \gamma$ .*

255 Then, we will reduce the study of the system to this set and so, using Definition 3.2, we  
 256 redefine (3.1) as

$$257 \quad (S') \quad \begin{cases} \dot{p} = E_M m - (p+1)w_R(p)(r-r_{\min}) \\ \dot{r} = (r_{\max}\alpha - r)w_R(p)(r-r_{\min}) \\ \dot{m} = (r_{\max}(1-\alpha) - m)w_R(p)(r-r_{\min}) \end{cases}$$

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

$$261 \quad \alpha_{\min}^* \doteq \frac{r_{\min}}{r_{\max}}.$$

262 Its importance will be analyzed throughout the current section.

267 **3.1. Local stability.**

 268 **Theorem 3.3.** *System (S') has the equilibria*

- 269 •
- $E_1 \doteq (p^*, r^*, m^*)$
- , locally stable if
- $\alpha^* > \alpha_{\min}^*$
- .
- 
- 270 •
- $E_2 \doteq (p, r_{\min}, 0)$
- , locally unstable if
- $\alpha^* > \alpha_{\min}^*$
- .
- 
- 271 •
- $E_3 \doteq (0, r, 0)$
- , locally unstable if
- $r \neq r_{\min}$
- .

272 with

273 (3.2) 
$$p^* \doteq \left\{ p \in \mathbb{R}_+ : (p+1)w_R(p) = \frac{E_M m^*}{r^* - r_{\min}} \right\},$$
 274 
$$r^* \doteq r_{\max} \alpha^*,$$
 275 
$$m^* \doteq r_{\max}(1 - \alpha^*).$$

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

278 (3.3) 
$$\begin{bmatrix} -(w_R(p) + (p+1)w'_R(p))(r - r_{\min}) & -(p+1)w_R(p) & E_M \\ (r_{\max}\alpha - r)w'_R(p)(r - r_{\min}) & (r_{\max}\alpha - 2r + r_{\min})w_R(p) & 0 \\ (r_{\max}(1 - \alpha) - m)w'_R(p)(r - r_{\min}) & (r_{\max}(1 - \alpha) - m)w_R(p) & -w_R(p)(r - r_{\min}) \end{bmatrix}.$$

 280 We first see that, if  $\alpha^* > \alpha_{\min}^*$ , the value  $p^*$  is unique since  $(p+1)w_R(p)$  is a monotone  
 281 increasing function satisfying  $w_R(0) = 0$  and  $\lim_{p \rightarrow \infty} (p+1)w_R(p) = \infty$  (as stated in Hypoth-  
 282 esis 2.1), and  $E_M m^*/(r^* - r_{\min}) > 0$ , so the set (3.2) yields a unique solution. For  $\alpha^* < \alpha_{\min}^*$ ,  
 283 the equation for  $p^*$  in (3.2) has no valid solution as  $E_M m^*/(r^* - r_{\min})$  becomes negative, and  
 284 therefore the equilibrium does not exist. The Jacobian (3.3) for  $E_1$  becomes

285 
$$J_1 = \begin{bmatrix} -(w_R(p^*) + (p^*+1)w'_R(p^*))(r^* - r_{\min}) & -(p^*+1)w_R(p^*) & E_M \\ 0 & -(r^* - r_{\min})w_R(p^*) & 0 \\ 0 & 0 & -w_R(p^*)(r^* - r_{\min}) \end{bmatrix}$$

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

291 
$$J_2 = \begin{bmatrix} 0 & -w_R(p) & E_M \\ 0 & (r^* - r_{\min})w_R(p) & 0 \\ 0 & r_{\max}(1 - \alpha^*)w_R(p) & 0 \end{bmatrix}$$

293 with characteristic polynomial

294 
$$P_2(\lambda) = \lambda^2(\lambda - (r^* - r_{\min})w_R(p)).$$
 295

296 If  $\alpha^* > \alpha_{\min}^*$ , then  $J_2$  has one positive eigenvalue and  $E_2$  becomes locally unstable. As for  $E_3$ ,  
 297 the Jacobian is

$$298 \quad J_3 = \begin{bmatrix} -w'_R(0)(r - r_{\min}) & 0 & E_M \\ (r_{\max}\alpha - r)w'_R(0)(r - r_{\min}) & 0 & 0 \\ 299 \quad r_{\max}(1 - \alpha)w'_R(0)(r - r_{\min}) & 0 & 0 \end{bmatrix}$$

300 with characteristic polynomial

$$301 \quad P_3(\lambda) = \lambda^2 \left( \lambda + w'_R(0)(r - r_{\min}) \right) - E_M r_{\max}(1 - \alpha)w'_R(0)(r - r_{\min})\lambda.$$

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

$$304 \quad \lambda^2 + \lambda w'_R(0)(r - r_{\min}) - E_M r_{\max}(1 - \alpha)w'_R(0)(r - r_{\min}) = 0.$$

306 By the Routh-Hurwitz criterion, the two remaining roots are in the open left half plane if  
 307 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  
 308 true. Consequently, for  $r \neq r_{\min}$ , there is at least one positive root, and so the equilibrium is  
 309 unstable. ■

310 **3.2. Global behavior.** We will study the global behavior of system (S') for the initial  
 311 conditions

$$312 \quad \text{(IC)} \quad p(0) > 0, \quad r(0) \in (r_{\min}, r_{\max}), \quad m(0) \in (0, r_{\max}), \quad r(0) + m(0) \leq 1.$$

314 and for a given constant allocation parameter

$$315 \quad \alpha(t) = \alpha^* \in (\alpha_{\min}^*, 1).$$

317 Under this constraint, we see that the dynamics of  $r$  and  $m$  become

$$318 \quad \dot{r} = (r^* - r)w_R(p)(r - r_{\min}), \quad \dot{m} = (m^* - m)w_R(p)(r - r_{\min}),$$

320 which means that, if  $p > 0$  and  $r > r_{\min}$ , the signs of  $\dot{r}$  and  $\dot{m}$  are given by the signs of  $r^* - r$   
 321 and  $m^* - m$ , respectively (and both  $\dot{r}$  and  $\dot{m}$  are zero if  $p = 0$  or  $r = r_{\min}$ ). Then, let us  
 322 divide  $\Gamma$  into the subsets

$$323 \quad \mathcal{R}^- \doteq \{(p, r, m) \in \Gamma : r \in (r_{\min}, r^*)\}, \quad \mathcal{M}^- \doteq \{(p, r, m) \in \Gamma : m \in (0, m^*)\}, \\ 324 \quad \mathcal{R}^+ \doteq \{(p, r, m) \in \Gamma : r \in (r^*, r_{\max})\}, \quad \mathcal{M}^+ \doteq \{(p, r, m) \in \Gamma : m \in (m^*, r_{\max})\},$$

325 such that  $\Gamma = \overline{\mathcal{R}^-} \cup \overline{\mathcal{R}^+} = \overline{\mathcal{M}^-} \cup \overline{\mathcal{M}^+}$ . In these sets, the following holds.

326 **Lemma 3.4.** *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*  
 327 *by (S'), and*

$$328 \quad \begin{cases} \dot{r} \geq 0 & \text{if } (p, r, m) \in \mathcal{R}^-, \\ \dot{r} \leq 0 & \text{if } (p, r, m) \in \mathcal{R}^+, \end{cases} \quad \begin{cases} \dot{m} \geq 0 & \text{if } (p, r, m) \in \mathcal{M}^-, \\ \dot{m} \leq 0 & \text{if } (p, r, m) \in \mathcal{M}^+. \end{cases}$$

330 Again, the invariance of the sets can be checked by evaluating the vector field over the  
 331 boundaries of the sets. Now we state a first result.

332 **Proposition 3.5.** For  $\alpha(t) = \alpha^* \in (\alpha_{\min}^*, 1)$  and initial conditions (IC), system (S') has a  
 333 lower bound

$$334 \quad (p, r, m) \geq (p_{\text{low}}, r_{\text{low}}, m_{\text{low}}) \text{ for all } t \geq 0,$$

336 with

$$337 \quad (3.4) \quad \begin{aligned} r_{\text{low}} &\doteq \min(r(0), r^*), & m_{\text{low}} &\doteq \min(m(0), m^*), \\ p_{\text{low}} &\doteq \left\{ p \in \mathbb{R}_+ : (p+1)w_R(p) = \frac{E_M m_{\text{low}}}{r_{\text{max}} - r_{\text{min}}} \right\}. \end{aligned}$$

339 *Proof.* For a trajectory emanating from  $\mathcal{R}^-$  (respectively,  $\mathcal{R}^+$ ), it follows that  $\dot{r} \geq 0$   
 340 (respectively,  $\dot{r} \leq 0$ ) for all  $t$  (according to Lemma 3.4), and so  $r \geq r(0)$  (respectively,  $r \geq r^*$ )  
 341 for all  $t$ . This proves that  $r \geq \min(r(0), r^*) > r_{\min}$  for all  $t$  (depending on whether the  
 342 trajectory starts in  $\mathcal{R}^-$  or  $\mathcal{R}^+$ ). Similarly, a trajectory starting in  $\mathcal{M}^-$  (respectively,  $\mathcal{M}^+$ )  
 343 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  
 344  $t$ . Then, it follows that  $m \geq \min(m(0), m^*)$  for all  $t \geq 0$ . The equation for  $p$  can thus be  
 345 lower-bounded to

$$346 \quad \dot{p} \geq E_M m_{\text{low}} - (p+1)w_R(p)(r_{\text{max}} - r_{\text{min}}),$$

348 which means  $p \geq p_{\text{low}}$  for all  $t \geq 0$ , with  $p_{\text{low}}$  the solution of (3.4), which is unique by the  
 349 same arguments as those used in Theorem 3.3. ■

350 A lower bound on system (S') is a stronger condition than the classical persistence for bio-  
 351 logical populations, as the bound is imposed not only for  $t \rightarrow \infty$  but for the whole trajectory.  
 352 As a consequence, the growth rate never vanishes, as it meets  $\mu(p, r) \geq w_R(p_{\text{low}})(r_{\text{low}} - r_{\text{min}}) >$   
 353 0 for all  $t \geq 0$ . Then, the global stability of the system is straightforward.

354 **Theorem 3.6.** For  $\alpha(t) = \alpha^* \in (\alpha_{\min}^*, 1)$  and initial conditions (IC), every solution of (S')  
 355 converges to the equilibrium  $E_1$ .

