Dynamical resource allocation models for bioreactor optimization

Guillaume Jeanne, Anne Goelzer, Sihem Tebbani, Didier Dumur, Vincent Fromion

Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 51-19 (2018) 20–23

Dynamical resource allocation models for bioreactor optimization

Guillaume Jeanne *,** **Anne Goelzer** * **Sihem Tebbani** **
Didier Dumur ** **Vincent Fromion** *

* *MaIAGE, INRA, UR1404, Université Paris-Saclay, 78350 Jouy-en-Josas, France (e-mail: firstname.name@inra.fr).*
** *L2S, CentraleSupélec - CNRS - Univ. Paris-Sud, Université Paris-Saclay, Control Department, Plateau du Moulon, 91190 Gif-sur-Yvette, France (e-mail: firstname.name@centralesupelec.fr)*

Abstract: Resource allocation based models were shown to capture the main bacterial cell design principles in steady-state, thus managing the resource allocation between cell processes, and in fine growth rate. This paper introduces the basis of extension of resource allocation models to dynamical conditions. The framework is applied to a bioreactor operating in batch mode for optimizing the production of added-value compounds by bacteria. The optimal predicted strategy is analyzed and discussed in light of the strategies obtained with other types of models. This framework should offer new opportunities for biotechnologies, especially for the simultaneous optimization of strain design and bioprocess control.

Keywords: Bioprocess optimization, Resource allocation model, Bacteria, Optimal control, Batch bioreactor

1. INTRODUCTION

Living cells are composed of interacting cellular processes, that perform specific tasks. Cellular processes consume resources for achieving their function, which creates operational constraints between these processes. The field of Systems Biology studied and exploited operational constraints both for theoretical and predictive purposes, within the constraint-based modeling framework. The first operational constraint – the mass balance of the metabolic network in steady-state – was formalized into an optimization problem known as Flux Balance Analysis (FBA) (Varma and Palsson, 1994). The FBA method predicts the flux distribution in the metabolic network and showed an outstanding predictive capability despite the simplicity of the cell description. This paved the way for a great number of works, to refine the cell description, to apply the framework for metabolic engineering or for the strain design (Lewis et al., 2012; Chowdhury et al., 2015), and finally to extend the theoretical framework, especially to dynamical conditions (dFBA) (Mahadevan et al., 2002).
In Gadkar et al. (2005) and Jabarivelisdeh and Waldherr (2016), dFBA formulations were used for strain and bioprocess design. Mathematically, dFBA correspond to bilevel optimization schemes that predict the metabolic time configuration maximizing growth or production of a product of interest.

Since 2009, a gap has been achieved in the cell description with the explicit integration of non-metabolic cellular processes within constraint-based models (reviewed in Goelzer and Fromion (2017)). These models formalize for a given growth rate, the resource sharing between all cellular processes at genome scale into a convex optimization problem and include explicitly the detailed building costs of proteins and molecular machines (e.g. ribosomes). This paper focuses on the RBA method (Goelzer et al., 2011), since all other methods can be reformulated under the RBA framework. In practice, the RBA method predicts in steady-state the metabolic fluxes and the whole-cell protein distribution maximizing growth rate and satisfying: (i) stochiometric constraints and mass conservation; (ii) viability of the cell, meaning that the cell has to produce all the cellular components in sufficient quantity compared to growth demand; (iii) limited capacity of the cellular processes, e.g. the proteins and the molecular machines have a limited efficiency to fulfill their function; (iv) limited density of cellular compartments, limiting the total amount of proteins allocated to one compartment.
The objective of this paper is to determine the outlines of extension of the RBA framework to dynamical conditions (dRBA), and like dFBA, apply the framework for bioreactor optimization.

