

## Optimal control of bacterial growth for metabolite production: The role of timing and costs of control

Eugenio Cinquemani, Francis Mairet, Ivan Yegorov, Hidde de Jong, Jean-Luc Gouzé

► **To cite this version:**

Eugenio Cinquemani, Francis Mairet, Ivan Yegorov, Hidde de Jong, Jean-Luc Gouzé. Optimal control of bacterial growth for metabolite production: The role of timing and costs of control. ECC 2019 - 18th European Control Conference, Jun 2019, Naples, Italy. pp.2657-2662, 10.23919/ECC.2019.8796079. hal-02384034

**HAL Id: hal-02384034**

**<https://hal.inria.fr/hal-02384034>**

Submitted on 6 Dec 2019

**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.

# Optimal control of bacterial growth for metabolite production: The role of timing and costs of control\*

E. Cinquemani<sup>1,†</sup>, F. Mairet<sup>2,†</sup>, I. Yegorov<sup>3</sup>, H. de Jong<sup>1</sup>, J.-L. Gouzé<sup>4</sup>

**Abstract**—The growth of microorganisms is controlled by strategies for the dynamical allocation of available resources over different cellular functions. Synthetic biology approaches are considered nowadays to artificially modify these strategies and turn microbial populations into biotechnological factories for the production of metabolites of interest. In our recent work [1], [2] we have studied dynamics of microbial resource allocation and growth in terms of coarse-grained self-replicator models described by ordinary differential equations, and proposed artificial control strategies for the optimization of metabolite production based on the reengineering of resource allocation. In this paper, we elaborate on our earlier results and further investigate synthetic resource allocation control strategies. Using numerical simulation, we study the effect on growth and bioproduction of the (biological or technological) costs associated with discontinuous control strategies, and of the time allotted to optimal substrate utilization. Results provide novel insight into the most favorable synthetic control strategies.

## I. INTRODUCTION

The growth of microorganisms requires the allocation of available resources, extracted from nutrients in the environments, to different cellular functions. Microorganisms have evolved strategies to achieve this in such a way as to survive and proliferate in a dynamically changing environment. In recent years several coarse-grained models of resource allocation in microorganisms have appeared in the literature, relating observed resource allocation strategies to molecular mechanisms implementing these strategies and to criteria that microorganisms are considered to optimize [3], [1], [4], [5], [6], [7], [8]. In addition to natural resource allocation strategies, there has been interest in the modification of these strategies for biotechnological purposes. In the latter case, the challenge consists in the development of synthetic resource allocation strategies maximizing the production of some metabolite or protein of interest. The advances in synthetic biology over the past decade have provided powerful experimental tools enabling the implementation of strategies for resource reallocation in microbial cells [9].

In recent work [1], [4] we have developed coarse-grained self-replicator models of bacterial growth and applied control-theoretical methods to study natural resource

allocation in a dynamic context. This work has been extended to the modification of natural resource allocation strategies by means of a so-called bacterial growth switch [10], with the objective of optimizing the production of a metabolite of interest [2]. Using techniques from optimal control theory [11] and numerical simulation, we showed, among other things, that the strategy optimizing the production of the metabolite over a given time interval from a given amount of nutrients supplied in the environment consists of a simple two-step control scheme: A first phase with maximal growth (and low production), followed by a second phase with low growth (and high production). A limitation of this study, possibly affecting its practical applicability, is that we neglected the biophysical constraints on and cellular costs of rapid control actions on the molecular level. Moreover, we did not study in depth the relation between the length of the production time-interval and the optimal control strategy.

Here, as a first contribution, we explore whether the costs of regulation inherent in natural resource allocation play a role in the synthesis of the target product, that is, whether external control design should account for these costs to maximize production. Based on a penalized optimal control approach, we show that conversion of substrate into product is very much insensitive to these costs, which only affect internal cellular dynamics to a mild extent. As a second contribution, we study how the time horizon allotted for the exploitation of a finite amount of substrate impacts the design of the external control action that maximizes the amount of product at final time. In view of the findings above, internal control costs can be safely neglected in this part of the study. Not surprisingly, we find that larger time horizons allow for larger total product synthesis. More interestingly, the analysis shows that two different regimes exist, depending on whether or not the time horizon is sufficient for complete utilization of the available substrate, corresponding to different control strategies and different results in terms of productivity, that is, amount of product synthesized per unit time. A third regime is also found for time horizons that are so large that no resource is invested into population growth, since the sole initial population suffices to fully convert substrate into product.

The models of cellular resource allocation and product synthesis that we developed in [1], [2] are reviewed in Sec. II. The optimal control problems that we addressed in the same publications are reviewed in Sec. III. Costs of regulation are investigated in Sec. IV, whereas the role of time horizon is discussed in Sec. V. These two sections provide the original contribution of this work. Conclusions are drawn in Sec. VI.

