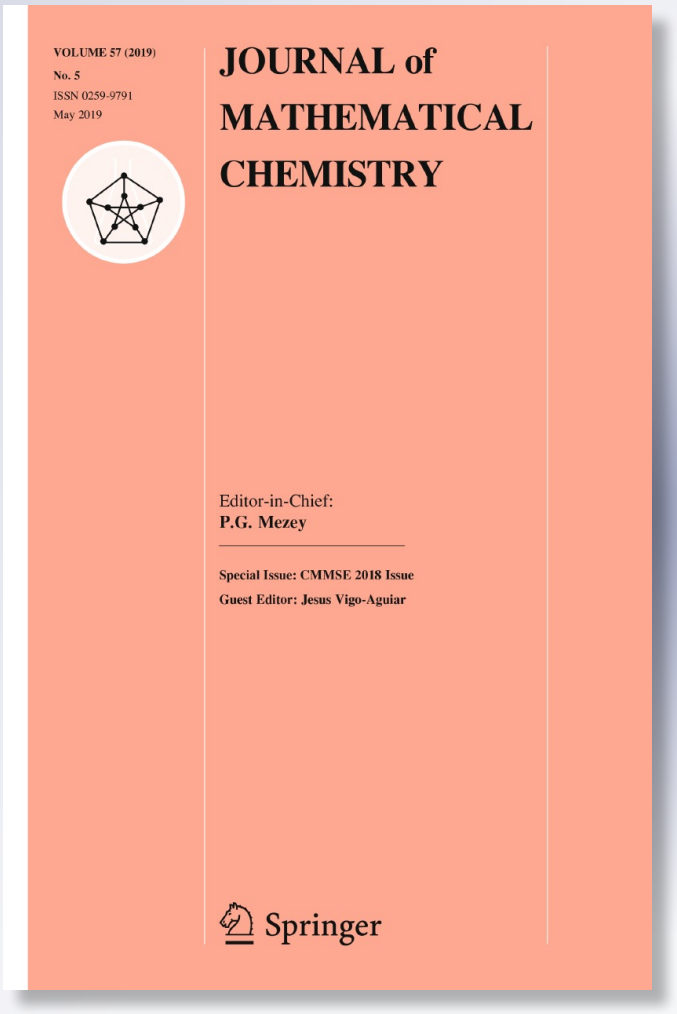
# *Influence of enzyme production dynamics on the optimal control of a linear unbranched chemical process*

**L. Bayón, P. Fortuny Ayuso, J. A. Otero, P. M. Suárez & C. Tasis**

**Journal of Mathematical Chemistry**

ISSN 0259-9791
Volume 57
Number 5

J Math Chem (2019) 57:1330-1343
DOI 10.1007/s10910-018-0969-3




Springer

Journal of Mathematical Chemistry (2019) 57:1330–1343
https://doi.org/10.1007/s10910-018-0969-3

ORIGINAL PAPER

# Influence of enzyme production dynamics on the optimal control of a linear unbranched chemical process

**L. Bayón[1] · P. Fortuny Ayuso[1] · J. A. Otero[1] · P. M. Suárez[1] · C. Tasis[1]**

Received: 11 June 2018 / Accepted: 23 October 2018 / Published online: 26 October 2018


**Abstract**
In this paper we consider the classic kinetic problem of the minimization of the operation time in which a substrate is transformed into a product in multi-step enzymatic reactions. We propose a realistic formulation that incorporates the enzyme dynamics, and present a method for obtaining the solution for an unbranched scheme and bi-linear kinetic model in an almost exclusively analytical way. The solution, which involves the exponential integral, is illustrated with several examples.

**Keywords** Optimal control · Chemical process · Enzymatic model

**Mathematics Subject Classification** 80A30 · 49M05 · 49J30

## 1 Introduction

The optimization of a chemical process involving multi-step enzymatic reactions is important in chemical kinetics. In this paper, we consider the influence of enzyme production dynamics in the minimization of the time during which the substrate is converted into the product.

When the objective is to minimize the transition time, an explicit solution for $n = 2$, can be found in [1], while, for $n = 5$, the authors solved the problem numerically. The solution for $n = 3$ is obtained quasi-analytically in [2]; in [3] we presented the quasi-analytical solution for the general case of $n$ steps. In [4] a combined optimization of the time taken to reach the new steady state and a measure of enzyme usage is considered. Later, in [5], and addressing as new objective the minimization of the operation time, we extended the theoretical analysis.

✉ L. Bayón
bayon@uniovi.es

[1] Department of Mathematics, University of Oviedo, Oviedo, Spain

In our previous papers [3,5] we assume that the enzymes can be switched on and off instantaneously. Here we propose a more realistic formulation that incorporates the enzyme dynamics.

Dynamic enzyme optimization for the activation of chemical pathways has been considered recently. In the majority of models in the literature, the enzyme degradation and dilution are modeled as linear functions of the enzyme concentration. In [4] the enzyme synthesis dynamics is considered to be linear with the expression rate. This model, also used in [6–8], will be adopted by us. We shall delve into this model and point out some details which we deem important to properly understand its physical context.