2. PROBLEM FORMULATION

2.1 Problem description

In this work, a single bacterial species is supposed to be immersed in a bioreactor, operating in batch mode, i.e., with a given and fixed volume of medium, V, assumed perfectly stirred and homogeneous. In order to keep the developments simple, we consider that only one substrate G is available in the medium.
We consider in the sequel a simplified model of the cell integrating the key ingredients of the RBA approach.

10.1016/j.ifacol.2018.09.020

The cell is then described through three main cellular processes: (*i*) the first one, denoted Σ_T, aggregates all the cellular processes related to the catabolism, i.e. processes (mostly enzymes) involved in the furniture of energy and elementary bricks necessary for bacteria survival and growth; (*ii*) the second one, denoted Σ_R, aggregates all the cellular subprocesses necessary for protein synthesis, i.e. the so-called translation apparatus, including ribosomes and a large number of accessory proteins, as e.g. the elongation factors and chaperones; (*iii*), the third one, denoted Σ_B, includes all cellular processes involved in the production of cellular macro-components that are not proteins as e.g. DNA, membrane, cell-wall, etc. In the sequel, for readability purpose and without loss of generality, we assumed that the molecular machines involved in all cellular processes are only composed of proteins.
In this simplified model, the substrate G is then imported and transformed by Σ_T, into an internal metabolite S. The metabolite S can be used indifferently by Σ_R and Σ_B in order to build respectively proteins and other cellular macro-components B necessary to the cell growth. We then associate for each unit of Σ_T, Σ_R and Σ_B a resource cost, corresponding to the number of S required for building each unit, and an efficiency coefficient per unit. Finally, since the purpose of the model is to optimize the production of a product of interest P, we add an additional cellular process, denoted Σ_P, to produce P from S and secrete it in the medium. We finally associate to Σ_P, its resource cost and its efficiency coefficient.

2.2 Cell model description

Following this preliminary description, a part of the metabolite S is converted into B and P by the fluxes ν_B and ν_P with stochiometric coefficients r_B and r_P respectively. Another part of S is consumed to produce the proteins contained in the four cellular processes. We introduce four fluxes of S, denoted by ν_{e_i} with $i \in \{T, B, P, R\}$, corresponding to the flux of S used to build the proteins of the i-th cellular process.
Dynamic evolution of S. The differential equation associated to the internal metabolite S is given by:

$$[\dot{S}](t) = \nu_T(t) - \nu_B(t) - \nu_P(t) - \nu_{e_T}(t) \cdots \\ -\nu_{e_B}(t) - \nu_{e_P}(t) - \nu_{e_R}(t) - \mu(t)[S](t) \tag{E.1}$$

In the previous system, $\mu(t)[S](t)$ takes into account the effects due to the cell volume increase where $\mu(t)$ is the bacterial growth rate (defined in § 2.3). ν_T is the nonnegative flux of G imported and transformed by Σ_T into S. The uptake flux is constrained by the maximum capability of Σ_T, leading to this first constraint:

$$\nu_T \leq k_T[e_T] = \frac{v_{m,T}[G]_{ext}}{[G]_{ext} + K_T + K_S[S]}[e_T] \tag{C.1}$$

where $[e_T]$ is the concentration of proteins of Σ_T. The efficiency coefficient k_T of Σ_T follows a first order kinetic reaction inhibited by the product 'S' (leading to assume that the uptake is inhibited by high internal concentration of S).
Synthesis of macro-components B by Σ_B. The differential equation associated to the production of the cellular macro-components B is given by:

$$[\dot{B}](t) = r_B \nu_B(t) - \mu(t)[B](t) \tag{E.2}$$

where ν_B is the nonnegative flux of production of cellular macro-components. We assume that $\nu_B(t) = \frac{[B]_0}{r_B}\mu(t)$ in order to ensure that cellular macro-components remain to a concentration close to $[B]_0$. The production flux of B is constrained by the maximum capability of Σ_B, leading to introduce this second constraint:

$$\nu_B \leq k_B[e_B] = \frac{v_{m,B}[S]}{[S] + K_B}[e_B] \tag{C.2}$$

where $[e_B]$ is the concentration of proteins of Σ_B and k_B, the efficiency coefficient of Σ_B, is a Michaelis-Menten like relation.
Synthesis of P by Σ_P. The differential equation associated to the production of P is given by:

$$[\dot{P}]_{ext}(t) = r_P \nu_P(t)[M_{tot}]_{ext}(t) \tag{E.3}$$

where $[M_{tot}]_{ext}$ is the concentration of cells in the bioreactor and ν_P is the nonnegative production flux. Like the other cellular processes, Σ_P is constrained by:

$$\nu_P \leq k_P[e_P] = \frac{v_{m,P}[S]}{[S] + K_P}[e_P] \tag{C.3}$$

where $[e_P]$ is the concentration of proteins of Σ_P and k_P is the efficiency coefficient of Σ_P, corresponding to a Michaelis-Menten like relation.
Synthesis of proteins by Σ_R. Finally, the proteins e_i for $i \in \{T, B, P, R\}$ involved in the cellular processes are produced by Σ_R, which leads to the following differential equations:

$$[\dot{e_i}](t) = r_{e_i}\nu_{e_i}(t) - \mu(t)[e_i](t) \tag{E.4}$$

where r_{e_i} is the resource cost necessary to build the protein e_i and $\mu[e_i]$ takes into account the effects due to the cell volume increase. The production flux of proteins is constrained by the maximum capability of Σ_R, leading to the constraint:

$$\sum_{P,B,R,T} \nu_{e_i} \leq k_R[e_R] = \frac{v_{m,R}[S]}{[S] + K_R}[e_R] \tag{C.4}$$

with $[e_R]$ the concentration of proteins of Σ_R and k_R the efficiency coefficient of Σ_R depending on S, as in Marr (1991). We further assume that dilution is the only phenomenon responsible for decreasing enzyme concentrations, i.e. the fluxes ν_{e_i} for $i \in \{T, B, P, R\}$ take only nonnegative values.
Moreover, all concentrations have also to take nonnegative values.

2.3 Population dynamics and growth rate evolution

Kubitschek et al. (1984) revealed that cell density is constant for any growth condition and along the cell cycle of bacteria as *E. coli* (or *B. subtilis*). We then follow the approach of Marr (1991), by relating the cell density, D_0, to the whole protein components:

$$\sum_{P,B,R,T} \frac{[e_i](t)}{r_{e_i}} = D_0.$$

In order to keep the density constant, the cell has to increase its volume when new proteins are produced. This leads to constrain the growth rate of bacteria. Such a mechanism can be described by this first order differential equation:

$$\dot{\mu}(t) = \frac{1}{\tau_\mu}\left(\frac{1}{D_0}\sum_{P,B,R,T} \nu_{e_i}(t) - \mu(t)\right) \tag{E.5}$$

where τ_μ is the constant related to the cell wall synthesis dynamics. Consequently, in each differential equation of § 2.2, growth rate is responsible of a dilution term, $-\mu[*]$ for all concentration dynamics.
Following this definition of the cell volume increase and since we have only one species of bacteria, and that their density remains constant, we can describe the evolution of the cell population through this differential equation:

$$[\dot{M_{tot}}]_{ext}(t) = \mu(t)[M_{tot}]_{ext}(t). \tag{E.6}$$

Finally, the dynamics of the external substrate $[G]_{ext}$ in the medium is given by the following differential equation:

$$[\dot{G}]_{ext}(t) = -r_G \nu_T(t)[M_{tot}]_{ext}(t) \tag{E.7}$$

where r_G is the stochiometric coefficient, corresponding to the number of G required to build S by the process Σ_T.