\*This work was supported in part by the Inria IPL project CoSy and by the French national funding agency under project Maximic (ANR-17-CE40-0024). Corresponding author E. Cinquemani, eugenio.cinquemani@inria.fr

<sup>1</sup>Univ. Grenoble Alpes, Inria, 38000 Grenoble, France

<sup>2</sup>Ifremer PBA, Nantes, France

<sup>3</sup>Also known as I. Egorov; North Dakota State University, USA

<sup>4</sup>Université Côte d'Azur, Inria, INRA, CNRS, Sorbonne Université, Biocore Project-Team, Sophia-Antipolis, 06902, France

<sup>†</sup>These authors contributed equally to the work.

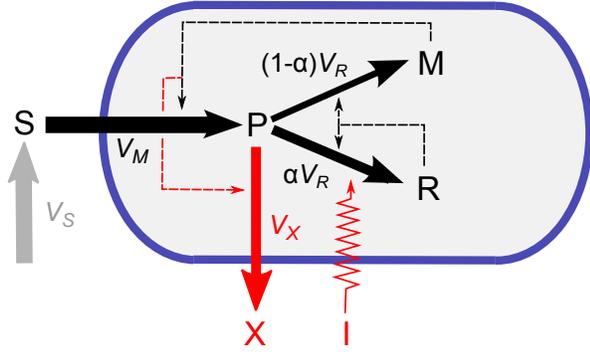


Fig. 1. Schematic representation of resource allocation with external control and synthesis of a product. The rod-shaped region represents the cellular population. Uptaken substrate  $S$  is transformed into precursors  $P$ , which are then allocated by internal cell control  $\alpha$  to metabolic machinery  $M$  or genetic machinery (ribosomes)  $R$  (biochemical fluxes: thick arrows; regulatory effects: thin arrows). In genetically modified cells,  $P$  can also be converted into product  $X$  under additional external control  $I$ . Black: Natural fluxes and regulation; Red: Additional fluxes and regulation of reengineered cells; Grey: Addition of substrate  $S$  into reaction volume during experiments (mostly neglected in this work). See text for more details.

## II. MODELLING

In this section we review the coarse-grained modelling of cellular growth and resource allocation developed in more generality in [1], [2]. The scenario considered is a population of cells growing on a single substrate in a fixed bioreactor volume (the volume occupied by cells being a negligible portion of the culture volume). The cellular population is seen as a growing biomass with identical behavior across cells. In accordance with this, reactions and their kinetics are described in terms of total population quantities and rates. Refer to Fig. 1. At a given experimental time  $t$ , let  $S(t)$  (in grams, [g]) be the quantity of substrate in the bioreaction volume. The substrate is imported by the cellular population at rate  $V_M(t)$  (grams per hour, [g h<sup>-1</sup>]) and converted into precursors (*e.g.* amino acids, the building blocks of proteins), whose total abundance over the population is denoted  $P(t)$  [g]. Cells can utilize  $P(t)$  to boost the genetic machinery, captured by the ribosomal abundance  $R(t)$  [g], or the metabolic machinery, captured by metabolic enzyme abundance  $M(t)$  [g]. The latter catalyzes substrate uptake and conversion into precursors, whereas the former catalyzes the synthesis of both ribosomes and metabolic enzymes, thus introducing dynamical feedback. In accordance with [1], cells can tradeoff conversion of  $P(t)$  into  $M(t)$  or  $R(t)$  via a control function  $\alpha(t) \in [0, 1]$ , leading to synthesis rates for  $M(t)$  and  $R(t)$  of the form  $(1 - \alpha(t))V_R(t)$  and  $\alpha(t)V_R(t)$  ([g h<sup>-1</sup>]), in the same order. The way  $V_M$  and  $V_R$  depend on  $S$ ,  $M$  and  $P$  is discussed later on. The growth of the population is expressed in terms of its volume  $\mathcal{V}(t)$  (in liters, [L]). It is assumed that  $\mathcal{V}(t) = \beta(M(t) + R(t))$ , with  $\beta$  [L g<sup>-1</sup>] a positive constant. The growth rate  $\mu(t) = \mathcal{V}^{-1}\dot{\mathcal{V}}(t)$  follows. In [1], an additional rate  $V_S(t)$  is considered to account for possible addition of  $S$  into the reaction volume.