356 *Proof.* Since  $p \geq p_{\text{low}} > 0$  and  $r \geq r_{\text{low}} > r_{\text{min}}$  for all  $t \geq 0$ , we have that  $\text{sign}(\dot{r}) =$   
 357  $\text{sign}(r^* - r)$  and  $\text{sign}(\dot{m}) = \text{sign}(m^* - m)$ , showing that  $r$  and  $m$  converge asymptotically to  $r^*$   
 358 and  $m^*$ , respectively. Consequently, the dynamical equation of  $p$  becomes  $\dot{p} = E_M m^* - (p +$   
 359  $1)w_R(p)(r^* - r_{\text{min}})$  and so  $\text{sign}(\dot{p}) = \text{sign}(p^* - p)$ , which means that  $p$  converges asymptotically  
 360 to the steady-state value  $p^*$ . ■

361 *Remark 3.7.* For the case over the invariant plane given by  $r(0) = r_{\text{min}}$  and  $m(0) > 0$ ,  
 362 concentrations  $m$  and  $r$  are constant along the whole trajectory, and  $p$  increases linearly with  
 363 time (as  $\dot{p} = E_M m(0)$ ). This is a degenerate case that contradicts the assumption  $p \ll 1$ , and  
 364 lacks biological relevance.

365 **3.3. Maximum steady-state growth rate.** A classical hypothesis in the literature is to  
 366 suppose bacterial populations in steady-state regimes maximize their growth rate ([10] and  
 367 references therein). We are interested in finding the static allocation strategy  $\alpha^*$  that produces  
 368 this situation. Since the only equilibrium that admits bacterial growth is  $E_1$ , we will express

369 the static optimization problem as

$$370 \quad \max_{\alpha^* \in [\alpha_{\min}^*, 1]} \mu(p^*, r^*),$$

372 which can be rewritten as  $\mu(p^*, r^*) = w_R(p^*)(r^* - r_{\min})$ . It is possible to express  $\alpha^*$  in terms  
373 of  $p^*$  through the relation

$$374 \quad (3.5) \quad \alpha^*(p^*) = \frac{E_M + (p^* + 1)w_R(p^*)\alpha_{\min}^*}{E_M + (p^* + 1)w_R(p^*)}.$$

376 Moreover, since the above function  $\alpha^*(p^*) : \mathbb{R}_+ \rightarrow (\alpha_{\min}^*, 1]$  is monotone decreasing, it is  
377 possible to write the optimization problem in terms of  $p^*$  instead of  $\alpha^*$ . The growth rate in  
378 terms of  $p^*$  writes

$$379 \quad (3.6) \quad w_R(p^*)(r^* - r_{\min}) = (r_{\max} - r_{\min}) \left( \frac{E_M w_R(p^*)}{E_M + (p^* + 1)w_R(p^*)} \right).$$

381 We differentiate w.r.t.  $p^*$  and we get the relation  $w_R(p^*)^2 = E_M w'_R(p^*)$ , which has a unique  
382 solution since, according to Hypothesis 2.1,  $w_R(p)^2$  is a monotone increasing function satisfying  
383  $w_R^2(0) = 0$  and  $\lim_{p \rightarrow \infty} w_R^2(p) = 1$ , and  $w'_R(p)$  is a monotone decreasing function satisfying  
384  $w'_R(0) > 0$  and  $\lim_{p \rightarrow \infty} w'_R(p) = 0$  (as  $w_R(p)$  is a strictly increasing upper-bounded function).  
385 Then, the condition for optimality can be expressed as

$$386 \quad (3.7) \quad \frac{w_R(p_{\text{opt}}^*)^2}{E_M w'_R(p_{\text{opt}}^*)} = 1.$$

388 Thus, the optimal allocation parameter  $\alpha^*$  is obtained by replacing  $p_{\text{opt}}^*$  in (3.5), and the  
389 maximal static growth rate can be calculated using (3.6). From (3.7), it can be seen that  $p_{\text{opt}}^*$   
390 depends neither on  $r_{\min}$  nor on  $r_{\max}$ , suggesting that the steady-state precursor concentration  
391 is independent of the housekeeping protein fraction  $q$  and of the growth rate-independent  
392 ribosomal fraction. Conversely, the precursor concentration is rather determined by the en-  
393 vironmental conditions and by the nature of the function  $w_R(p)$ . It can be proven that the  
394 latter result is not a consequence of assumption (2.1): when considering a definition of the  
395 bacterial volume as  $\beta(P + R + M + Q)$ , which takes into account the mass  $P$ , the optimal  
396 precursor concentration amounts to  $p_{\text{opt}}^*/(1 + p_{\text{opt}}^*)$ .

397 In addition, from  $\dot{p} = 0$  in (S'), we get:

$$398 \quad \frac{r^* - r_{\min}}{m^*} = \frac{E_M}{(p^* + 1)w_R(p^*)}.$$

400 This shows that, for the optimal steady state, the concentration ratio of the active gene expres-  
401 sion machinery over the metabolic machinery does not depend on  $r_{\max}$  either. Thus, a cellular  
402 strategy regulating the precursor concentration and the balance between gene expression and  
403 metabolism could lead to the optimal equilibrium, regardless of the demand for  $Q$ .

404 **4. Model calibration.** Whereas the parameter values do not affect the results above and  
 405 the optimal control analysis in the next section, they are nevertheless important for simulations  
 406 illustrating the dynamics and optimal allocation strategies of system (S'). Below, we derive  
 407 such parameters for the model bacterium *Escherichia coli*, using published sources. The  $\beta$   
 408 constant used in the definition of the bacterial volume (2.1) corresponds to the inverse of the  
 409 protein density, which is set to 0.003 [L g<sup>-1</sup>] based on [10]. According to [6], the ribosomal  
 410 fraction of the proteome<sup>1</sup> can vary between 6% and 55%. In more recent studies [27], this sector  
 411 is divided into growth-rate dependent and independent fractions. The maximal growth-rate  
 412 dependent ribosomal fraction of the proteome is estimated to be 41%, and the growth rate-  
 413 independent fraction is 9%. Based on these experimental estimations, we set  $r_{\max} = 0.41 +$   
 414  $0.09 = 0.5$ . We performed further calibrations using data sets from [29, 30, 6, 27] containing  
 415 measurements of various strains of *E. coli* growing in different media. The data sets are  
 416 composed essentially of data points (*growth rate, RNA/protein mass ratio*) measured at steady  
 417 state. Most RNA is ribosomal RNA found overwhelmingly in ribosomes, the main constituent  
 418 of the gene expression machinery. In order to adjust the measurements to model (S'), the  
 419 observed RNA/protein ratios can be converted to mass fractions  $r$  through multiplication  
 420 with a conversion factor  $\rho = 0.76$   $\mu\text{g}$  of protein/ $\mu\text{g}$  of RNA [6]. As a result, we have  $n$   
 421 measurements of form  $(\tilde{\mu}_k, \tilde{r}_k)$  which are assumed to follow a linear relation [6], as seen in  
 422 Figure 2a. From the vertical intercept of the linear regression performed using the data  
 423 points, we obtain  $r_{\min} = 0.07$ , in agreement with previous studies [6, 13, 27]. Each data  
 424 point, composed of an observed growth rate and its associated ribosomal mass fraction, can  
 425 be related to an optimal steady state of system (S') for a certain environmental condition  $e_M$ .  
 426 Thus, each  $k$ th pair  $(\tilde{\mu}_k, \tilde{r}_k)$  of the  $n$  measurements should yield a constant environmental  
 427 condition  $e_{M,k}$ , and all pairs should simultaneously adjust the rate constant  $k_R$ . Such fitting  
 428 can be done by resorting to the Michaelis-Menten kinetic form introduced in (2.4). Based on  
 429 [10], we fix the half-saturation constant of protein synthesis  $K_R = 1$  g L<sup>-1</sup>. We then define  
 430 the parameter vector  $\theta = (k_R, e_{M,1}, \dots, e_{M,n})$  which is computed by solving a least-squares  
 431 regression problem. Using the relation (2.6), the cost function to minimize is

$$432 \quad \min_{\theta \in \mathbb{R}_+^{n+1}} \sum_{k=1}^n (\tilde{\mu}_k - \mu_{\text{opt}}^*(k_R, e_{M,k}))^2 + (\tilde{r}_k - r_{\text{opt}}^*(k_R, e_{M,k}))^2,$$

434 where the non-dimensional growth rate  $\mu_{\text{opt}}^*$  is calculated using (3.6), and the optimal steady  
 435 state  $(r_{\text{opt}}^*, p_{\text{opt}}^*, m_{\text{opt}}^*)$  is expressed in terms of  $\alpha_{\text{opt}}^*$  (using Theorem 3.3) which is, at the same  
 436 time, a function of  $k_R$  and  $e_{M,k}$ . The numerical solution yields  $k_R = 6.23$  h<sup>-1</sup>, and different  
 437 values of  $e_M$  matching different nutrients from the dataset (see Figure 2b). We can validate  
 438 these results by computing the maximal growth rate  $k_R(r_{\max} - r_{\min}) = 2.68$  h<sup>-1</sup> based on  
 439 the adjusted parameters, which is a value that corresponds well with literature values of the  
 440 maximal growth rate of *E. coli* in rich media [30].

## 441 5. Optimal resource allocation.

442 **5.1. Problem definition.** In this section we formulate the dynamic optimization problem  
 443 under the hypothesis that microbial populations have evolved resource allocation strategies

---

<sup>1</sup>The proteome is the total amount of protein in the cell

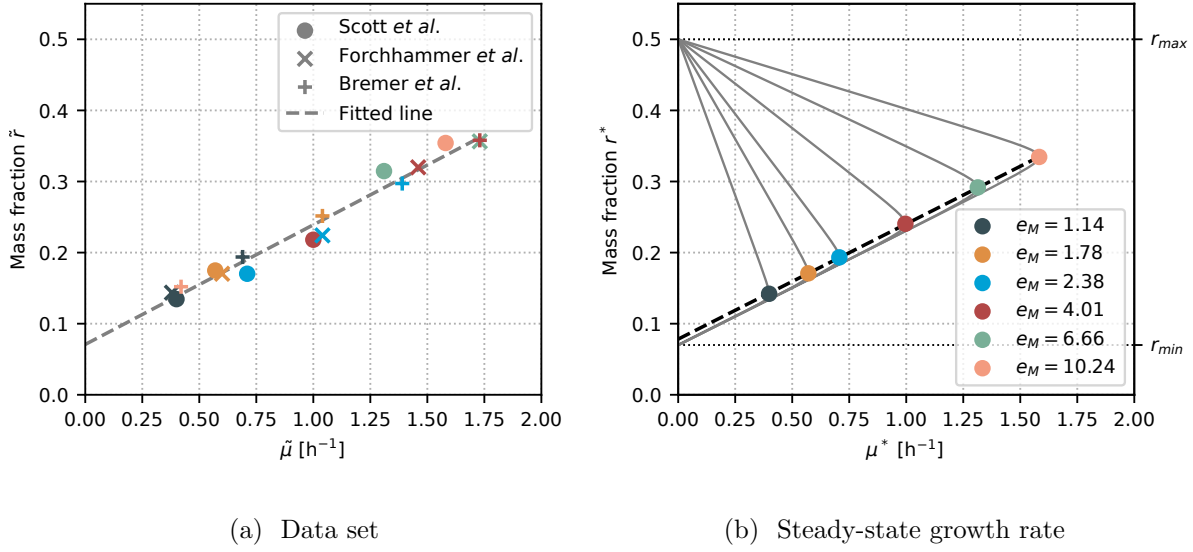


Figure 2: Experimental data from [6, 29, 30] plotted in (a) shows a linearity of  $r^2 = 0.9739$  (dashed line, fitted to data) with a vertical intercept  $r_{\min} = 0.07$  and slope  $k_R = 6.23 \text{ h}^{-1}$ . In (b), steady-state growth rate curves  $\mu^*$  are shown in terms of the mass fraction  $r^* \in (r_{\min}, r_{\max})$  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.