Then we present a method for obtaining the solution of an $n$-step system with unbranched scheme and bi-linear kinetic model in an almost exclusively analytical way. We develop this method using Pontryagin's Minimum Principle, as a sequence of phases or steps. We show how the solution involves, in some cases, the mathematical function called the Exponential integral. As a byproduct, we also obtain several general laws both for the enzymes and the metabolites, from which the computation of a solution becomes much easier. Finally we will present numerical examples for a test-case.

## 2 Enzymatic model

In past research [1,3,5] (and in the mayor part of [4]), the enzymes are assumed to be switched on and off instantaneously. This is a simplification because after the onset of gene expression, the enzyme is produced gradually and, after down-regulation of the encoding gene, the enzyme is gradually degraded. Therefore the enzyme synthesis cannot be as fast as required by the bang-bang profiles. A more realistic solution is obtained when the model of enzyme production dynamics is included.

The enzyme dynamic equation is a simplified representation of the processes involved in the enzyme synthesis and destruction. It is known that the concentration of enzymes is driven by two factors: the enzyme synthesis rate and the dilution through growth; it is generally assumed that the latter is the major source for protein degradation. There is also an upper bound on the rate at which it can be synthesized.

On the other hand, enzyme degradation and dilution are typically modelled as linear functions of the enzyme concentration (see, for example, [4,6–8]). In all these works, the models of the enzyme dynamics are considered to be linear in the expression rate ($u_i$). Specifically, the production of enzyme $e_i$ is described as:

$$\dot{e}_i(t) = u_i(t) - \lambda e_i(t) \tag{1}$$

where $u_i(t)$ is the expression rate of $e_i(t)$, $(i = 1, \ldots, n)$ and $\lambda$ (the same for all $i$) accounts for the dilution by cell growth and the constituent protein degradation rate. In [4,8], the following constraints are imposed to limit the amount of enzymes and their rates:

$$0 \leq u_i(t) \leq u_{i\,\max} \tag{2}$$

$$0 \leq e_i(t) \leq E_T \tag{3}$$

At this point we think it necessary to point out some details concerning this model. First of all, we shall assume, without loss of generality, following [4,8], that all the values of $u_{i\,\max}$ are the same: $u_{i\,\max} = u_{\max}\ Ms^{-1}$. Secondly, we carry out a normalization of the problem. We divide all enzyme levels by the maximum enzyme concentration; we also divide the substrate, intermediate and product levels by the initial substrate concentration. That is, after normalization, we have $E_T = 1\ M$.

From (1), one immediately deduces that the maximum of $e_i(t)$ is reached when $\dot{e}_i(t) = 0$. Therefore, this maximum is attained for:

$$E_T = \frac{u_i(t)}{\lambda} \tag{4}$$

So that, after normalization, the model must satisfy the following condition:

$$u_i(t) = \lambda \tag{5}$$

We shall see in the forthcoming sections that in our problem, the expression rate will have a bang-bang profile, so that at each time only one of them will be active, with constant value $u_{max}$ and the remaining ones with value zero:

$$u_i^*(t) = \begin{cases} u_{\max} & t \in [t_{i-1}, t_i) \\ 0 & t \notin [t_{i-1}, t_i) \end{cases} ; \quad i = 1, \ldots, n \tag{6}$$

where $t_0 = 0$. From this, Condition (5) can be restated, actually, as:

$$u_{\max} = \lambda \tag{7}$$

Figure 1 shows the concentration profile for an enzyme for different values of $\lambda$ (and, as a consequence, of $u_{max}$, due to (7)). We see how the model satisfies the constraint $e_i(t) \leq E_T$ with $E_T = 1\ M$, allowing for a simple way of adjusting the rate of increase.

However, this model presents another remarkable property. If we add all the equations in (1) for $i = 1, \ldots, n$ and use (6), we get:

$$\sum_{i=1}^{n} \dot{e}_i(t) = \sum_{i=1}^{n} u_i(t) - \lambda \sum_{i=1}^{n} e_i(t) \Rightarrow \sum_{i=1}^{n} \dot{e}_i(t) = u_{\max} - \lambda \sum_{i=1}^{n} e_i(t) \tag{8}$$

so that, if we call $S(t) = \sum_{i=1}^{n} e_i(t)$ the total sum of the concentrations of the enzymes, the following equality also holds:

$$\dot{S}(t) = u_{\max} - \lambda S(t) = \lambda(1 - S(t)) \tag{9}$$

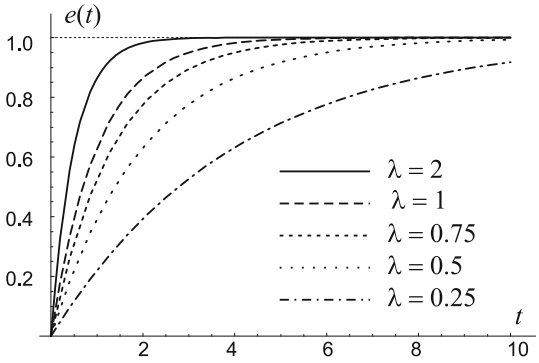



**Fig. 1** Model of the enzyme dynamics

so that the maximum value of $S(t)$ is also bounded:

$$S(t) = \sum_{i=1}^{n} e_i(t) \leq 1 \tag{10}$$

Other authors consider more complex models for particular examples. In [9,10] the enzyme degradation and dilution are also modelled as linear functions of the enzyme concentration. But in [9] the synthesis incorporates parameters over the maximal promoter activity and repression coefficient of gene and the active repressor level is also considered. On the other hand, in [10], the synthesis is modelled as the sum of a constant (or basal) expression rate: a set of Hill functions representing the activating or repressing effects of the transcription factors on the enzyme expression. We shall not consider these specific models, as our focus will be to analyze, as much as possible, the general model (1).

## 3 Statement of the problem

We consider the case of an unbranched metabolic pathway composed of $n$ irreversible reaction steps converting substrate $x_1$ into product $p$. The substrate concentration at time $t$ is denoted $x_1(t)$; $p(t)$ is the concentration of the final product; $x_i(t)$, $(i = 2, \ldots, n)$ the concentrations of the intermediate compounds, and $e_i(t)$, $(i = 1, \ldots, n)$ the concentrations of the enzymes catalyzing the $i$-th reaction:

$$x_1 \xrightarrow{e_1} x_2 \xrightarrow{e_2} x_3 \xrightarrow{e_3} \cdots \xrightarrow{e_{n-2}} x_{n-1} \xrightarrow{e_{n-1}} x_n \xrightarrow{e_n} p \tag{11}$$

We assume that the rate of the $i$-th reaction, $v_i(x_i(t), e_i(t))$ is linear in the enzyme concentrations, $e_i$:

$$v_i(x_i(t), e_i(t)) = w_i(x_i(t)) \cdot e_i(t) \tag{12}$$

and, in this paper, we use the mass action kinetic model:

$$w_i(x_i) = k_i x_i \tag{13}$$

where $k_i$ are the catalytic efficiencies of the enzymes. As the dynamical model for the pathway is driven by the law of conservation of mass:

$$\dot{x}_i(t) = v_{i-1}(x_{i-1}(t), e_{i-1}(t)) - v_i(x_i(t), e_i(t)) \tag{14}$$

and employing normalized quantities, the dynamic model is given by the set of differential equations:

$$\begin{cases} \dot{x}_1 = -k_1 e_1 x_1 & x_1(0) = 1 \\ \dot{e}_1 = u_1 - \lambda e_1 & e_1(0) = 0 \\ \dot{x}_2 = k_1 e_1 x_1 - k_2 e_2 x_2 & x_2(0) = 0 \\ \dot{e}_2 = u_2 - \lambda e_2 & e_2(0) = 0 \\ \cdots & \\ \dot{x}_n = k_{n-1} e_{n-1} x_{n-1} - k_n e_n x_n & x_n(0) = 0 \\ \dot{e}_n = u_n - \lambda e_n & e_n(0) = 0 \end{cases} \tag{15}$$

where we shall consider as initial condition, for $t = 0$, the concentrations of the intermediate compounds and of the product to be zero. Moreover, the concentrations of the compounds, $x_i$, and the enzymes, $e_i$, are positive quantities and, after normalization, the upper bound on the value of the enzymatic concentration is 1.

The objective is to minimize the operation time, i.e. transform $x_1$ into $p$ as fast as possible; however, the product $p(t)$ needs not be fully synthesized. We impose a predefined aim concentration $p(t_f) = C_f$ $(0 < C_f < 1)$, where $t_f$ is the final time, to be minimized. The initial substrate $x_1$ is assumed exhaustible. With these conditions, we obtain:

$$x_1(t_f) + x_2(t_f) + \cdots + x_n(t_f) = 1 - C_f \tag{16}$$

So that the optimization problem may be stated as:

$$\begin{array}{ll} \text{(Pr)}: & \tau_{C_f} = \min\limits_{u_1,\ldots u_n} \int_0^{t_f} dt = \min\limits_{u_1,\ldots u_n} t_f \\ \text{subject to:} & (15), (16) \\ \text{and:} & u_1(t) \geq 0, \ldots u_n(t) \geq 0; \quad u_1(t) + \cdots + u_n(t) \leq u_{\max} \end{array} \tag{17}$$

### 3.1 The exponential-integral Ei function

In the next sections a closed form solution will be found, for the first time. In some cases, it includes the Ei-function. Abramowitz and Stegun [11] first define the exponential integral as:

$$E_n(x) = \int_1^{\infty} \frac{e^{-xt}}{t^n} dt; \quad x > 0,\ n = 0, 1, \ldots \tag{18}$$

where $n$ is the order of the integral. An alternative definition is:

$$\mathrm{Ei}(x) = -\int_{-x}^{\infty} \frac{e^{-t}}{t} dt; \quad x > 0 \tag{19}$$