In this paper we are interested in the scenario where cells are genetically modified so as to synthesize a metabolite of

interest under external control. This scenario is modelled in [2] as an extension of the above. In accordance with [2], let us denote  $X(t)$  [g] the abundance of the product of interest and  $I(t) \in [0, 1]$  the action of external control. Product  $X(t)$  is synthesized at a rate  $V_X(t)$  [g h<sup>-1</sup>] determined by metabolic activity  $M(t)$  and precursor availability  $P(t)$  as expressed later on. External control  $I(t)$  acts on the natural control  $\alpha(t)$ , resulting in the new resource allocation function  $u(t) = \alpha(t)I(t)$ . Throughout the rest of the paper, we additionally assume that the system is closed, namely no addition of  $S$  ( $V_S(t) = 0$ ) or removal of  $X$  occurs after the beginning of the experiment. Over a given time period  $t \in [0, T]$ , the dynamics of the extensive variables introduced so far can thus be written as the differential equation system

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

In order to specify the expression of the rate functions, it is convenient to replace  $P$ ,  $M$  and  $R$  by the intensive variables  $p(t) = P(t)/\mathcal{V}(t)$ ,  $m(t) = M(t)/\mathcal{V}(t)$  and  $r(t) = R(t)/\mathcal{V}(t)$  [g L<sup>-1</sup>], representing the respective concentrations in the (strictly positive) population volume. Thanks to mass conservation, one finds that  $m(t) = \beta^{-1} - r(t)$ . The above equations are then equally represented by the system

$$\begin{cases} \dot{S}(t) = -v_M(t)\mathcal{V}(t), \\ \dot{p}(t) = v_M(t) - v_R(t) - v_X(t) - \mu(t)p(t), \\ \dot{r}(t) = u(t)v_R(t) - \mu(t)r(t), \\ \dot{X}(t) = v_X(t)\mathcal{V}(t), \\ \dot{\mathcal{V}}(t) = \mu(t)\mathcal{V}(t), \end{cases} \quad (2)$$

for rates  $v_M(t) = V_M(t)/\mathcal{V}(t)$ ,  $v_R(t) = V_R(t)/\mathcal{V}(t)$  and  $v_X(t) = V_X(t)/\mathcal{V}(t)$  [g L<sup>-1</sup> h<sup>-1</sup>] that can now be expressed in terms of  $r(t)$  and  $p(t)$ . In [2], these rates are first defined and then renormalized along with a rescaling of concentration variables and time axis to eliminate some inessential parameters. Here, for notational simplicity and conciseness, we do not detail this step but rather assume that all concentrations and rates defined above refer to their normalized counterparts. For these quantities, consistent with [2], system (2) still holds, provided the definitions

$$\begin{aligned} v_M(t) &= k_2 \frac{s(t)m(t)}{K_2 + s(t)}, & v_R(t) &= \frac{p(t)r(t)}{K + p(t)}, \\ v_X(t) &= k_1 \frac{p(t)m(t)}{K_1 + p(t)}, & \mu(t) &= v_R(t), \end{aligned} \quad (3)$$

where  $k_1$ ,  $k_2$ ,  $K$ ,  $K_1$  and  $K_2$  are (renormalized) nonnegative constants, and  $m(t) = 1 - r(t)$  (renormalization eliminates parameter  $\beta$  introduced earlier on). See [1], [2] for an extensive discussion of this choice of reaction rate expressions. We will typically assume that the initial conditions

$$S(0) = S_0, p(0) = p_0, r(0) = r_0, X(0) = X_0, \mathcal{V}(0) = \mathcal{V}_0 \quad (4)$$

obey  $S_0 > 0$ ,  $\mathcal{V}_0 > 0$ ,  $X_0 = 0$ ,  $p_0 > 0$ , as well as  $0 < r_0 \leq 1$  (the latter bounds follow from renormalization and the need of a nonempty ribosome pool to start their own synthesis and growth).

### III. OPTIMAL CONTROL PROBLEMS

The model of the previous section allows one to address in the first place questions related with the natural strategies that cells implement to optimize growth. The question is first addressed in [1], for the case of wild-type cells (that is, no  $X$  and  $I$ , and hence  $v_X = 0$ ) exposed to sudden availability of  $S$  at an invariant concentration in the reaction volume ( $V_S = V_M$ ). For the case of this paper ( $V_S = 0$  and synthesis of  $X$ ) the problem is addressed in [2]. In both cases, the question is formulated as an optimal control problem for the maximization of the population volume  $\mathcal{V}$  at final time  $T$ ,

$$\max_{u \in \mathcal{U}} \log \frac{\mathcal{V}(T)}{\mathcal{V}_0} = \int_0^T \mu(t) dt \quad \text{subject to (2)–(4),} \quad (5)$$

with  $\mathcal{U}$  the class of (measurable) functions taking values in  $[0, 1]$ . Here external control is absent (captured by  $I$  fixed to 1) and  $u(t)$  has the interpretation of internal control law ( $u(t) = \alpha(t)$ ). It can be shown [1], [2] that the optimal control profile  $u^*(t)$  that cells should implement to maximize biomass production in finite time is a bang-bang-singular profile. That is, upon administration of  $S$ , cells exit their quiescence state by abruptly alternating maximal synthesis of  $M$  ( $\alpha = 0$ ) with maximal synthesis of  $R$  ( $\alpha = 1$ ), until a singular intermediate steady-state is achieved ( $\alpha$  in-between). A biological discussion based on literature is provided in [1], showing that a similar control law may correspond, at least qualitatively, to a known biochemical circuit existing in cells.