2.4 dRBA optimization problem

The maximization of production can be formulated in several ways. The final quantity or concentration of product of interest is the most natural criterion, but the time necessary to obtain the maximal quantity is also an essential issue. As suggested in Gadkar et al. (2005), a solution could be to find the Pareto front of maximum final quantity of product versus final time t_f. Here we follow the approach proposed in Jabarivelisdeh and Waldherr (2016) by defining a criterion given by the ratio between the concentration of the product at the final time and the culture duration. The dynamics and the constraints of the optimization problem have been defined in § 2.2 and § 2.3. It remains to introduce the initial concentrations in G, P and M_{tot} (respectively denoted $[G]^0_{ext}$, $[P]^0_{ext}$ and $[M_{tot}]^0_{ext}$) and to define the final time t_f of the culture duration within the bioreactor as the time where G is less than or equal to 1% of its initial concentration:

$$\max_{\nu_{e_P},\nu_{e_B},\nu_{e_R},\nu_{e_T}} \frac{[P]_{ext}(t_f)}{t_f}$$
$$w.r.t. \begin{cases} (\text{E.1}),\dots,(\text{E.7}) \\ (\text{C.1}),\dots,(\text{C.4}) \\ \nu_i \geq 0 \text{ for all } i \in \{P,B,T\} \\ \nu_{e_i}, [e_i] \geq 0 \text{ for all } i \in \{P,B,R,T\} \\ [S],[B],[G]_{ext},[P]_{ext} \geq 0 \\ \text{Initial conditions, } [G]^0_{ext},[P]^0_{ext},[M_{tot}]^0_{ext} \\ [G]_{ext}(t_f) \leq 1\%[G]^0_{ext} \end{cases} \tag{1}$$

The problem formulated in (1), called hereafter dRBA, is a Mayer problem, see e.g. Lee and Markus (1967) or Bryson and Ho (1975). We invoke the necessary conditions given by the Pontryagin's maximum principle to identify the trajectory candidates for optimality.
From practical point of view, for speed and simplicity, we use Bocop (Bonnans et al. (2017)) in order to compute efficiently the open-loop control and the associated optimal trajectory (the computation time for solving our problem is around ten seconds with a laptop).

3. RESULTS

The dRBA optimization problem (1) is solved using model parameters deduced from Jeanne et al. (2016), and initial conditions (10,0,0.045) for $[G]^0_{ext}$, $[P]^0_{ext}$ and $[M_{tot}]^0_{ext}$ respectively. The time evolutions of state variables are given in Fig. 1. The optimal trajectory presents three main phases:

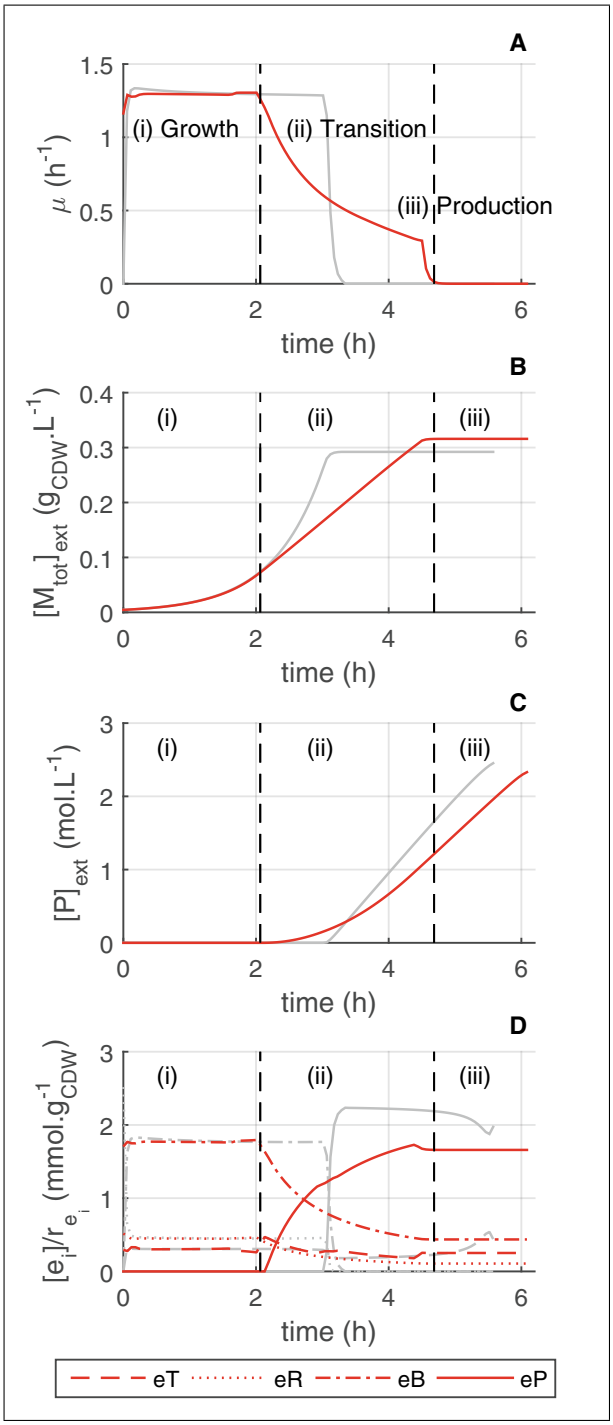

Fig. 1. Optimal time evolution of: A- growth rate, B- concentration of micro-organisms, C- concentration of product of interest, D- intracellular concentration of enzymes. In red, the optimal solution predicted by dRBA. In grey, the dRBA^u case (see text) mimicking dFBA formulation. g_{CDW} is gram per cell dry weight.

(i) Proteins are first allocated towards biomass synthesis only: cells grow at constant growth rate without production of P. Enzyme concentrations are constant

and this phase coincides with a so-called balanced exponential regime;

(ii) Cell mass is increasing linearly while proteome configuration is switching towards the synthesis of P. The synthesis of enzyme e_P and the secretion of P begin while the syntheses of e_B and e_R stop (ν_{e_B} and ν_{e_R} are null). The cell accumulates e_P and e_T while keeping the cell integrity, i.e. synthesis of biomass in agreement with the volume increase.

(iii) The production of all proteins stops, leading to the arrest of growth. The uptake flux of G is fully rerouted towards the synthesis of P. The proteome configuration is not totally devoted to the production of P since low concentrations of ribosomes and anabolic enzymes remain in the cell. This is in agreement with the fact that cells are not able to degrade proteins.

We would like to compare the difference between dRBA and the dFBA formulation of Gadkar et al. (2005). However, the formulation of both methods are far different. Actually dFBA does not integrate the dynamics of proteins, which allows for instantaneous changes of flux distribution over time. We thus chose to mimic dFBA conditions by removing the constraints on the positiveness of e_i for $i \in \{P, B, R, T\}$ and by constraining μ to take nonnegative values within dRBA (referred as dRBA^u in Fig. 1 and Table 1). This relaxation leads to convert 'instantaneously' a type of protein into another one. We solved the optimization problem dRBA^u for the same initial conditions. We obtained two phases, close to phase (i) and (iii) of dRBA (in grey on Fig. 1, Table 1). The cell switched abruptly from a phase of biomass production ($e_P = 0$) to a phase of P production ($e_B = e_R = 0$). This strategy is close to the bang-bang optimal strategy obtained using dFBA in Gadkar et al. (2005).