This definition must be understood in terms of the Cauchy principal value, due to the singularity of the integrand at zero. Note that the two previous definitions are related as follows:

$$\mathrm{Ei}(-x) = -E_1(x) \tag{20}$$

A good algorithm for evaluating Ei is to use the power series for small $x$ and the asymptotic series for large $x$. The power series is:

$$\mathrm{Ei}(x) = \gamma + \ln(x) + \sum_{k=1}^{\infty} \frac{x^k}{k \cdot k!} \tag{21}$$

where $\gamma$ is the Euler–Mascheroni constant. The asymptotic expansion is:

$$\mathrm{Ei}(x) \sim \frac{e^x}{x} \sum_{k=0}^{\infty} \frac{k!}{x^k} \tag{22}$$

## 4 Optimal solution

Problem (17) stated above is a multi-dimensional optimal control problem, with free end-time $t_f$ and free end state $\mathbf{x}(t_f)$, $\mathbf{e}(t_f)$. In general Lagrange form, it is:

$$\min_{\mathbf{u}(t)} J = \int_0^{t_f} F(\mathbf{x}(t), \mathbf{e}(t), \mathbf{u}(t), t) dt \tag{23}$$

subject to:

$$\dot{x}_i(t) = f_i(\mathbf{x}(t), \mathbf{e}(t), \mathbf{u}(t), t); \quad x_i(0) = x_{i0}; \quad i = 1, \ldots, n \tag{24}$$

$$\dot{e}_i(t) = g_i(\mathbf{x}(t), \mathbf{e}(t), \mathbf{u}(t), t); \quad e_i(0) = e_{i0}; \quad i = 1, \ldots, n \tag{25}$$

$$\mathbf{u}(t) \in U(t), \quad 0 \le t \le t_f \tag{26}$$

where $\mathbf{x}(t) = (x_1(t), \ldots, x_n(t)) \in \mathbb{R}^n$ and $\mathbf{e}(t) = (e_1(t), \ldots, e_n(t)) \in \mathbb{R}^n$ are the two state vectors, and $\mathbf{u}(t) = (u_1(t), \ldots, u_n(t)) \in \mathbb{R}^n$ the control vector. The optimum $t_f^*$ is unknown and to be determined. To solve it, we use Pontryagin's Minimum Principle (PMP) [12].

In our case, Problem (17) presents some specific properties which permit its solution in a quasi-analytic way:

– First, $F(\mathbf{x}(t), \mathbf{e}(t), \mathbf{u}(t), t) = 1$, so that the functional is

$$\min_{\mathbf{u}(t)} J = \min_{\mathbf{u}(t)} t_f \tag{27}$$

– Secondly, and this is the most remarkable property, the state equations have the form:

$$\dot{x}_i(t) = f_i(\mathbf{x}(t), \mathbf{e}(t)); \quad x_1(0) = 1, x_i(0) = 0,\ i = 2, \ldots, n \tag{28}$$

$$\dot{e}_i(t) = g_i(\mathbf{e}(t), \mathbf{u}(t)); \quad e_i(0) = 0, i = 1, \ldots, n \tag{29}$$

So that there is a decoupling between the state vectors $\mathbf{x}(t)$, $\mathbf{e}(t)$ and the control vector $\mathbf{u}(t)$. The compounds do not appear in the state equation of the enzymes, while in the state equation of the latter, the control does not appear explicitly.

– Finally, the problem is linear in the controls, wiht $u_i(t) \geq 0$ and $u_1(t) + \ldots + u_n(t) \leq u_{\max}$.

These properties allow us to propose an algorithm leading to the quasi-analytic solution of (17). This procedure has the following steps:

**Step 1: Determination of the structure of the optimal solution.** When the control appears linearly, as is the case, the optimal control is said to be of bang-bang type, because it switches between its upper and lower bounds at discrete instants (which are called the switching times). The general form of the solution can be described as follows: there exist $n$ switching times (as many as enzymes) and the optimal expression rate of the $i$-th enzyme is of bang-band type and satisfies:

$$u_i(t) = \begin{cases} u_{\max} & \text{for } t \in [t_{i-1}, t_i) \\ 0 & \text{for } t \notin [t_{i-1}, t_i) \end{cases}; \quad i = 1, \ldots, n \tag{30}$$

where $\{t_0, t_1, t_2, \ldots, t_n\}$ are the switching times, with $t_0 = 0$ and $t_n = t_f$. Certainly, the switching times are as yet unknown: their values will be kept as variables during the whole process, and they will be calculated in the last step of the algorithm.

**Step 2. Resolution of the state equations of the enzymes** Once the optimal values of the controls $u_i$ are known, we can proceed to solve the state equation of the enzymes:

$$\dot{e}_i(t) = u_i - \lambda e_i; \quad e_i(0) = 0; \quad i = 1, \ldots, n \tag{31}$$

and obtain their optimal value. Notice that (31) is not strictly speaking a *system* of differential equations but a *list* of mutually independent equations.