444 enabling them to maximize their biomass [31, 32]. This is represented by an OCP (Optimal  
 445 Control Problem) where the objective is to maximize the final volume at time  $T$  given by  
 446  $\mathcal{V}(T)$ . For the sake of convenience, we propose to maximize the quantity  $\log \mathcal{V}(T)$  (since  $\log$   
 447 is an increasing function) given by

$$448 \quad \log \mathcal{V}(T) = \int_0^T \mu(p, r) dt + \log \mathcal{V}(0).$$

449

450 As the initial condition  $\mathcal{V}(0)$  is fixed, we define the cost function

$$451 \quad J(u) \doteq \int_0^T \mu(p, r) dt = \int_0^T w_R(p)(r - r_{\min}) dt.$$

452

453 Since  $\mathcal{V}$  appears neither in the dynamics nor in the cost function, the optimal problem will  
 454 be written considering the reduced system introduced in (S') with initial conditions given by

455 (IC). We write the optimal control problem

$$456 \quad (\text{OCP}) \quad \left\{ \begin{array}{l} \text{maximize} \quad J(u) = \int_0^T w_R(p)(r - r_{\min}) dt \\ \text{subject to} \quad \text{dynamics (S')}, \\ \quad \quad \quad \text{initial conditions (IC)}, \\ \quad \quad \quad \alpha(\cdot) \in \mathcal{U}, \end{array} \right.$$

457

458 with  $\mathcal{U}$  the set of admissible controllers, which are Lebesgue measurable real-valued functions  
 459 defined on the interval  $[0, T]$  and satisfying the constraint  $\alpha(t) \in [0, 1]$ . Problem (OCP) has  
 460 neither final state constraints nor path constraints. In the context of dynamical optimization,  
 461 the use of path constraints can be useful to restrict the solutions to those meeting certain  
 462 physical and biological limitations, especially when dealing with more complex models. While  
 463 enforcing additional constraints on the OCP increases the dimension of the problem, standard  
 464 optimal control solvers are able to handle such formulations. In this work, imposing initial  
 465 conditions (IC) guarantees that every trajectory of the system stays within the set  $\Gamma$  defined  
 466 in Lemma 3.1, which ensures that the solutions are consistent with the biological assumptions.  
 467 In principle, this formulation of the problem resembles the optimal control problem proposed  
 468 in [10]: the objective is to maximize the accumulation of a certain quantity within the system  
 469 during a fixed time interval  $[0, T]$ . The main difference lies in the dynamics of the system,  
 470 as the introduction of the protein Q increases the system dimension by one, which yields a  
 471 more relevant (and more complex) associated OCP. We will see in following sections that the  
 472 problem raised in this work can be solved by a generalization of Giordano *et al.*'s approach.

473 **5.2. Pontrjagin Maximum Principle.** Existence of a solution for this class of OCPs is  
 474 rather trivial. Given that there are no terminal constraints, there is no controllability issue.  
 475 Moreover, the dynamics is affine in the control with the latter included in a compact and con-  
 476 vex set (a closed interval), and one can easily check that every finite-time trajectory remains  
 477 bounded. So existence is guaranteed by Filippov's theorem [33]. Then, for an optimal con-  
 478 trol problem (OCP) with state  $\varphi \in \mathbb{R}^n$ , Pontrjagin maximum principle (PMP) ensures that  
 479 there exist  $\lambda^0 \leq 0$  and a piecewise absolutely continuous mapping  $\lambda(\cdot) : [0, T] \rightarrow \mathbb{R}^n$ , with  
 480  $(\lambda(\cdot), \lambda^0) \neq (0, 0)$ , such that the extremal  $(\varphi, \lambda, \lambda^0, \alpha)$  satisfies the generalized Hamiltonian  
 481 system

$$482 \quad (\text{PMP}) \quad \left\{ \begin{array}{l} \dot{\varphi} = \frac{\partial}{\partial \lambda} H(\varphi, \lambda, \lambda^0, \alpha), \\ \dot{\lambda} = -\frac{\partial}{\partial \varphi} H(\varphi, \lambda, \lambda^0, \alpha), \\ H(\varphi, \lambda, \lambda^0, \alpha) = \max_{\alpha \in [0, 1]} H(\varphi, \lambda, \lambda^0, \alpha), \end{array} \right.$$

483

484 for almost every  $t \in [0, T]$ . For our particular case, we have the state vector  $\varphi \doteq (p, r, m)$  and  
 485 adjoint vector  $\lambda \doteq (\lambda_p, \lambda_r, \lambda_m)$  and the Hamiltonian given by

$$486 \quad (5.1) \quad H(\varphi, \lambda, \lambda^0, \alpha) = \lambda^0 w_R(p)(r - r_{\min}) + \langle \lambda, F(\varphi, u) \rangle,$$

487



488 where  $F$  represents the right-hand side of system (S'). Given that in (OCP) there is no terminal  
 489 condition on the state  $\varphi(T)$ , the transversality condition for the adjoint state is  $\lambda(T) = 0$ ,  
 490 and we can discard abnormal extremals from the analysis. In other words, any extremal  
 491  $(\varphi, \lambda, \lambda^0, \alpha)$  satisfying PMP is normal, so  $\lambda^0 \neq 0$ . Developing (5.1) yields the Hamiltonian

$$492 \quad H = \left( E_M m - (p+1)w_R(p)(r - r_{\min}) \right) \lambda_p + (r_{\max}\alpha - r)w_R(p)(r - r_{\min})\lambda_r \\ 493 \quad + (r_{\max}(1 - \alpha) - m)w_R(p)(r - r_{\min})\lambda_m - \lambda^0 w_R(p)(r - r_{\min}),$$

495 and the adjoint system is

$$(5.2) \quad \left\{ \begin{array}{l} \dot{\lambda}_p = w_R(p)(r - r_{\min})\lambda_p + (p+1)w'_R(p)(r - r_{\min})\lambda_p - (r_{\max}\alpha - r)w'_R(p)(r - r_{\min})\lambda_r \\ \quad - (r_{\max}(1 - \alpha) - m)w'_R(p)(r - r_{\min})\lambda_m + \lambda^0 w'_R(p)(r - r_{\min}), \\ 496 \quad \dot{\lambda}_r = (p+1)w_R(p)\lambda_p + w_R(p)(r - r_{\min})\lambda_r - (r_{\max}\alpha - r)w_R(p)\lambda_r \\ \quad - (r_{\max}(1 - \alpha) - m)w_R(p)\lambda_m + \lambda^0 w_R(p), \\ 497 \quad \dot{\lambda}_m = -E_M \lambda_p + w_R(p)(r - r_{\min})\lambda_m. \end{array} \right.$$

498 Since the Hamiltonian is linear in the control  $\alpha$ , we rewrite it in the input-affine form  $H =$   
 499  $H_0 + \alpha H_1$  with

$$500 \quad H_0 = \left( E_M m - (p+1)w_R(p)(r - r_{\min}) \right) \lambda_p - r w_R(p)(r - r_{\min})\lambda_r \\ 501 \quad + \left( r_{\max} - m \right) w_R(p)(r - r_{\min})\lambda_m - \lambda^0 w_R(p)(r - r_{\min}), \\ 502 \quad (5.3) \quad H_1 = r_{\max} w_R(p)(r - r_{\min})(\lambda_r - \lambda_m).$$

504 The constrained optimal control  $\alpha$  should maximize the Hamiltonian, so the solution is

$$505 \quad \alpha(t) = \begin{cases} 0 & \text{if } H_1 < 0, \\ 1 & \text{if } H_1 > 0, \\ 506 \quad \alpha_{\text{sing}}(t) & \text{if } H_1 = 0, \end{cases}$$

507 where  $\alpha_{\text{sing}}(t)$  is called a singular control, showing that any optimal control is a concatenation  
 508 of bangs ( $\alpha = \pm 1$ ) and singular arcs, depending on the sign of the switching function  $H_1$ . As  
 509 obtained in [21, 23], a bang arc  $\alpha = 0$  (respectively  $\alpha = 1$ ) corresponds to a pure allocation  
 510 strategy where the production of  $R$  (respectively  $M$ ) is completely switched off. While a full  
 511 description of the optimal control is often difficult to obtain through PMP, there are certain  
 512 analyses that can be performed to help understand its structure. We will first see that the  
 513 final bang of the optimal control is an upper bang  $\alpha = 1$ .

514 **Lemma 5.1.** *There exists  $\epsilon$  such that the optimal control solution of (OCP) is  $\alpha(t) = 1$  for*  
 515 *the interval of time  $[T - \epsilon, T]$ .*

516 *Proof.* We define  $\lambda_z = \lambda_r - \lambda_m$ , where its dynamics can be obtained from (5.2). It can be  
517 seen that, when evaluating its dynamics at final time, we get

$$518 \quad \dot{\lambda}_z(T) = \lambda^0 w_R(p(T)) < 0,$$

520 due to the whole adjoint state being null at final time except for  $\lambda^0$ . As  $\lambda_z(T)$  also vanishes  
521 due to the transversality conditions, we have  $\lambda_z(T - \epsilon) > 0$  for a certain  $\epsilon$ . Then,  $H_1 > 0$  for  
522 the interval  $[T - \epsilon, T]$ , which corresponds to a bang arc  $\alpha = 1$ . ■

523 A control  $\alpha = 1$  implies a strategy in which all resources are allocated to ribosome synthe-  
524 sis, thus favoring the synthesis of proteins. An intuitive interpretation of Lemma 5.1 is that,  
525 when approaching the final time  $T$ , the most efficient strategy is to exploit as much as possible  
526 the available precursors. This is achieved by maximizing the proteins catalyzing  $v_R$ , at the  
527 expense of arresting the uptake of nutrients  $v_M$  from the environment. In order to further  
528 describe the optimal control, we can analyze the singular extremals. A singular arc occurs  
529 when the switching function  $H_1$  vanishes over a subinterval of time. A detailed description  
530 of the singular arcs can be done by differentiating succesively the switching function  $H_1$  until  
531 the singular control  $\alpha_{\text{sing}}$  can be obtained as a function of the state  $\varphi$  and the adjoint state  $\lambda$ .