Additionally, in [2], maximization of product synthesis  $X$  is considered, based on a finite substrate amount  $S$  and using external control  $I$ . The question is again formulated as an optimal control problem,

$$\max_{u \in \mathcal{U}} X(T) = \int_0^T v_X(t) \mathcal{V}(t) dt \quad \text{subject to (2)–(4).} \quad (6)$$

Once again, the solution  $u^*$  found in [2] is bang-bang-singular (see illustration later on in Fig. 2). Here, however,  $u(t)$  corresponds to the product  $\alpha(t)I(t)$ . In the light of our model, this is the control action driving the internal cell dynamics after compensation by the external induction  $I(t)$ . If  $u^*$  is a solution to the problem and  $\alpha(t)$  is the natural strategy that the cell would implement, the optimal external compensation is  $I(t) = u^*(t)/\alpha(t)$ , which explicitly depends on the knowledge of  $\alpha(t)$ .

In all of these problems, the abruptness of the optimal dynamics of  $u^*$  is questionable. From a biophysical standpoint, this implies (repeated) instantaneous reorganizations of the cellular physiology that are not realistic. From a biotechnological viewpoint, however, the question is whether accounting precisely for these internal costs and the details of the natural policy  $\alpha$  has any relevant effect on the design of external induction  $I(t)$  and the achievable  $X(T)$ . This question is the object of Sec. IV. Moreover, it is an open question

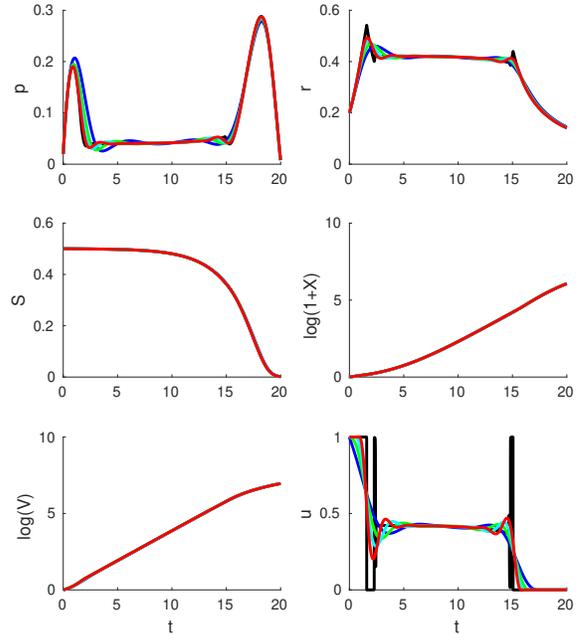


Fig. 2. Solutions of the optimal control problem (7), with control cost  $\lambda = 10^{-1}$  (blue),  $\lambda = 10^{-2}$  (green),  $\lambda = 10^{-3}$  (cyan) and  $\lambda = 10^{-4}$  (red), and of the optimal control problem (6), without control cost (black lines).

whether the choice of  $T$  has itself relevant effects on the attainable  $X(T)$ , depending on how substrate consumption is distributed over  $[0, T]$ , and on the corresponding optimal control strategies. Also based on the results of Sec. IV, this is the object of Sec. V.

### IV. COST OF CONTROL

In this section we investigate the sensitivity of optimal control policies to possible costs associated with abrupt changes of the control action. We notably wish to assess whether the achievable product synthesis  $X(T)$  is sensitive to these costs. To formalize the problem we introduce a modified version of problem (6). In analogy with regularized estimation approaches [12], we express the cost of rough control actions by the penalization term

$$\lambda \int_0^T \ddot{u}(t)^2 dt, \quad \lambda \geq 0,$$

the faster the changes of  $u$ , the larger the cost. Subtracting this cost from the objective function  $X(T)$  implies favoring smoother control solutions (see also [13] for a related approach to regularized control). Yet the very definition of this cost requires the second derivative of  $u$  to be well-defined, which is not guaranteed by the problem statement (6) with  $u \in \mathcal{U}$ . To tackle this, we modified the problem as follows. Let us define  $v = \dot{u}$  and  $w = \dot{v} = \ddot{u}$ . Let  $\mathcal{W}$  be the space of functions over  $[0, T]$  that are measurable and uniformly

bounded. We now modify (6) into the optimization problem

$$\begin{aligned}
& \max_{w \in \mathcal{W}, u_0, v_0} X(T) - \lambda \int_0^T w(t)^2 dt = \\
& \int_0^T (v_X(t) \mathcal{V}(t) - \lambda w(t)^2) dt \\
& \text{subject to (2)–(4)} \\
& \text{and } \begin{cases} \dot{u} = v, & u(0) = u_0, \\ \dot{v} = w, & v(0) = v_0, \end{cases} \\
& \text{and } 0 \leq u(t) \leq 1, \quad t \in [0, T].
\end{aligned} \tag{7}$$