The way to solve these equations consists solving all the equations (31) on each interval $[t_{i-1}, t_i)$, $i = 1, \ldots, n$. We shall denote by $e_{ji}(t)$ (for $i, j = 1, \ldots, n$) the optimal $j$-th enzyme concentration in the $i$-th interval $[t_{i-1}, t_i]$. Notice that if $u_i = 1$,

from the condition: $u_1 + \cdots + u_n \leq 1$, we get: $u_j = 0,\ j \neq i,\ (i, j = 1, \ldots, n)$. By way of illustration, let us write the first steps explicitly:

1. In the first interval $[0, t_1]$, we have $u_1 = u_{\max}$; $u_i = 0,\ (i = 2, \ldots, n)$, and:

$$\begin{array}{lll} \dot{e}_1 = u_{\max} - \lambda e_1 & e_1(0) = 0 & \Longrightarrow\ e_{11}(t) = 1 - e^{-\lambda t} \\ \dot{e}_i = -\lambda e_i & e_i(0) = 0 & \Longrightarrow\ e_{i1}(t) = 0; \quad i = 2, \ldots, n \end{array} \tag{32}$$

2. Second interval: $[t_1, t_2]$. Now $u_1 = 0$, $u_2 = u_{\max}$; $u_i = 0,\ (i = 3, \ldots, n)$, and:

$$\begin{array}{lll} \dot{e}_1 = -\lambda e_1 & e_1(t_1) = e_{11}(t_1) & \Longrightarrow\ e_{12}(t) = e^{-\lambda(t-t_1)} - e^{-\lambda t} \\ \dot{e}_2 = u_{\max} - \lambda e_2 & e_2(t_1) = 0 & \Longrightarrow\ e_{22}(t) = 1 - e^{-\lambda(t-t_1)} \\ \dot{e}_i = -\lambda e_i & e_i(t_1) = 0 & \Longrightarrow\ e_{i2}(t) = 0; \quad i = 3, \ldots, n \end{array} \tag{33}$$

3. Third interval: $[t_2, t_3]$. Here $u_1 = 0$, $u_2 = 0$, $u_3 = u_{\max}$; $u_i = 0,\ (i = 4, \ldots, n)$, and:

$$\begin{array}{lll} \dot{e}_1 = -\lambda e_1 & e_1(t_2) = e_{12}(t_2) & \Longrightarrow\ e_{13}(t) = e^{-\lambda(t-t_1)} - e^{-\lambda t} \\ \dot{e}_2 = -\lambda e_2 & e_2(t_2) = e_{22}(t_2) & \Longrightarrow\ e_{23}(t) = e^{-\lambda(t-t_2)} - e^{-\lambda(t-t_1)} \\ \dot{e}_3 = u_{\max} - \lambda e_3 & e_3(t_2) = 0 & \Longrightarrow\ e_{33}(t) = 1 - e^{-\lambda(t-t_2)} \\ \dot{e}_i = -\lambda e_i & e_i(t_2) = 0 & \Longrightarrow\ e_{i3}(t) = 0; \quad i = 4, \ldots, n \end{array} \tag{34}$$

The values for the successive intervals are similarly obtained, by concatenating the solutions. The optimal solution of the complete system can be described on each interval, as on the $i$-th interval, $[t_{i-1}, t_i]$ (for $i = 2, \ldots, n-1$), there are 3 laws governing the enzyme concentrations:

(a) The concentration of the enzymes before the $i$-th are always:

$$e_{ji}(t) = e^{-\lambda(t-t_j)} - e^{-\lambda(t-t_{j-1})} \ \text{ for } j = 1, \ldots, i-1 \tag{35}$$

(b) The concentration of the $i$-th enzyme is:

$$e_{ji}(t) = 1 - e^{-\lambda(t-t_{i-1})} \ \text{ for } j = i \tag{36}$$

(c) Finally, enzymes from the $i+1$-th on have not been activated yet, so that their concentration is zero:

$$e_{ji}(t) = 0 \ \text{ for } j = i+1, \ldots, n \tag{37}$$

**Step 3. Resolution of the state equation of the compounds.** At this point, we can solve the state equation of each compound for every interval:

$$\begin{cases} \dot{x}_1 = -k_1 e_1 x_1 & x_1(0) = 1,\ x_1(t) \geq 0 \\ \dot{x}_2 = k_1 e_1 x_1 - k_2 e_2 x_2 & x_2(0) = 0,\ x_2(t) \geq 0 \\ \cdots & \\ \dot{x}_n = k_{n-1} e_{n-1} x_{n-1} - k_n e_n x_n & x_n(0) = 0,\ x_n(t) \geq 0 \end{cases} \tag{38}$$

This is truly a system of differential equations: except in the first one, the $i-1$-th compound influences the evolution of the $i$-th one. However, the special configuration of the system allows for its recursive resolution, starting at the first interval and sequentially substituting the previous computed values. Denote by $x_{ji}(t)$ (for $i,\ j=1,\ldots,n$) the optimal $j$-th metabolite concentration, in the $i$-th interval $[t_{i-1}, t_i)$. We show how this step goes in the first two intervals:

1. First interval: $[0, t_1]$.

$$\begin{array}{lll} \dot{x}_1 = -k_1 e_1 x_1 & x_1(0)=1 & \Longrightarrow\ x_{11}(t) = e^{\frac{k_1}{\lambda}(1-e^{-\lambda t}-\lambda t)} \\ \dot{x}_2 = k_1 e_1 x_1 - k_2 e_2 x_2 & x_2(0)=0 & \Longrightarrow\ x_{21}(t) = 1 - x_{11}(t) \\ \dot{x}_i = k_{i-1} e_{i-1} x_{i-1} - k_i e_i x_i & x_i(0)=0 & \Longrightarrow\ x_{i1}(t) = 0; \quad i=3,\ldots,n \end{array} \tag{39}$$

Notice that, remarkably, solving the second equation is unnecessary on this interval, as there already exists a relation between $x_{21}(t)$ and $x_{11}(t)$. Summing the first two equations and taking into account that $e_{21}(t)=0$, one gets:

$$\begin{array}{l} \dot{x}_1 = -k_1 e_1 x_1 \\ \dot{x}_2 = k_1 e_1 x_1 - k_2 e_2 x_2 \end{array} \Longrightarrow \dot{x}_1 + \dot{x}_2 = -k_2 e_2 x_2 \Rightarrow \dot{x}_1 + \dot{x}_2 = 0 \tag{40}$$

so that:

$$x_{11}(t) + x_{21}(t) = K \tag{41}$$

from which we infer, as $x_1(0)=1$:

$$x_{21}(t) = K - x_{11}(t) = 1 - x_{11}(t) \tag{42}$$

2. Second interval: $[t_1, t_2]$. In the same way as above, we get:

$$\begin{array}{lll} \dot{x}_1 = -k_1 e_1 x_1 & x_1(t_1) = x_{11}(t_1) \Rightarrow & x_{12}(t) = e^{\frac{k_1}{\lambda}(e^{-\lambda(t-t_1)}-e^{-\lambda t})-k_1 t_1} \\ \dot{x}_2 = k_1 e_1 x_1 - k_2 e_2 x_2 & x_2(t_1) = x_{21}(t_1) \Rightarrow & x_{22}(t) = \mathbb{F}(t, k_1, k_2, \lambda, t_1) \\ \dot{x}_3 = k_2 e_2 x_2 - k_3 e_3 x_3 & x_3(t_1) = 0 & x_{32}(t) = K - [x_{12}(t) + x_{22}(t)] \end{array} \tag{43}$$

for some constant $K$ to be computed later, and

$$\dot{x}_i = k_{i-1} e_{i-1} x_{i-1} - k_i e_i x_i \quad x_i(t_1) = 0 \ \Longrightarrow\ x_{i1}(t) = 0; \quad i = 4,\ldots,n \tag{44}$$

The value of the function $\mathbb{F}$ has been computed, in our case, using the symbolic algebra system Mathematica™. Notice again how, in this case, the third equation needs not be solved, as:

$$\dot{x}_1 + \dot{x}_2 + \dot{x}_3 = 0 \tag{45}$$

which implies that:

$$x_{12}(t) + x_{22}(t) + x_{32}(t) = K \tag{46}$$

taking into account that $x_{21}(t) + x_{11}(t) = 1$ by (39), this implies that $K = 1$, so that:

$$x_{32}(t) = [x_{11}(t_1) + x_{21}(t_1)] - [x_{12}(t) + x_{22}(t)] = 1 - [x_{12}(t) + x_{22}(t)] \quad (47)$$

This remarkable result can be generalized for the successive intervals with the general law (obtained with an elementary computation):

$$x_{ji}(t) = 1 - \sum_{k=1}^{j-1} x_{ki}(t) \quad \text{for } i = 1, \ldots, n-1, \quad j = i+1, \ldots, n \quad (48)$$

The values for each successive interval are similarly obtained, and the global solutions are concatenations of these partial ones.

Another adequate computation provides a new general law.

3. In the next interval $[t_2, t_3]$, one can easily verify that:

$$\dot{x}_1 = -k_1 e_1 x_1 \; x_1(t_2) = x_{12}(t_2) \implies x_{12}(t) = e^{\frac{k_1}{\lambda}(e^{-\lambda(t-t_1)} - e^{-\lambda t}) - k_1 t_1} \quad (49)$$

so that:

$$x_{12}(t) = x_{13}(t) \quad (50)$$

and so on, on each interval. Thus, the optimal profile of the $j$-th metabolite follows the same function in the intervals $[t_{i-1}, t_i]$ for $i = j+1, \ldots, n$:

$$x_{ji}(t) = x_{ji+1}(t) = \cdots = x_{jn}(t) \quad (51)$$

In the end, on each interval, one needs only calculate 3 formulas for the $j$-th metabolite (for $j = 2, \ldots, n$), as $x_{ji}(t) = 0$ for $i \leq j-2$ and $x_{ji}(t) = x_{jj+1}(t)$ for $i \geq j+2$:

$$x_{jj-1}(t), \quad x_{jj}(t) \text{ and } x_{jj+1}(t) \quad (52)$$

which greatly simplifies the computation of the solution.