### 532 5.3. Study of the singular arcs.

533 **5.3.1. Introduction.** We assume  $H_1$  vanishes on a whole sub-interval  $[t_1, t_2] \subset [0, T]$ , so  
534 the extremal belongs to the singular surface

$$535 \quad \Sigma \doteq \{(\varphi, \lambda) \in \mathbb{R}^6 : H_1(\varphi, \lambda) = 0\}.$$

537 Since  $H_1$  vanishes identically, so does its derivative with respect to time. Differentiating along  
538 an extremal  $(\varphi, \lambda)$  amounts to taking a Poisson bracket<sup>2</sup> with the Hamiltonian  $H$  [33]. Indeed,  
539 along the singular arc,

$$540 \quad 0 = \dot{H}_1 = \frac{\partial H_1}{\partial \varphi} \dot{\varphi} + \frac{\partial H_1}{\partial \lambda} \dot{\lambda} = \sum_{i=1}^n \left( \frac{\partial H}{\partial \lambda_i} \frac{\partial H_1}{\partial \varphi_i} - \frac{\partial H}{\partial \varphi_i} \frac{\partial H_1}{\partial \lambda_i} \right) = \{H, H_1\} = \{H_0, H_1\}.$$

541 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  
542 Lie bracket of the vector fields  $F_0$  and  $F_1$ . Differentiating again we obtain

$$543 \quad 0 = \dot{H}_{01} = H_{001} + \alpha H_{101}.$$

544 Again,  $H_{001} \doteq \langle \lambda, F_{001} \rangle$  where, with the same notation as before,  $F_{001}$  is the Lie bracket of  $F_0$   
545 with  $F_{01}$ . If, on the set

$$546 \quad \Sigma' \doteq \{(\varphi, \lambda) \in \mathbb{R}^6 : H_1(\varphi, \lambda) = H_{01}(\varphi, \lambda) = 0\},$$

---

<sup>2</sup>The Poisson bracket  $\{f, g\}$  of two functions  $f$  and  $g$  along an extremal  $(\varphi, \lambda)$  is defined as

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

547 the bracket  $H_{101}$  is also zero, the control disappears from the previous equality, and one has  
 548 to differentiate at least two more times to retrieve the control:  $H_{0001}$  is also zero, and

$$549 \quad (5.4) \quad 0 = H_{00001} + \alpha H_{10001}.$$

550 If the length-five bracket  $H_{10001}$  is not zero, the singular arc is of *order two*. When  $H_{101}$   
 551 vanishes not only on  $\Sigma'$  but on all  $\mathbb{R}^6$ , the order is said to be *intrinsic* and connections  
 552 between bang and singular arcs can only occur through an infinite number of switchings [34],  
 553 the so-called Fuller phenomenon. Otherwise, the order is termed *local*, and Fuller phenomenon  
 554 may or may not occur. Using (5.4), the singular control  $u_s$  is obtained as a function of both  
 555 the state  $\varphi$  and the adjoint state  $\lambda$  as

$$556 \quad \alpha_s(\varphi, \lambda) \doteq -\frac{H_{00001}}{H_{10001}}.$$

557 In our low-dimensional situation, there exists the possibility that the singular control is in  
 558 feedback form, that is, as a function of the state only. The latter can be verified by rewriting  
 559 the system in dimension four (Mayer optimal control formulation where the final volume is  
 560 maximized), in terms of  $\tilde{\varphi} \doteq (p, r, m, \mathcal{V})$  and its adjoint  $\tilde{\lambda} \doteq (\lambda_p, \lambda_r, \lambda_m, \lambda_{\mathcal{V}})$ . The dynamics is  
 561 affine in the control,

$$562 \quad \dot{\tilde{\varphi}} = \tilde{F}_0(\tilde{\varphi}) + \alpha \tilde{F}_1(\tilde{\varphi}),$$

563 and so is the Hamiltonian:

$$564 \quad \tilde{H}(\tilde{\varphi}, \tilde{\lambda}, \alpha) = \tilde{H}_0 + \alpha \tilde{H}_1,$$

565 with  $\tilde{H}_i = \langle \tilde{\lambda}, \tilde{F}_i \rangle$ ,  $i = 0, 1$ . The same computation as before leads to the following relations  
 566 along a singular arc of order two:

$$567 \quad 0 = \dot{\tilde{H}}_1 = \dot{\tilde{H}}_{01} = \dot{\tilde{H}}_{001} = \dot{\tilde{H}}_{0001},$$

568 and

$$569 \quad 0 = \tilde{H}_{00001} + \alpha \tilde{H}_{10001}.$$

570 **Proposition 5.2.** *Assume that, for all  $\varphi$ ,  $\tilde{F}_1$ ,  $\tilde{F}_{01}$ , and  $\tilde{F}_{001}$  are independent. Then, an*  
 571 *order two singular control depends only on the state  $\tilde{\varphi}$ , and can be expressed as*

$$572 \quad \alpha_s(\tilde{\varphi}) = -\frac{\det\left(\tilde{F}_1, \tilde{F}_{01}, \tilde{F}_{001}, \tilde{F}_{00001}\right)}{\det\left(\tilde{F}_1, \tilde{F}_{01}, \tilde{F}_{001}, \tilde{F}_{10001}\right)}.$$

573 *Proof.* The previous relations imply that  $\tilde{\lambda}$  is orthogonal to  $\tilde{F}_1$ ,  $\tilde{F}_{01}$ ,  $\tilde{F}_{001}$ , and also to  
 574  $\tilde{F}_{00001} + \alpha \tilde{F}_{10001}$ . If these four vector fields were independent at some point along the singular

575 arc,  $\tilde{\lambda} \in \mathbb{R}^4$  would vanish: for a problem in Mayer form, this would contradict the maximum  
 576 principle. So their determinant must vanish everywhere along the arc and

$$577 \quad \det \left( \tilde{F}_1, \tilde{F}_{01}, \tilde{F}_{001}, \tilde{F}_{00001} \right) + \alpha \det \left( \tilde{F}_1, \tilde{F}_{01}, \tilde{F}_{001}, \tilde{F}_{10001} \right) = 0.$$

578 If the second determinant was zero, given the rank assumption on the first three vector fields,  
 579  $F_{10001}$  would belong to their span; but this is impossible since it would imply  $H_{10001} = 0$ ,  
 580 contradicting the fact that the singular is of order two. ■

581 Going back to the three-dimensional formulation, one can explicit the computations by  
 582 successively differentiating the expression (5.3).

583 **5.3.2. Singular arc in feedback form.** The condition  $H_1 = 0$  could be a consequence of  
 584 the growth rate  $w_R(p)(r - r_{\min})$  vanishing over the whole interval  $[t_1, t_2]$ . We will see this is  
 585 not possible given the dynamics of the system.

586 **Proposition 5.3.** *The growth rate  $\mu(p, r) = w_R(p)(r - r_{\min})$  cannot vanish along the optimal*  
 587 *solution of (OCP).*

588 *Proof.* For any trajectory of (S') with initial conditions (IC), control  $\alpha(\cdot) \in \mathcal{U}$  and  
 589  $t \in [0, T]$ , we have  $\dot{p} \leq E_M r_{\max}$ , which means  $p \leq p_{\max}^T \doteq E_M r_{\max} T + p(0)$ . Then,  
 590  $\dot{r} \geq -r_{\max} w_R(p_{\max}^T)(r - r_{\min})$ . Additionally, since  $w_R(p)$  is continuously differentiable, there  
 591 exists  $c$  such that  $cp \geq w_R(p)$ , which means that  $\dot{p} \geq -cp(p_{\max}^T + 1)(r_{\max} - r_{\min})$ . Then,  
 592 at worst, the state  $p$  (respectively,  $r$ ) decays exponentially towards the value 0 (respectively,  
 593  $r_{\min}$ ), which cannot be attained in finite time. ■

594 As a consequence of Proposition 5.3, the condition  $H_1 = 0$  becomes

$$595 \quad \text{(Condition 1)} \quad \lambda_r - \lambda_m = 0.$$

597 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  
 598 of (Condition 1) is

$$599 \quad \text{(Condition 2)} \quad \phi(\varphi, \lambda)w_R(p) - E_M \lambda_p = 0.$$

601 Along a singular arc, the Hamiltonian can be rewritten as

$$602 \quad \text{(5.5)} \quad H = E_M m \lambda_p + \phi(\varphi, \lambda)w_R(p)(r - r_{\min}),$$

604 and, using (Condition 1) and (Condition 2), the adjoint system becomes

$$605 \quad \begin{cases} \frac{d\lambda_p}{dt} = w_R(p)(r - r_{\min})\lambda_p - \phi(\varphi, \lambda)w'_R(p)(r - r_{\min}), \\ \frac{d\lambda_r}{dt} = w_R(p)(r - r_{\min})\lambda_r - \phi(\varphi, \lambda)w_R(p). \end{cases}$$

606

607 **Proposition 5.4.** *Neither  $\phi(\varphi, \lambda)$  nor  $\lambda_p$  can vanish along a singular arc.*

608 *Proof.* According to (Condition 2), if either  $\phi(\varphi, \lambda)$  or  $\lambda_p$  are null, then both of them are  
 609 null. Then, if  $\phi(\varphi, \lambda) = \lambda_p = 0$ , equation (5.5) would imply that the Hamiltonian vanishes  
 610 in  $\Sigma$ , and therefore it would vanish for the whole interval  $[0, T]$  (as it is constant along the  
 611 solution). However, one can see in (5.1) that the Hamiltonian evaluated at final time is  
 612  $-\lambda^0 w_R(p(T))(r(T) - r_{\min})$  which cannot be 0 due to Proposition 5.3 and  $\lambda^0 \neq 0$ . ■

613 We differentiate (Condition 2) w.r.t. time and we get  $\dot{\phi}(\varphi, \lambda)w_R(p) + \phi(\varphi, \lambda)w'_R(p)\dot{p} - E_M\dot{\lambda}_p$   
 614  $= 0$ . Replacing the latter and using Proposition 5.4 allows us to reduce the expression to

$$615 \text{ (Condition 3)} \quad -(r_{\max} - r_{\min})w_R(p)^2 + E_M(m + r - r_{\min})w'_R(p) = 0,$$

617 which allows us to express  $m + r$  in terms of  $p$ .