This way, double-differentiability of  $u$  is guaranteed by construction. If  $(w^*, u_0^*, v_0^*)$  is a solution to this problem, the corresponding solution in  $u$  is then fixed by  $\ddot{u}^*(t) = w^*(t)$ ,  $t \in [0, T]$ , with  $u^*(0) = u_0^*$ ,  $\dot{u}^*(0) = v_0^*$ . We seek solutions for  $\lambda > 0$ , the larger  $\lambda$ , the smoother the resulting control  $u$  at the price of a generally smaller production  $X(T)$ .

To answer the question that opens this section, we computed solutions of the optimal control problem (7) for different weights  $\lambda$  of the control cost and compared them with the solution of the original problem (6), that is, in absence of control costs. Calculations were performed for parameter values borrowed from [2], namely  $(k_1, k_2, K, K_1, K_2) = (0.139, 1, 0.003, 0.003, 0.1)$ ,  $T = 20$  and initial conditions  $(S_0, p_0, r_0, X_0, \mathcal{V}_0) = (0.5, 0.02, 0.2, 0, 1)$ , in the respective units. Numerical solutions were obtained by the software package Bocop [14].

Results are shown in Fig. 2. The different levels of smoothness of the control action associated with the different values of  $\lambda$  tested are apparent. For  $\lambda$  approaching 0, as expected, the solution approaches the cost-unaware bang-bang-singular strategy. Very interestingly, the different control strategies  $u$  obtained affect the internal resource allocation profiles  $p$  and  $r$  to a noticeable extent, but they hardly affect the production profile of  $X$  and the consumption profile of  $S$ . This can be explained in part by the (realistic) initial conditions of the system. Indeed, the greatest impact of control costs appears to be on the initial transient of the optimal control strategy, whereby cells need to adjust their physiology to the sudden availability of substrate. Differences in this transient modify internal resource allocation from an initially small pool of precursors  $p$  and over a small population  $\mathcal{V}$ , such that their net effect on growth and productivity is negligible.

To sum up the section, we showed by a regularized optimal control approach that the effects of control costs on the attainable synthesis of a product are negligible in realistic conditions, despite detectable changes in internal dynamics.

## V. ROLE OF TIME HORIZON

In this section we address the question of how optimal control for the maximal production of a compound of interest depends on the allotted production time  $T$ . Intuitively speaking, the question is how to best balance substrate utilization for population growth versus its utilization for the synthesis of the product by the existing cellular population. Time clearly enters this question since a longer horizon allows

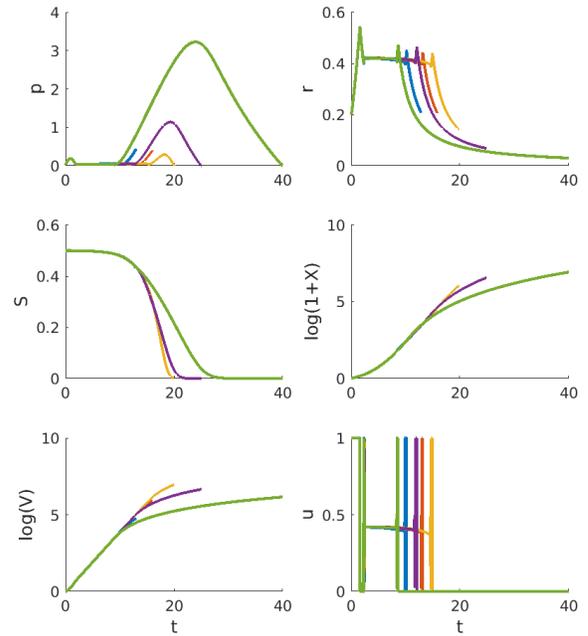


Fig. 3. Time dynamics of  $p$ ,  $r$ ,  $S$ ,  $X$ ,  $V$  and control action  $u$  associated with the solution of problem (6) for the different time horizons  $T = 13$  (blue),  $T = 16$  (red),  $T = 20$  (yellow),  $T = 25$  (purple) and  $T = 40$  (green).