Table 1. Comparison of dRBA and dRBA^u

	dRBA Strategy	dRBA^u Strategy
$J = [P]_{ext}(t_f)/t_f$	0.38	0.44
$[P]_{ext}(t_f)$ (in $mmol.L^{-1}$)	2.33	2.45
t_f (in h)	6.08	5.58
$[M_{tot}]_{ext}$ (in $g_{CDW}.L^{-1}$)	0.32	0.29

4. CONCLUSION & PERSPECTIVES

This paper proposes a dynamical model integrating a simplified resource allocation model of cells and the bioreactor operating in batch mode. Compared to other predictive methods, using resource allocation models enables predictions, where the cell behavior over time results from a trade-off between operational constraints on cellular processes and the bioreactor. The predicted optimal strategy for maximizing the product of interest is surprisingly close to a standard steering politics (Cuthrell and Biegler, 1989). Perspectives of this work cover two main topics. The first one is to explore other bioreactor operating modes (e.g. fedbatch). The second one is to go further within the cell description since we have already a validated genome-scale cellular model of *B. subtilis* (Goelzer et al., 2015). This perspective paves the way towards a rational strategy of strain modifications where the strain and the bioreactor are designed together.

REFERENCES

Bonnans, F., Giorgi, D., Grelard, V., Heymann, B., Maindrault, S., Martinon, P., Tissot, O., and Liu, J. (2017). Bocop A collection of examples. Technical report, INRIA. URL `http://www.bocop.org`.

Bryson, A.E. and Ho, Y.C. (1975). *Applied Optimal Control*. Hemisphere Publishing, New York.

Chowdhury, A., Zomorrodi, A.R., and Maranas, C.D. (2015). Bilevel optimization techniques in computational strain design. *Computers & Chemical Engineering*, 72, 363–372.

Cuthrell, J.E. and Biegler, L.T. (1989). Simultaneous optimization and solution methods for batch reactor control profiles. *Computers & Chemical Engineering*, 13(1-2), 49–62.

Gadkar, K.G., Doyle, F.J., Edwards, J.S., and Mahadevan, R. (2005). Estimating optimal profiles of genetic alterations using constraint-based models. *Biotechnology and bioengineering*, 89(2), 243–251.

Goelzer, A. and Fromion, V. (2017). Resource allocation in living organisms. *Biochemical Society Transactions*, BST20160436.

Goelzer, A., Fromion, V., and Scorletti, G. (2011). Cell design in bacteria as a convex optimization problem. *Automatica*, 47(6), 1210–1218.

Goelzer, A., Muntel, J., Chubukov, V., Jules, M., Prestel, E., Nölker, R., Mariadassou, M., Aymerich, S., Hecker, M., Noirot, P., et al. (2015). Quantitative prediction of genome-wide resource allocation in bacteria. *Metabolic engineering*, 32, 232–243.

Jabarivelisdeh, B. and Waldherr, S. (2016). Improving bioprocess productivity using constraint-based models in a dynamic optimization scheme. *IFAC-PapersOnLine*, 49(26), 245–251.

Jeanne, G., Tebbani, S., Goelzer, A., Fromion, V., and Dumur, D. (2016). Modelling and optimization of metabolic pathways in bacteria. In *20th International Conference on System Theory, Control and Computing (ICSTCC), 2016*, 312–317. IEEE.

Kubitschek, H.E., Baldwin, W.W., Schroeter, S.J., and Graetzer, R. (1984). Independence of buoyant cell density and growth rate in *Escherichia coli*. *Journal of bacteriology*, 158(1), 296–299.

Lee, E.B. and Markus, L. (1967). *Foundations of Optimal Control Theory*. John Wiley and Sons.

Lewis, N.E., Nagarajan, H., and Palsson, B.O. (2012). Constraining the metabolic genotype–phenotype relationship using a phylogeny of in silico methods. *Nature Reviews Microbiology*, 10(4), 291.

Mahadevan, R., Edwards, J.S., and Doyle, F.J. (2002). Dynamic flux balance analysis of diauxic growth in *Escherichia coli*. *Biophysical journal*, 83(3), 1331–1340.

Marr, A.G. (1991). Growth rate of *Escherichia coli*. *Microbiological reviews*, 55(2), 316–333.

Varma, A. and Palsson, B.O. (1994). Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type *Escherichia coli* w3110. *Applied and environmental microbiology*, 60(10), 3724–3731.