618 **Lemma 5.5.** *Along a singular arc over the interval  $[t_1, t_2]$ ,*

$$619 \quad m + r = x(p)$$

621 *with  $x(p) : \mathbb{R}_+ \rightarrow [r_{\min}, \infty)$  defined as*

$$622 \quad x(p) \doteq (r_{\max} - r_{\min}) \frac{w_R(p)^2}{E_M w'_R(p)} + r_{\min},$$

624 *which, using (3.7), yields  $x(p_{\text{opt}}^*) = r_{\max}$ .*

625 The fact that the control does not show up in (Condition 3)—which is obtained by differentiat-  
 626 ing (Condition 1) twice—means that the singular arc is *at least* of order two. We differentiate  
 627 (Condition 3) and we get

$$628 \text{ (Condition 4)} \quad \left( r_{\max} - x(p) + (p+1)x'(p) \right) w_R(p)(r - r_{\min}) - E_M m x'(p) = 0.$$

630 We define the function

$$631 \text{ (5.6)} \quad y(p) \doteq w_R(p) \left( r_{\max} - x(p) + (p+1)x'(p) \right).$$

633 Using (Condition 3) and (5.6) in (Condition 4) yields

$$634 \quad (x(p) - r_{\min})y(p) - \left( E_M x'(p) + y(p) \right) m = 0,$$

636 which means we can express  $m$  and  $r$  in terms of  $p$  along the singular arc.

637 **Lemma 5.6.** *Along a singular arc over the interval  $[t_1, t_2]$ ,*

$$638 \text{ (5.7)} \quad m = (x(p) - r_{\min}) \frac{y(p)}{E_M x'(p) + y(p)},$$

$$639 \text{ (5.8)} \quad r = x(p) - (x(p) - r_{\min}) \frac{y(p)}{E_M x'(p) + y(p)}.$$

640

641 We differentiate (Condition 4) and we get

$$\begin{aligned}
642 \quad (5.9) \quad & -(r_{\max}(1 - \alpha) - m)w_R(p)(r - r_{\min}) + x'(p)\frac{y(p)}{E_M x'(p) + y(p)}\dot{p} \\
643 \quad & + (x(p) - r_{\min})\left(\frac{y'(p)}{E_M x'(p) + y(p)} - \frac{y(p)}{(E_M x'(p) + y(p))^2}(E_M x''(p) + y'(p))\right)\dot{p} = 0,
\end{aligned}$$

644 meaning that we can express

$$645 \quad \alpha_{\text{sing}}(p) = 1 - \frac{m}{r_{\max}} \left( \left( \frac{x'(p)}{x(p) - r_{\min}} + \frac{y'(p)}{y(p)} - \frac{E_M x''(p) + y'(p)}{E_M x'(p) + y(p)} \right) \frac{\dot{p}}{w_R(p)(r - r_{\min})} + 1 \right).$$

647 While (Condition 3) showed that the order of the singular arc is *at least* two, the latter relation  
648 proves that it is *exactly* two. Indeed, the coefficient before  $\alpha$  in (5.9) is  $-r_{\max}w_R(p)(r - r_{\min})$ ,  
649 which cannot vanish as proven in Proposition 5.3. The singular arc is said to be *locally of order*  
650 *two*, as the coefficient of  $\alpha$  in (Condition 3) is zero along the singular arc, but not everywhere  
651 on the cotangent bundle [34]. In this case, the presence of the Fuller phenomenon (i.e., the  
652 junctions between bang and singular arcs constituting an infinite number of switchings) is  
653 not guaranteed. However, this turns out to be the case as it will be shown in the numerical  
654 computations. Besides, in accordance with Proposition 5.2, the order two singular control can  
655 be expressed in feedback form, i.e., as a function of the state only. We performed a numerical  
656 rank test using Singular Value Decomposition, which confirmed that the rank condition is  
657 fulfilled. More precisely, the actual computation proves that the singular control can be  
658 expressed as a function of  $p$  only (Lemma 5.6 entails that  $r$ ,  $m$  and therefore  $\dot{p}$  can be expressed  
659 in terms of  $p$ ), which allows to retrieve the turnpike behaviour as described in the following  
660 section.

661 **5.3.3. The turnpike phenomenon.** Using (5.7) and (5.8), we see that, along a singular  
662 arc, the dynamical equation of  $p$  becomes

$$663 \quad \dot{p} = E_M w_R(p) \frac{x(p) - r_{\min}}{E_M x'(p) + y(p)} (r_{\max} - x(p)),$$

665 which is only equal to 0 when  $r_{\max} = x(p)$ . This is only true at  $p = p_{\text{opt}}^*$ , and so

$$666 \quad \text{sign}(\dot{p}) = \text{sign}(p_{\text{opt}}^* - p),$$

668 meaning that, in a singular arc over the interval  $[t_1, t_2]$ , the concentration  $p$  converges asymp-  
669 totically to the optimal value  $p_{\text{opt}}^*$ . This means that  $m$  and  $r$  would also converge to the optimal  
670 values  $m_{\text{opt}}^*$  and  $r_{\text{opt}}^*$ , respectively, and the singular control  $\alpha_{\text{sing}}$  to  $\alpha_{\text{opt}}^*$ . We formalize this in  
671 the following theorem.

672 **Theorem 5.7.** *On a singular arc, the system states and singular control tend asymptotically*  
673 *to*

$$\begin{aligned}
674 \quad (p, r, m) &= (p_{\text{opt}}^*, r_{\text{opt}}^*, m_{\text{opt}}^*), \\
675 \quad \alpha_{\text{sing}}(t) &= \alpha_{\text{opt}}^*.
\end{aligned}$$

677 The above theorem is an explicit proof of the presence of the turnpike property: an opti-  
 678 mal control characterized by a singular arc that stays exponentially close to the steady-state  
 679 solution of the static optimal control problem [35]. This phenomenon has been considerably  
 680 studied in econometry [36], and more recently in biology [37, 10, 20]. It has been shown that,  
 681 for large final times, the trajectory of the system spends most of the time near the optimal  
 682 steady state, and that in infinite horizon problems, it converges to this state.

683 **5.4. Numerical results.** The computations of the optimal trajectories were performed  
 684 with Bocop [38], which solves the optimal control problem through a direct method. An  
 685 online version of the numerical computations can be visualized and executed on the gallery  
 686 of the `ct` (Control Toolbox) project<sup>3</sup>. The time discretization algorithm used is Lobato IIIC  
 687 (implicit, 4-stage, order 6) with 2000 time steps. Figure 3 shows an optimal trajectory with  
 688  $r(0) + m(0) < r_{\max}$ , where most of the bacterial mass corresponds to class  $Q$  proteins. The  
 689 obtained optimal control confirms the conclusions of the latter section: a large part of the  
 690 time, the optimal control remains near the optimal steady-state allocation  $\alpha_{\text{opt}}^*$ , according  
 691 to the turnpike theory (Theorem 5.7). The solution presents chattering after and before the  
 692 singular arc, as expected in the presence of Fuller’s phenomenon (even if only a finite number  
 693 of bangs is computed by the numerical method), and the final bang corresponds to  $\alpha = 1$   
 694 (Lemma 5.1). In order to verify the optimality of the singular arc, we performed a numerical  
 695 computation of the derivatives of  $H_1$ , which is shown in Figure 4. The fact that the factor  
 696 of  $\alpha$  in the fourth derivative is different from 0 confirms that the singular arc is of order 2.  
 697 Moreover, its negativity complies with the *generalized Legendre-Clebsch* condition given by

$$698 \quad (5.10) \quad (-1)^k \frac{\partial}{\partial \alpha} \left( \frac{d^{2k}}{dt^{2k}} H_1 \right) < 0,$$

700 along the singular arc, which is a necessary condition for optimality. As we state in [23], even  
 701 if there exist no available sufficient condition to verify local optimality of extremals with Fuller  
 702 arcs, a check of the Legendre-Clebsch condition along the singular arc can ensure that the  
 703 extremal obtained is not a too crude local minimizer. For the second-order singular arc case,  
 704 the condition corresponds to the case  $k = 2$ . The initial conditions used in Figure 3 were only  
 705 chosen to confirm the theoretical results found throughout this section, by emphasizing the  
 706 main features of the solution. However, from a biological perspective, a situation where  $r + m$   
 707 is significantly different from its steady-state value  $r_{\max}$  is not to be expected: a common  
 708 assumption in these classes of coarse-grained models is that the transcription of  $Q$  proteins  
 709 is autoregulated around stable levels [39], which translates into a constant  $q = 1 - r_{\max}$  (and  
 710 therefore  $m + r = r_{\max}$ ) for the whole interval  $[0, T]$ . We will see in next section that this  
 711 hypothesis produces a very particular structure of the optimal control solution.