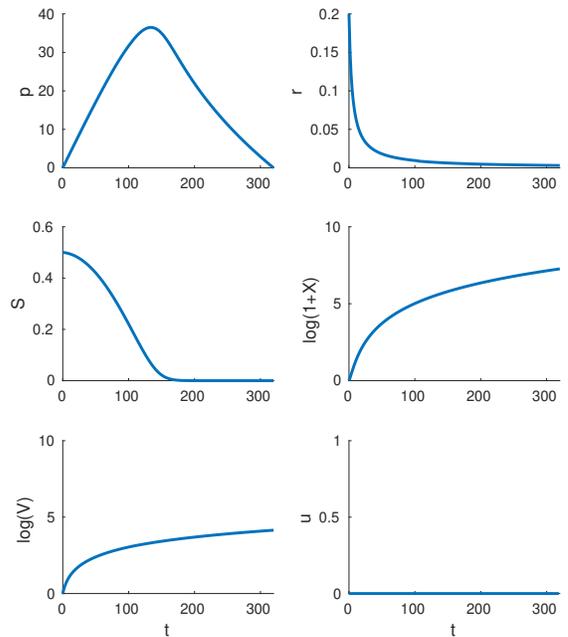


Fig. 4. Same as in Fig. 3, but with  $T = 320$ .

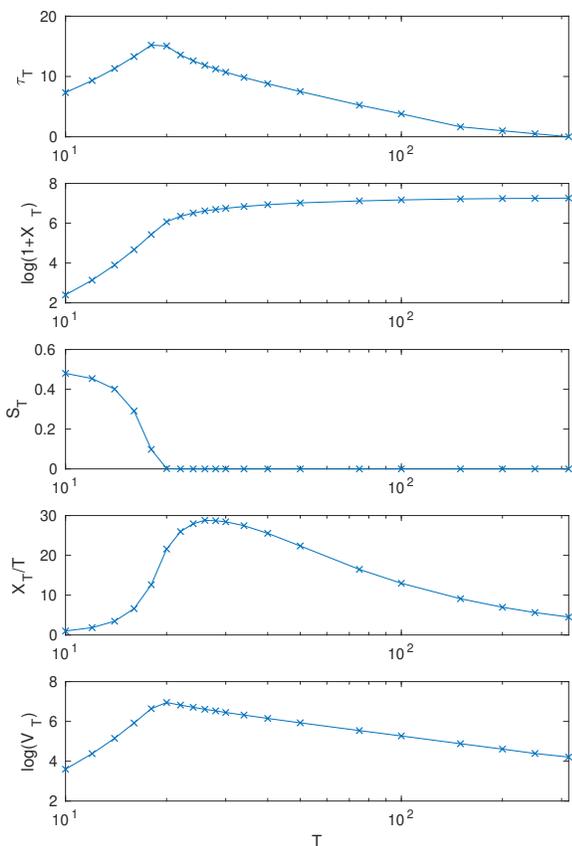


Fig. 5. Dependency of optimal switch time  $\tau_T$ , total production  $X_T$ , substrate residual  $S_T$ , average production rate  $X_T/T$ , and final population volume  $\mathcal{V}_T$  on the control time horizon  $T$  (see text for the mathematical definition of these quantities).

one to build up the cell factory (cell population) at a slower pace, thus consuming less resources. If one may expect that a larger  $T$  enables a larger production  $X(T)$  (just because control is given more flexibility), it is *a priori* unclear what control can achieve this, and what are its implications on the cellular population dynamics.

In the light of the results of the previous section, as long as  $X$  is the variable of interest, cellular control costs need not be accounted for in much detail. To address the question of this section, we therefore rely on the solution of the maximum production control problem in the original form (6). For the same system parameters and initial conditions of the previous section, we numerically solved problem (6) for different time horizons  $T$ , again using Bocop [14].

Time-profiles of the different optimal control actions and resulting system dynamics are reported in Fig. 3 and 4. Dependency on the time horizon  $T$  of a few useful indicators of the optimal control action is shown in Fig. 5. In the latter figure and below, for any given value of  $T$ , quantities  $X_T$ ,  $S_T$  and  $\mathcal{V}_T$  denote the values of  $X(T)$ ,  $S(T)$  and  $\mathcal{V}(T)$ , in the same order, corresponding to the solution of (6). (An

additional quantity reported in the figure,  $\tau_T$ , is introduced further below). As a first confirmation, not surprisingly,  $X_T$  increases with  $T$  (see in Fig. 3 the different profiles of  $X$  at the respective final times, and, in Fig. 5, the plot of  $\log(1 + X_T)$  as a function of  $T$ ). For the smallest values of  $T$ , time is insufficient to fully utilize the available substrate. For the given system parameters, the minimum horizon  $\bar{T}$  to deplete the substrate is about 20. The control actions corresponding to these cases show a similar initial transient to get into a growth phase. The larger  $T$ , the longer this phase (compare the profiles of  $\mathcal{V}$  and  $u$  in Fig. 3), after which all resources are redirected for the residual time into utilization of  $S$  for the production of  $X$  (the residual growth is explained by the existing pool of ribosomes  $R$  built up during the growth phase). Note that ignoring internal control costs, that is, obtaining bang-bang solutions for  $u$ , has the advantage that this phase change can be defined precisely as the time  $\tau_T$  when  $u$  drops to 0. In Fig. 5, consistent with the above discussion,  $\tau_T$  is shown to be an increasing function of  $T$  up to  $T = \bar{T}$ .

For  $T > \bar{T}$ , a change of regime is observed. The time of arrest of ribosome synthesis,  $\tau_T$ , starts receding. Indeed, an increased time horizon  $T$  allows for a smaller population (as obtained with a smaller  $\tau_T$ ) to fully transform the available substrate  $S$  into product  $X$ . It is therefore convenient to spare  $S$  from the growth of a larger population in favor of a longer phase ( $t \geq \tau_T$ ) where fewer cells convert  $S$  into  $X$  until depletion. This is clearly illustrated in Fig. 5, where  $\tau_T$  is shown to be a decreasing function of  $T$  for  $T > \bar{T}$  (in Fig. 3, one sees that  $\tau_T$  is such that the precursors are exhausted precisely at the final time). Correspondingly (Fig. 5), the total production  $X_T$  increases with  $T$  although at a lower rate compared to its increase for  $T$  up to  $\bar{T}$  (for  $T \geq \bar{T}$ , the amount of substrate utilized is invariant and equal to  $S_0$ ). Yet the productivity, that is, the average rate of product synthesis  $X_T/T$ , decreases over  $T$  (see again Fig. 5), since production is diluted out in time. Should production be made more robust to possible interruption of the process, or compliant with extraction of  $X$  in the course of the process, a modified optimal control objective should be considered taking productivity explicitly into account.

In particular, for  $T \gg \bar{T}$ , one observes in Fig. 5 that  $\tau_T$  converges to zero. In essence, for  $T$  large enough (in our simulations, for  $T \geq 320$ ), the optimal control strategy collapses into one where the entire conversion of  $S$  into  $X$  is operated by a minimal cell population, as obtained from the initial pool of cells  $X_0$  by the growth resulting from the initial pool of ribosomes  $r_0$  (see Fig. 4). In this regime,  $u$  is identically 0, until either substrate and precursors are exhausted (and it becomes irrelevant thereafter). This is easy to explain: With arbitrarily large  $T$ , there is no need to waste resources  $S$  into the growth of a larger factory of cells, since a small population will anyhow complete the job although at a slower pace.

To conclude this section, we have shown in simulation that two different regimes exist depending on whether the allotted production time allows for a complete utilization

of the substrate. These two regimes yield different optimal control strategies. A third regime exists corresponding to full conversion of substrate into product by the smallest possible population (no resource allocated to population growth). It is worth recalling that the strategies discussed above represent the optimal regulatory signal  $u^*$  internal to the cell. Based on this and the (approximate) knowledge of the natural control signal  $\alpha(t)$  (as discussed in Sec. IV), the optimal external induction  $I(t)$  can be determined as  $I(t) = u^*(t)/\alpha(t)$ . Analysis of compatibility of the resulting induction strategies with possible technological and physical constraints (boundaries on induction strength, etc.), along with easily adapted problem reformulations, are out of the scope of this paper.

## VI. CONCLUSIONS

In this paper we have addressed open questions pertaining to the biotechnological synthesis of compounds of interest by a growing population of cells via an optimal control approach. Starting from earlier work by the same authors and collaborators, we have first reviewed the published modelling and optimal control results. The use of a coarse-grained, whole-cell model of resource allocation and microbial growth distinguishes our work from previous application of (optimal) control in biotechnology [15], [16], [17]. We have then isolated and addressed two open questions, the possible impact of physiological costs of internal cellular resource reallocation onto the design of optimal control strategies, and the role of the duration of the production process in control design. We have first shown by a regularized control approach that the cellular costs incurred in resource reallocation have a negligible net impact on the conversion of substrate into product. Then, we have shown that the biotechnological control rationale should be different depending on whether or not the time for production allows for the complete depletion of the substrate.

The results obtained are of direct interest for the biotechnological applications they refer to. The little sensitivity of production to details of internal cellular dynamics (costs) suggests that optimal control of the production process does not require precise or finer-grained models. At a genetic engineering level, it also suggests that a certain degree of inaccuracy in the design of the engineered strains can be tolerated. The different regimes observed for optimal control agree to some extent with the well-established practice in dynamic bioprocess operation of following up a growth phase by a production phase [18]. However, such empirical control strategies could be further optimized using the kind of optimal control techniques explored here, especially in the context of batch processes. Yet, the analysis also implicitly suggests that the situation for fed-batch and continuous cultivation processes, where substrate is replenished in the course of production, may be different, and deserves further study.