712 **6. Biologically relevant scenarios.** Despite their simplicity, self-replicator models have  
 713 been capable of accounting for a number of observable phenomena during steady-state mi-  
 714 crobial growth, under the assumption that bacteria allocate their resources in such a way  
 715 as to maximize growth. Here, we apply the general optimal allocation strategy derived in  
 716 the previous section to predict the bacterial response to certain environmental changes. We

---

<sup>3</sup><https://ct.gitlabpages.inria.fr/gallery/bacteria/bacteria.html>

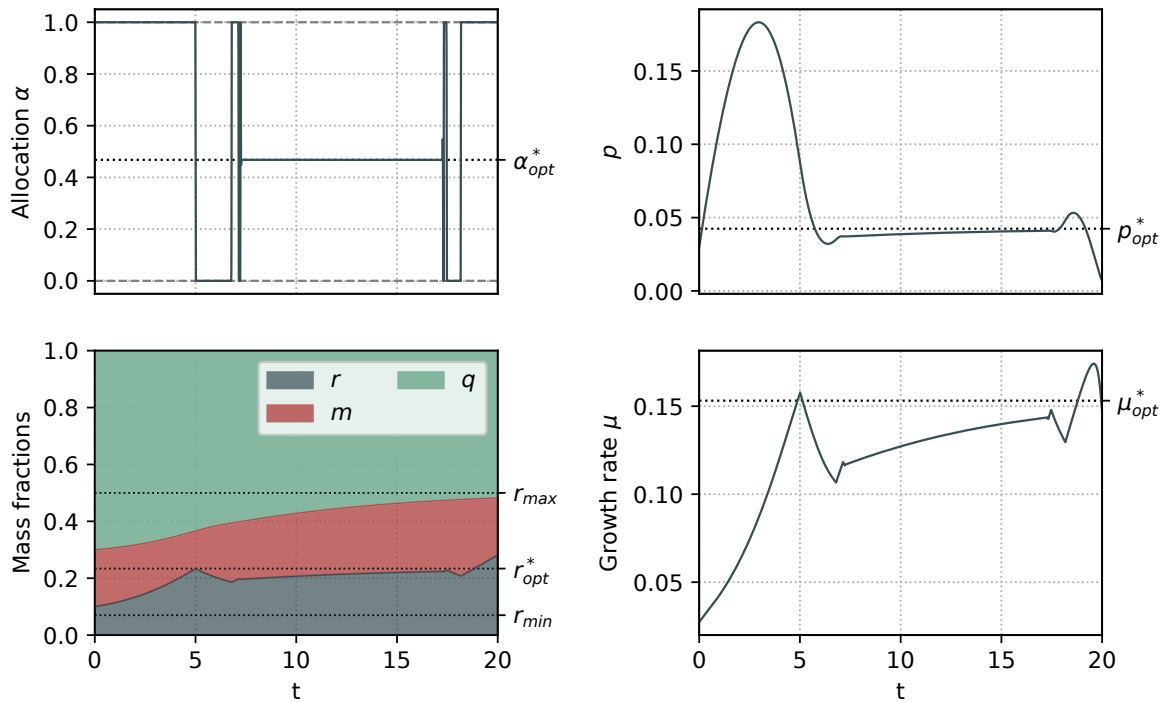


Figure 3: Numerical simulation of (OCP) obtained with Bocop, for the parameter values derived in Section 4. Initial state is  $p(0) = 0.03$ ,  $r(0) = 0.1$ ,  $m(0) = 0.2$  with  $E_M = 0.6$ . As predicted, the optimal control  $\alpha$  involves chattering after and before the singular arc. The mass 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  $(p_{opt}^*, r_{opt}^*, m_{opt}^*)$ .

717 consider two situations that commonly affect bacteria: changes in the nutrient concentration  
 718 in the medium, and changes in the environment submitting the cell to a particular stress.

719 **6.1. Nutrient shift.** Bacteria are known to traverse different habitats throughout their  
 720 lifetime, experiencing fluctuating nutrient concentrations in the medium. In [10], we explored  
 721 how bacteria dynamically adjust their allocation strategy when facing a nutrient upshift. In  
 722 this work, we show that considering a class of growth rate-independent proteins in the model  
 723 refines these previous results. We consider the optimal control problem with the initial state  
 724 being the optimal steady state for a low value of  $E_M$ , and we set a higher  $E_M$  for the time  
 725 interval  $[0, T]$ , representing a richer medium. Setting initial conditions at steady state has an  
 726 impact on the singular arc of the optimal control: it holds that  $m + r = r_{max}$  and  $q = 1 - r_{max}$   
 727 for the whole trajectory, which yields a constant singular arc.

728 **Theorem 6.1.** *If  $r(0) + m(0) = r_{max}$  (i.e.,  $q$  starts from a steady-state value), then any*  
 729 *singular arc over the interval  $[t_1, t_2]$  of the optimal control corresponds to the optimal steady*  
 730 *state.*



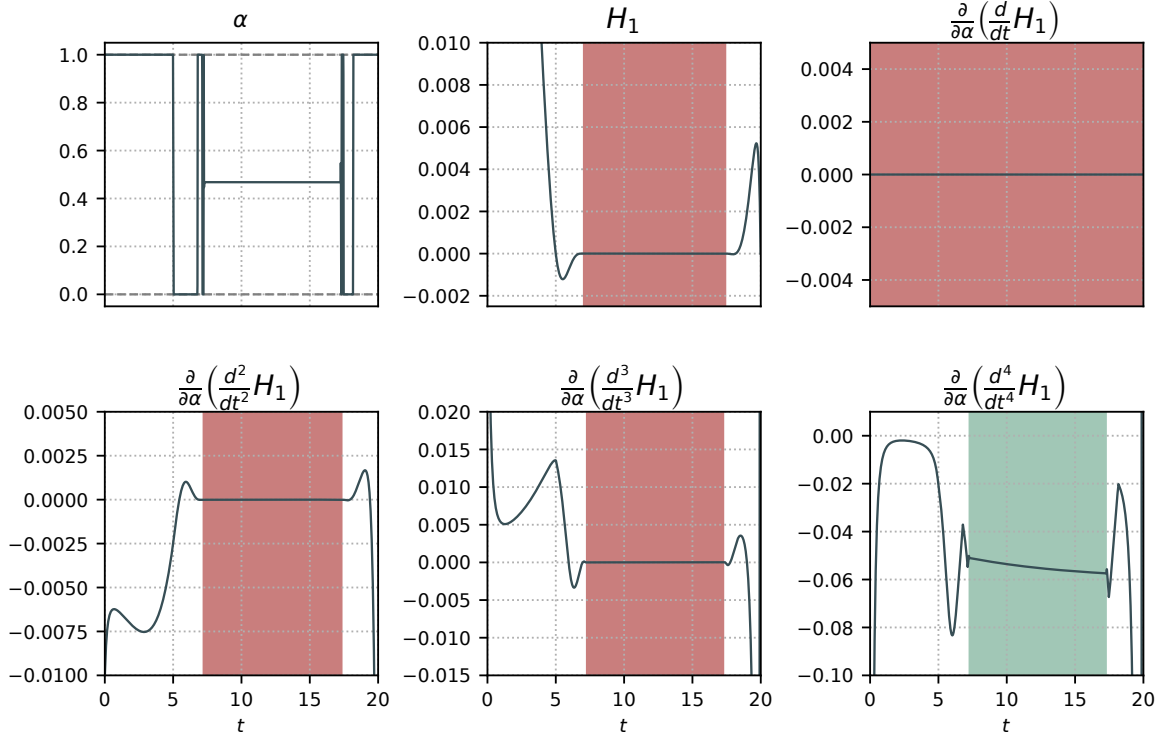


Figure 4: Factors of  $\alpha$  in the derivatives of  $H_1$  evaluated over the trajectory plotted in Figure 3. The intervals where the functions vanish are marked in red. As expected, all functions vanish along the singular arc except for the factor in the fourth derivative (highlighted in green) which is negative according to the Legendre-Clebsch condition (5.10).

731 *Proof.* The dynamical equation for  $q$  is  $\dot{q} = ((1 - r_{\max}) - q)w_R(p)(r - r_{\min})$ , where it can  
 732 be seen that the set  $q = 1 - r_{\max}$  is invariant. This means that, for any trajectory emanating  
 733 from a steady state,  $q$  remains constant even under changes of the nutrient quality  $E_M$ . Then,  
 734 by using the relation (2.5), we obtain

$$735 \quad (6.1) \quad m + r = r_{\max}.$$

737 Along the singular arc, it holds that  $m + r = x(p)$ , which, using (6.1), implies that  $p = p_{\text{opt}}^*$ ,  
 738 meaning that the precursor concentration along the singular arc is constant and optimal.  
 739 Then,  $\alpha_{\text{sing}} = \alpha_{\text{opt}}^*$ ,  $m = m_{\text{opt}}^*$  and  $r = r_{\text{opt}}^*$  for the whole singular arc. ■

740 A numerical simulation of this scenario is shown in Figure 5. As expected, the increase  
 741 in  $E_M$  produces a higher ribosomal mass fraction  $r$ , which translates into an increase of the  
 742 growth rate, stabilizing at the maximal steady-state growth rate  $\mu_{\text{opt}}^*$  through an oscillatory  
 743 phase. It is noteworthy that, in comparison to Giordano *et al.*'s model, the relative changes  
 744 in mass fractions  $r$  and  $m$  are much lower, which corresponds well with the relative changes  
 745 observed in [6]. Additionally, while the presence of  $r_{\min}$  does not noticeably affect the solution

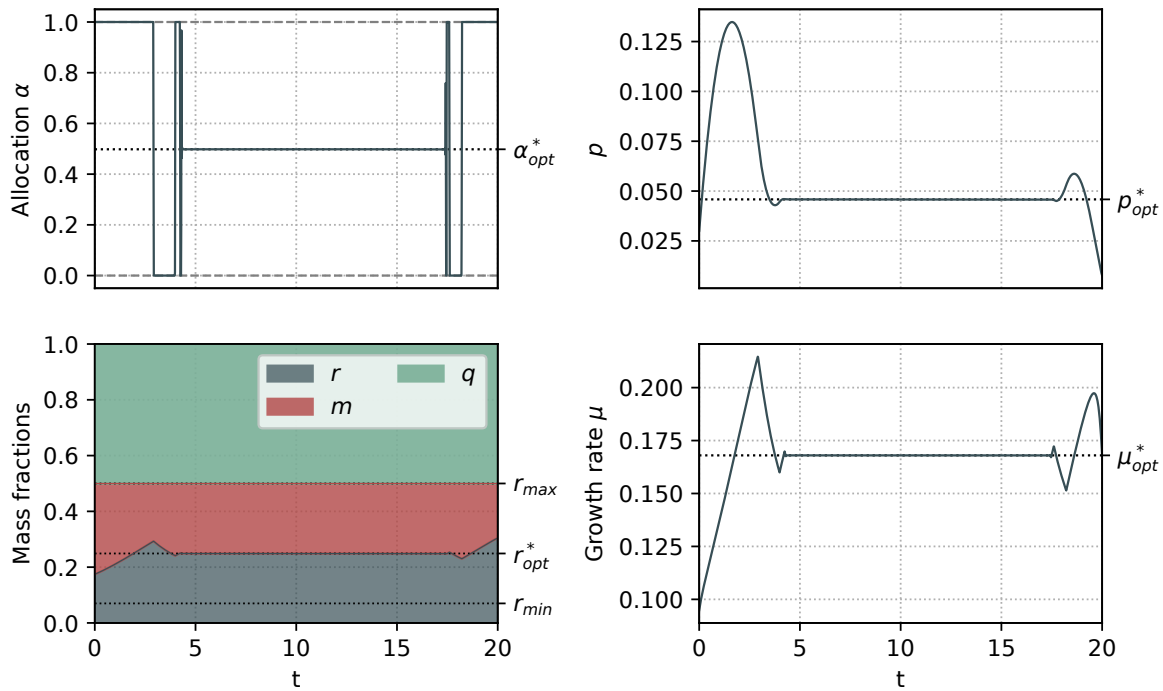


Figure 5: Numerical simulation of the optimal control problem starting from a steady state. The initial state corresponds to the optimal steady state for  $E_M = 0.3$  (poor medium), and the new environmental constant is fixed to  $E_M = 0.7$  (rich medium). As predicted,  $m + r$  ( $= 1 - q$ ) remains 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  $\mu_{opt}^*$  with respect to the growth rate before the upshift.