From a biological viewpoint, the results show that the observed cellular growth and metabolite secretion dynamics are rather robust to internal regulatory mechanisms and

associated costs. In a different interpretation, this means that reconstructing internal mechanisms from external observations is hard. Targeted experimental work is thus needed to pinpoint the internal cellular dynamics and the costs of resource reallocation, and further mathematical developments other than those of Sec. IV may be needed to quantitatively characterize them. Finally, from a mathematical viewpoint, an analytical study of what has been shown in simulation here constitutes an intriguing direction of research. Theoretical results on the optimal control problems addressed in this work would further boost the application of control-theoretic approaches to cellular biology [9].

## REFERENCES

- [1] N. Giordano, F. Mairet, J. Gouzé, J. Geiselmann, and H. de Jong, "Dynamical allocation of cellular resources as an optimal control problem: Novel insights into microbial growth strategies," *PLoS Comput. Biol.*, vol. 12, no. 3, p. e1004802, 2016.
- [2] I. Yegorov, F. Mairet, H. de Jong, and J.-L. Gouzé, "Optimal control of bacterial growth for the maximization of metabolite production," *J Math Biol.* 2018. [Online]. Available: <https://doi.org/10.1007/s00285-018-1299-6>
- [3] D. Molenaar, R. van Berlo, D. de Ridder, and B. Teusink, "Shifts in growth strategies reflect tradeoffs in cellular economics," *Mol. Syst. Biol.*, vol. 5, p. 323, 2009.
- [4] I. Yegorov, F. Mairet, and J. Gouzé, "Optimal feedback strategies for bacterial growth with degradation, recycling and effect of temperature," *Optim. Control Appl. Methods*, vol. 39, no. 2, pp. 1084–1109, 2018.
- [5] H. van den Berg, Y. Kiselev, S. Kooijman, and M. Orlov, "Optimal allocation between nutrient uptake and growth in a microbial trichome," *J. Math. Biol.*, vol. 37, no. 1, pp. 28–48, 1998.
- [6] M. Scott, S. Klumpp, E. Mateescu, and T. Hwa, "Emergence of robust growth laws from optimal regulation of ribosome synthesis," *Mol. Syst. Biol.*, vol. 10, p. 747, 2014.
- [7] M. Pavlov and M. Ehrenberg, "Optimal control of gene expression for fast proteome adaptation to environmental change," *Proc. Natl. Acad. Sci. USA*, vol. 110, no. 51, pp. 20527–32, 2013.
- [8] E. Bosdriesz, D. Molenaar, B. Teusink, and F. Bruggeman, "How fast-growing bacteria robustly tune their ribosome concentration to approximate growth-rate maximization," *FEBS J.*, vol. 282, no. 10, pp. 2029–44, 2015.
- [9] D. D. Vecchio, A. Dy, and Y. Qian, "Control theory meets synthetic biology," *J. R. Soc. Interface*, vol. 13, no. 120, p. 20160380, 2016.
- [10] J. Izard, C. Gomez Balderas, D. Ropers, S. Lacour, X. Song, Y. Yang, A. Lindner, J. Geiselmann, and H. de Jong, "A synthetic growth switch based on controlled expression of RNA polymerase," *Mol. Syst. Biol.*, vol. 11, no. 11, p. 840, 2015.
- [11] H. Schattler and U. Ledzewicz, *Geometric Optimal Control: Theory, Methods and Examples*, ser. Interdisciplinary Applied Mathematics 38. New York: Springer-Verlag, 2012.
- [12] G. Wahba, *Spline Models for Observational Data*. Philadelphia, USA: SIAM, 1990.
- [13] M. Caponigro, R. Ghezzi, B. Piccoli, and E. Trélat, "Regularization of chattering phenomena via bounded variation controls," *IEEE Trans Autom Control*, vol. 63, pp. 2046–2060, 2018.
- [14] F. Bonnans, P. Martinon, D. Giorgi, V. Grélard, S. Maindrault, O. Tissot, and J. Liu, *Bocop 2.0.5 – User Guide*, 2017. [Online]. Available: <http://bocop.org>
- [15] G. Bastin and D. Dochain, *On-line Estimation and Adaptive Control of Bioreactors*. Amsterdam: Elsevier, 1990.
- [16] J. Banga, E. Balsa-Canto, C. Moles, and A. Alonso, "Dynamic optimization of bioprocesses: Efficient and robust numerical strategies," *J. Biotechnol.*, vol. 117, no. 4, pp. 407–19, 2005.
- [17] C. Venkateswarlu, "Advances in monitoring and state estimation of bioreactors," *J. Sci. Indus. Res.*, vol. 63, pp. 491–8, 2005.
- [18] N. Venayak, N. Anesiadis, W. Cluett, and R. Mahadevan, "Engineering metabolism through dynamic control," *Curr. Opin. Biotechnol.*, vol. 34, pp. 142–52, 2015.