746 of the optimal control problem, it contributes to a model that more accurately reproduces the  
 747 experimental data (Figure 2a), representing a significant improvement from the modeling  
 748 point of view.

749 **6.2. Bacterial response to stress.** The other scenario of interest is an environmental  
 750 change imposing a certain stress on the microbial population, which is counteracted through  
 751 the synthesis of a stress response protein  $W$ . This protein is also growth rate-independent like  
 752  $Q$ , and its production can be triggered by many different situations. For instance, when subject  
 753 to extreme temperatures, the production of so-called molecular chaperones helps bacteria  
 754 counter the effect of protein unfolding [40, 14]. Likewise, the production of other proteins is  
 755 known to protect bacteria like *E. coli* against acid stress [41]. Another possible scenario is the  
 756 response to metabolic load imposed by the induced overexpression of a heterologous protein  
 757 [42]. All of these situations are known to reduce the resources available for growth-associated  
 758 proteins (Figure 6), consequently decreasing the maximal growth rate attainable. Here, we

759 model a general stress response through the production of the W protein that takes up a  
 760 fraction  $w$  of the proteome, thus reducing  $r_{\max}$  to a certain  $r_{\max}^w < r_{\max}$ .

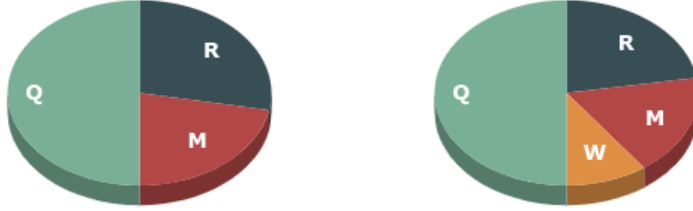


Figure 6: Left: original case. Right: new proposed case, where  $q$  remains unchanged, but the maximal allocation  $m + r$  is restricted to a  $r_{\max}^w < r_{\max}$ .

761 As before, we assume  $q$  takes up a constant fraction  $1 - r_{\max}$  of the proteome, but the  
 762 proportions of resources allocated to M and R are now  $r_{\max}^w \alpha$  and  $r_{\max}^w (1 - \alpha)$  respectively.  
 763 By construction, we have  $w = r_{\max} - m - r$ , which means we can express

$$764 \quad \dot{w} = (r_{\max} - r_{\max}^w - w) w_R(p)(r - r_{\min}),$$

766 showing that the mass fraction  $w$  converges asymptotically to the difference  $r_{\max} - r_{\max}^w$ . The  
 767 remaining mass fractions  $p$ ,  $r$  and  $m$  obey the dynamics of system (S'), so the application of  
 768 the optimal solution found in last section is straightforward. An example is shown in Figure  
 769 7. As predicted,  $m + r$  converges to the reduced  $r_{\max}^w$ ,  $q$  remains constant at  $1 - r_{\max}$  and  $w$   
 770 converges to  $r_{\max}^w - r_{\max}$ . The reduction of resources available for growth-associated proteins  
 771 (M and R) causes the growth rate to drop, as was shown experimentally [6].

772 **7. Conclusion.** In this work, we proposed a dynamical self-replicator model of bacterial  
 773 growth based on the work of [10], which introduces a growth rate-independent class of pro-  
 774 tein. As a consequence, the proteome of the bacterial cell can be divided into the metabolic  
 775 machinery M, the gene expression machinery R, and the housekeeping machinery Q. While Q  
 776 is growth rate-independent, this is also the case for a fraction of R required for cell replication  
 777 to occur. As a consequence of this hypothesis, a maximum ribosomal concentration  $r_{\max}$  ap-  
 778 pears in the model kinetics, limiting the allocation of resources to M and R. We studied the  
 779 asymptotic behavior of the system, showing that, under certain conditions, all solutions con-  
 780 verge towards the only globally attractive equilibrium. We then explored the optimal dynamic  
 781 allocation strategies that consider maximizing the bacterial population volume in terms of the  
 782 resource allocation parameter  $\alpha$ . This involved a study of the static and dynamic aspects  
 783 of optimal strategies. For the first one, we showed there is a unique optimal steady state,  
 784 which corresponds to experimental observations of growing cultures of *E. coli* [29, 30, 6, 27].  
 785 The dynamic problem is approached through optimal control theory, by application of the  
 786 Pontrjagin's Maximum Principle. The obtained optimal control has a Fuller-singular-Fuller  
 787 structure with a non-constant singular arc, in contrast to the constant singular arc obtained  
 788 in Giordano *et al.*'s approach. We performed a detailed analysis of the OCP in both analytic  
 789 and numerical ways. In particular, the singular arc of the optimal solution is characterized

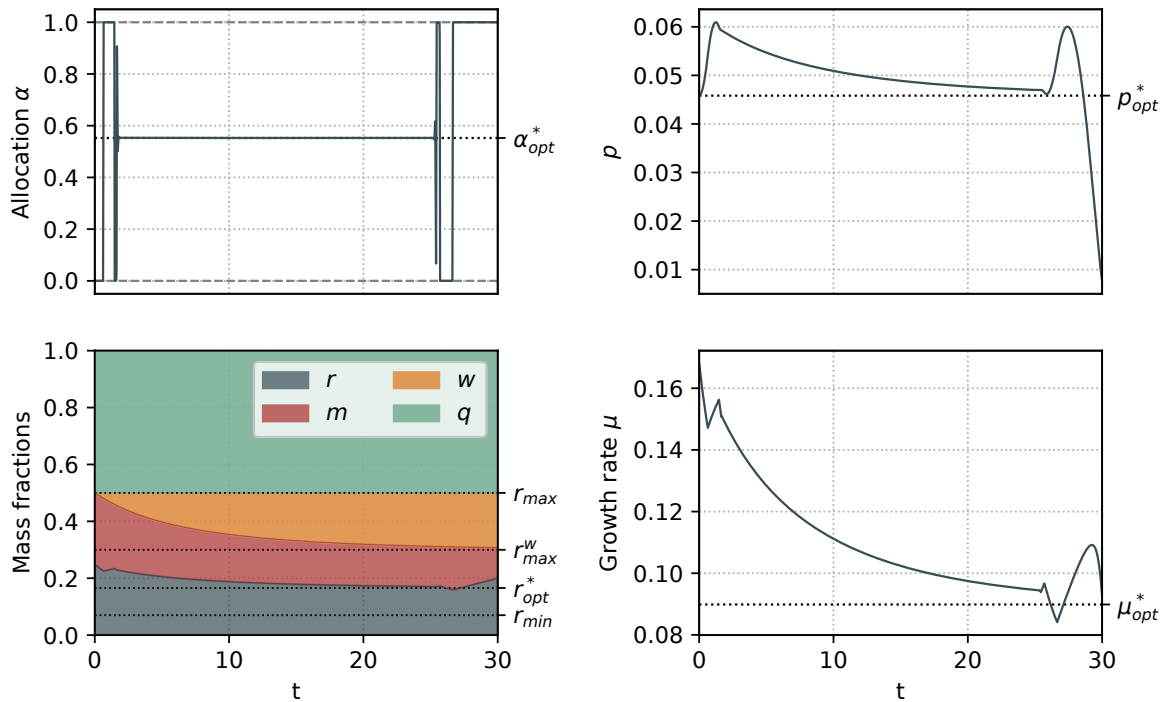


Figure 7: Numerical simulation of an optimal trajectory where the initial conditions are the optimal steady state for  $E_M = 0.7$  and  $r_{\max} = 0.5$ . A certain stress is induced at  $t = 0$ , which triggers the synthesis of the growth rate-independent protein  $w$ , reducing the fraction  $r_{\max}$  to  $r_{\max}^w = 0.3$ . As a result, the steady-state growth rate is significantly reduced.

790 by i) its feedback form (i.e., being expressed as a function of the state only), ii) being exactly  
 791 of order 2, and iii) the turnpike phenomenon (where the state trajectory and optimal control  
 792 converge asymptotically towards the optimal steady state and control). Moreover, we showed  
 793 that, when the mass fraction of class  $Q$  proteins is at steady state, the singular arc of the  
 794 optimal solution corresponds to the optimal steady state. Additionally, we showed that the  
 795 dynamical approach can be used to predict the behavior of the system when subject to stress.  
 796 The latter is modeled through a reduction of the fraction of growth rate-dependent protein  
 797 synthesis as the production of a  $w$  protein that reduces  $r_{\max}$ .

798 While the main features of Giordano *et al.*'s work are present in this approach, our gen-  
 799 eralization shows a better agreement with the experimental data given by the introduction  
 800 of the parameters  $r_{\max}$  and  $r_{\min}$  in the model. Additionally, the proposed partitioning of  
 801 the proteome in a dynamic setting can account for certain natural phenomena known to re-  
 802 duce the fraction of growth rate-dependent proteins in the cell. These modifications yield  
 803 interesting optimal control problems, which could potentially help understand the internal  
 804 decision-making mechanisms evolved by bacteria.

805 Our approach was built on the joint exploitation of theoretical and numerical results.

806 When tackling more complex problems as proposed, e.g., in Tsiantis & Banga [43], a PMP  
 807 perspective tends to yield very complicated mathematical formulations. Using direct methods  
 808 has the advantage of avoiding these issues, but it often requires some knowledge to initialize  
 809 the optimization algorithm or to check the validity of the solutions. In order to investigate  
 810 complex biological systems, we advocate the development and theoretical analysis of simple  
 811 models, in line with the question to be investigated, coupled with numerical exploration of  
 812 optimal solutions (using larger models if necessary).

813 **Acknowledgments.** We would like to acknowledge the help of S. Psalmon and B. Schall  
 814 from [Polytech Nice Sophia](#) for the numerical simulations.

815

## REFERENCES

- 816 [1] M. SCHAECHTER, J. L. INGRAHAM, AND F. C. NEIDHARDT, *Microbe*, ASM Press, Washington, DC,  
 817 2006.
- 818 [2] F. NEIDHARDT, *Bacterial growth: constant obsession with  $dN/dt$* , *Journal of Bacteriology*, 181 (1999),  
 819 pp. 7405–7408.
- 820 [3] S. BRUL, S. VAN GERWEN, AND M. ZWIETERING, *Modelling Microorganisms in Food*, Elsevier, 2007.
- 821 [4] A. BRAUNER, O. FRIDMAN, O. GEFEN, AND N. Q. BALABAN, *Distinguishing between resistance, toler-*  
 822 *ance and persistence to antibiotic treatment*, *Nature Reviews Microbiology*, 14 (2016), pp. 320–330.
- 823 [5] J. C. LIAO, L. MI, S. PONTRELLI, AND S. LUO, *Fuelling the future: microbial engineering for the*  
 824 *production of sustainable biofuels*, *Nature Reviews Microbiology*, 14 (2016), pp. 288–304.
- 825 [6] M. SCOTT, C. W. GUNDERSON, E. M. MATEESCU, Z. ZHANG, AND T. HWA, *Interdependence of cell*  
 826 *growth and gene expression: origins and consequences*, *Science*, 330 (2010), pp. 1099–1102.
- 827 [7] H. VAN DEN BERG, Y. KISELEV, AND M. ORLOV, *Optimal allocation of building blocks between nutrient*  
 828 *uptake systems in a microbe*, *Journal of Mathematical Biology*, 44 (2002), pp. 276–296.
- 829 [8] D. MOLENAAR, R. VAN BERLO, D. DE RIDDER, AND B. TEUSINK, *Shifts in growth strategies reflect*  
 830 *tradeoffs in cellular economics*, *Molecular Systems Biology*, 5 (2009), p. 323.
- 831 [9] A. Y. WEISSE, D. A. OYARZÚN, V. DANOS, AND P. S. SWAIN, *Mechanistic links between cellular trade-*  
 832 *offs, gene expression, and growth*, *Proceedings of the National Academy of Sciences*, 112 (2015),  
 833 pp. E1038–E1047.
- 834 [10] N. GIORDANO, F. MAIRET, J.-L. GOUZÉ, J. GEISELMANN, AND H. DE JONG, *Dynamical allocation*  
 835 *of cellular resources as an optimal control problem: novel insights into microbial growth strategies*,  
 836 *PLoS Computational Biology*, 12 (2016), p. e1004802.
- 837 [11] D. W. ERICKSON, S. J. SCHINK, V. PATSALO, J. R. WILLIAMSON, U. GERLAND, AND T. HWA, *A*  
 838 *global resource allocation strategy governs growth transition kinetics of Escherichia coli*, *Nature*, 551  
 839 (2017), pp. 119–123.
- 840 [12] H. DOURADO AND M. J. LERCHER, *An analytical theory of balanced cellular growth*, *Nature Communi-*  
 841 *cations*, 11 (2020), pp. 1–14.
- 842 [13] M. SCOTT, S. KLUMPP, E. MATEESCU, AND T. HWA, *Emergence of robust growth laws from optimal*  
 843 *regulation of ribosome synthesis*, *Molecular Systems Biology*, 10 (2014), p. 747.
- 844 [14] F. MAIRET, J.-L. GOUZÉ, AND H. DE JONG, *Optimal proteome allocation and the temperature depen-*  
 845 *dence of microbial growth laws*, *npj Systems Biology and Applications*, 7 (2021), pp. 1–11.
- 846 [15] M. BASAN, S. HUI, H. OKANO, Z. ZHANG, Y. SHEN, J. WILLIAMSON, AND T. HWA, *Overflow*  
 847 *metabolism in Escherichia coli results from efficient proteome allocation*, *Nature*, 528 (2015), pp. 99–  
 848 104.
- 849 [16] A. MAITRA AND K. DILL, *Bacterial growth laws reflect the evolutionary importance of energy efficiency*,  
 850 *Proceedings of the National Academy of Sciences of the USA*, 112 (2015), pp. 406–411.
- 851 [17] D. A. CARLSON, A. B. HAURIE, AND A. LEIZAROWITZ, *Infinite Horizon Optimal Control*, Springer,  
 852 Berlin, Heidelberg, 1991.
- 853 [18] I. YEGOROV, F. MAIRET, AND J.-L. GOUZÉ, *Optimal feedback strategies for bacterial growth with degra-*

- 854            *ation, recycling, and effect of temperature*, Optimal Control Applications and Methods, 39 (2018),  
855            pp. 1084–1109.
- 856 [19] J. IZARD, C. G. BALDERAS, D. ROPERS, S. LACOUR, X. SONG, Y. YANG, A. LINDNER, J. GEISELMANN,  
857            AND H. DE JONG, *A synthetic growth switch based on controlled expression of RNA polymerase*,  
858            Molecular Systems Biology, 11 (2015), p. 840.
- 859 [20] I. YEGOROV, F. MAIRET, H. DE JONG, AND J.-L. GOUZÉ, *Optimal control of bacterial growth for the*  
860            *maximization of metabolite production*, Journal of Mathematical Biology, 78 (2019), pp. 985–1032.
- 861 [21] A. G. YABO, J.-B. CAILLAU, AND J.-L. GOUZÉ, *Singular regimes for the maximization of metabolite*  
862            *production*, in 2019 IEEE 58th Conference on Decision and Control (CDC), IEEE, 2019, pp. 31–36.
- 863 [22] A. G. YABO AND J.-L. GOUZÉ, *Optimizing bacterial resource allocation: metabolite production in con-*  
864            *tinuous bioreactors*, IFAC-PapersOnLine, 53 (2020), pp. 16753–16758.
- 865 [23] A. G. YABO, J.-B. CAILLAU, AND J.-L. GOUZÉ, *Optimal bacterial resource allocation: metabolite pro-*  
866            *duction in continuous bioreactors*, Mathematical Biosciences and Engineering, 17 (2020), pp. 7074–  
867            7100.
- 868 [24] J. EWALD, M. BARTL, AND C. KALETA, *Deciphering the regulation of metabolism with dynamic opti-*  
869            *mization: an overview of recent advances*, Biochemical Society Transactions, 45 (2017), pp. 1035–  
870            1043.
- 871 [25] E. KLIPP, R. HEINRICH, AND H.-G. HOLZHÜTTER, *Prediction of temporal gene expression: Metabolic*  
872            *optimization by re-distribution of enzyme activities*, European Journal of Biochemistry, 269 (2002),  
873            pp. 5406–5413.
- 874 [26] D. A. OYARZÚN, B. P. INGALLS, R. H. MIDDLETON, AND D. KALAMATIANOS, *Sequential activation of*  
875            *metabolic pathways: a dynamic optimization approach*, Bulletin of Mathematical Biology, 71 (2009),  
876            p. 1851.
- 877 [27] S. HUI, J. M. SILVERMAN, S. S. CHEN, D. W. ERICKSON, M. BASAN, J. WANG, T. HWA, AND J. R.  
878            WILLIAMSON, *Quantitative proteomic analysis reveals a simple strategy of global resource allocation*  
879            *in bacteria*, Molecular Systems Biology, 11 (2015), p. 784.
- 880 [28] M. BASAN, M. ZHU, X. DAI, M. WARREN, D. SÉVIN, Y.-P. WANG, AND T. HWA, *Inflating bacterial*  
881            *cells by increased protein synthesis*, Molecular Systems Biology, 11 (2015), p. 836.
- 882 [29] J. FORCHHAMMER AND L. LINDAHL, *Growth rate of polypeptide chains as a function of the cell growth*  
883            *rate in a mutant of Escherichia coli 15*, Journal of Molecular Biology, 55 (1971), pp. 563–568.
- 884 [30] H. BREMER, P. P. DENNIS, ET AL., *Modulation of chemical composition and other parameters of the*  
885            *cell by growth rate*, Escherichia coli and Salmonella: Cellular and Molecular Biology, 2 (1996),  
886            pp. 1553–69.
- 887 [31] J. S. EDWARDS, R. U. IBARRA, AND B. O. PALSSON, *In silico predictions of Escherichia coli metabolic*  
888            *capabilities are consistent with experimental data*, Nature Biotechnology, 19 (2001), pp. 125–130.
- 889 [32] N. E. LEWIS, K. K. HIXSON, T. M. CONRAD, J. A. LERMAN, P. CHARUSANTI, A. D. POLPITIYA,  
890            J. N. ADKINS, G. SCHRAMM, S. O. PURVINE, D. LOPEZ-FERRER, ET AL., *Omic data from evolved*  
891            *E. coli are consistent with computed optimal growth from genome-scale models*, Molecular Systems  
892            Biology, 6 (2010), p. 390.
- 893 [33] A. A. AGRACHEV AND Y. SACHKOV, *Control Theory from the Geometric Viewpoint*, vol. 87, Springer  
894            Science & Business Media, 2013.
- 895 [34] M. I. ZELIKIN AND V. F. BORISOV, *Theory of Chattering Control: with Applications to Astronautics,*  
896            *Robotics, Economics, and Engineering*, Springer Science & Business Media, 2012.
- 897 [35] E. TRÉLAT AND E. ZUAZUA, *The turnpike property in finite-dimensional nonlinear optimal control*,  
898            Journal of Differential Equations, 258 (2015), pp. 81–114.
- 899 [36] D. CASS, *Optimum growth in an aggregative model of capital accumulation*, The Review of Economic  
900            Studies, 32 (1965), pp. 233–240.
- 901 [37] J.-M. CORON, P. GABRIEL, AND P. SHANG, *Optimization of an amplification protocol for misfolded*  
902            *proteins by using relaxed control*, Journal of Mathematical Biology, 70 (2015), pp. 289–327.
- 903 [38] I. S. TEAM COMMANDS, *BOCOP: an open source toolbox for optimal control*. <http://bocop.org>, 2017.
- 904 [39] S. KLUMPP, Z. ZHANG, AND T. HWA, *Growth rate-dependent global effects on gene expression in bacteria*,  
905            Cell, 139 (2009), pp. 1366–1375.
- 906 [40] D. RONCARATI AND V. SCARLATO, *Regulation of heat-shock genes in bacteria: from signal sensing to*  
907            *gene expression output*, FEMS Microbiology reviews, 41 (2017), pp. 549–574.

- 908 [41] N. GUAN AND L. LIU, *Microbial response to acid stress: mechanisms and applications*, Applied Micro-  
909 biology and Biotechnology, 104 (2020), pp. 51–65.
- 910 [42] H. DONG, L. NILSSON, AND C. G. KURLAND, *Gratuitous overexpression of genes in Escherichia coli*  
911 *leads to growth inhibition and ribosome destruction.*, Journal of Bacteriology, 177 (1995), pp. 1497–  
912 1504.
- 913 [43] N. TSIANTIS AND J. R. BANGA, *Using optimal control to understand complex metabolic pathways*, BMC  
914 Bioinformatics, 21 (2020), pp. 1–33.