**Step 4. Calculation of the switching times.** Once the optimum values for states and controls have been obtained, the values of the unknowns (the switching times $t_1, \ldots, t_{n-1}$ and the operation time $t_f$) need to be computed. In order to do this, we use restriction (16) which remains unused. Following the Lagrange multipliers method, we define the augmented functional:

$$L(t_1, t_2, \ldots, t_{n-1}, t_f, \beta) = t_f + \beta(x_{1n}(t_f) + x_{2n}(t_f) + \cdots + x_{nn}(t_f) - C_f) \quad (53)$$

where the values of the compounds $x_{1n}(t_f), x_{2n}(t_f), \ldots, x_{nn}(t_f)$ are already known (as functions of the switching times). The nonlinear system of $n+1$ equations:

$$\frac{\partial L}{\partial t_1} = 0; \frac{\partial L}{\partial t_2} = 0; \ldots; \frac{\partial L}{\partial t_{n-1}} = 0; \frac{\partial L}{\partial t_f} = 0; \frac{\partial L}{\partial \beta} = 0 \tag{54}$$

gives the final solution to the optimization problem.

As a matter of fact, the special structure of the previous system permits the elimination of the last-but-one equation so that there are only $n$ of them:

$$\frac{\partial L}{\partial t_1} = 0; \frac{\partial L}{\partial t_2} = 0; \ldots; \frac{\partial L}{\partial t_{n-1}} = 0; \frac{\partial L}{\partial \beta} = 0 \tag{55}$$

This eliminates one equation, thus simplifying somehow the resolution of the (itself already quite complicated) nonlinear system.

## 5 Numerical example

We now give the complete solution to a numerical example and then analyze the influence of the degradation and dilution parameter $\lambda$.

### 5.1 Base case

This example is inspired by the test case presented in [2] and later in [8], and which is a *three-step linear pathway with mass action kinetics* (LPN3B).

$$x_1 \xrightarrow{e_1} x_2 \xrightarrow{e_2} x_3 \xrightarrow{e_3} p \tag{56}$$

The pathway consists of three reactions with mass action kinetics. Each reaction is catalyzed by a specific enzyme $e_i$; $x_1$ corresponds to the substrate, $x_2$ and $x_3$ to the intermediate metabolites and $p$ to the product. The objective is the minimization of the time needed to reach a 90% amount of product (i.e. $C_f = 0.9\ M$).

The authors of [2,8], consider the ideal case of an unalterable substrate (buffered substrate concentration) and in which the enzymes can be instantly activated. We propose here, as discussed above, two improvements in the modelization, in order to obtain a more realistic chemical model: (i) we consider the consumption of the substrate and (ii) we take into account the enzyme dynamics.

We set, in our computations, $k_i = 1.0\ s^{-1}$ and $\lambda = 1.0 s^{-1}$. In Fig. 2, the optimal enzyme profile is shown and in Fig. 3, the optimal metabolite and product profiles. The three optimal switching times, in sec, are:

$$t_1 = 3.39406; \quad t_2 = 6.82934; \quad t_3 = 11.1909 \tag{57}$$

where the last one is, obviously, $t_f$. At that time, 90% of the product $p(t)$ is obtained, as seen in Fig. 3. Notice, also in that figure, how the substrate $x_1(t)$ is exhaustible. In

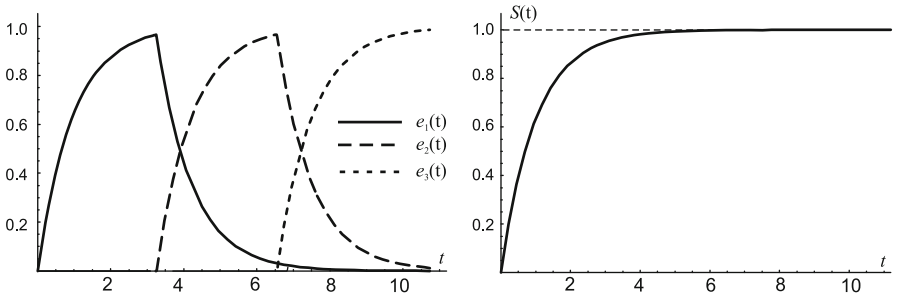



**Fig. 2** Optimal enzyme profile

**Fig. 3** Optimal metabolite and product profile

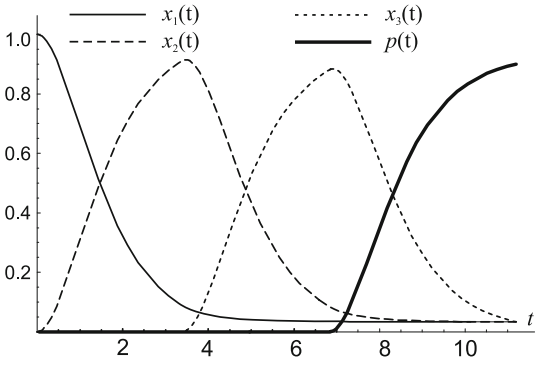



Fig. 2 one can also realize how the restriction:

$$S(t) = e_1(t) + e_2(t) + e_3(t) \leq 1 \tag{58}$$

is obeyed and how the maximum is asymptotically sought in order to reach the maximum speed. By way of example, we show the expression of the function $\mathbb{F}$ for $x_{22}(t)$ in this case:

$$\begin{aligned} x_{22}(t) = e^{-e^{t_1}-t-e^{-t_1}-t_1-t} \cdot \Big[ & -e^{e^{-t_1}} \left(e^{t_1} - 1\right) \mathrm{Ei}\left(e^{-t}\left(-1 + 2e^{t_1}\right)\right) \\ & + e^{e^{-t_1}} \left(e^{t_1} - 1\right) \mathrm{Ei}\left(2 - e^{-t_1}\right) + e^{t_1+1}\left(e^{t_1+e^{-t_1}} - e\right)\Big] \end{aligned} \tag{59}$$

The remaining ones are computed similarly.

### 5.2 Sensitivity analysis

Next, we analyze the influence of the value of the degradation and dilution parameter $\lambda$ which, as we stated above, permits us to gauge the growth speed of the enzymes. Table 1 shows the different values of the optimal switching and operation times for the case $n = 2$.

**Table 1** Switching times and operation time of the optimal solution, $n = 2$

| $\lambda\ (s^{-1})$ | $t_1\ (s)$ | $t_2 \equiv t_f\ (s)$ |
|---|---|---|
| 2 | 2.93786 | 6.40382 |
| 1 | 3.02187 | 6.99172 |
| 0.75 | 3.07630 | 7.35311 |
| 0.50 | 3.21402 | 8.04757 |
| 0.25 | 3.66502 | 9.85349 |

**Table 2** Switching times and operation time of the optimal solution, $n = 3$

| $\lambda\ (s^{-1})$ | $t_1\ (s)$ | $t_2\ (s)$ | $t_3 \equiv t_f\ (s)$ |
|---|---|---|---|
| 1 | 3.39406 | 6.82934 | 11.1909 |
| 0.5 | 3.51875 | 7.18176 | 12.4201 |
| 0.25 | 3.85307 | 8.04811 | 14.6923 |

As in Figure 1, the less the value of $\lambda$, the less the growth speed of the enzymes, which obviously influences the value of the operation time $t_f$. In Table 2, similar data is shown for the case $n = 3$ and the same slowing down of the enzymes is visible for decreasing $\lambda$.

Notice how the values $t_f = 5.93948(s)$ for $n = 2$, computed in [5] and $t_f = 10.0995(s)$ for $n = 3$, in the cases where the enzymes can be instantaneously activated and their dynamics are not considered, are still (logically) smaller.

## 6 Conclusions

The influence of enzyme production dynamics over the optimal control of a linear unbranched chemical process has been addressed. The fact that the enzymes are not instantly activated provides a more realistic modelling at the cost of a more complex optimization problem. A resolution algorithm which allows tackling and solving the problem in a quasi-analytic way has been provided. General laws for the enzyme evolution and the metabolites have been found which provide simplifications and reduce the computational cost of the resolution process. These laws have been found with the symbolic algebra system Mathematica$^{TM}trademark$ and in some cases involve the function Ei. The model of enzyme dynamics studied conjoins simplicity and versatility when trying to properly study a large family of chemical reactions.

## References

1. E. Klipp, R. Heinrich, H.G. Holzhutter, Prediction of temporal gene expression. Metabolic optimization by re-distribution of enzyme activities. Eur. J. Biochem. **269**(22), 5406–5413 (2002)
2. M. Bartl, P. Li, S. Schuster, Modelling the optimal timing in metabolic pathway activation—use of Pontryagin's Maximum Principle and role of the Golden section. BioSystems **101**, 67–77 (2010)
3. L. Bayon, J.M. Grau, M.M. Ruiz, P.M. Suarez, Optimal control of a linear unbranched chemical process with steps: the quasi-analytical solution. J. Math. Chem. **52**(4), 1036–1049 (2014)

4. D. Oyarzun, B. Ingalls, R. Middleton, D. Kalamatianos, Sequential activation of metabolic pathways: a dynamic optimization approach. Bull. Math. Biol. **71**(8), 1851–1872 (2009)
5. L. Bayon, J.A. Otero, M.M. Ruiz, P.M. Suarez, C. Tasis, Sensitivity analysis of a linear unbranched chemical process with n steps. J. Math. Chem. **53**(3), 925–940 (2015)
6. D. Oyarzun, R.H. Middleton, Optimal adaptation of metabolic networks in dynamic equilibrium, in *Proceedings of 2011 American Control Conference*, San Francisco, CA, USA, pp. 2897–2902 (2011)
7. M. Bartl, M. Kotzing, S. Schuster, P. Li, C. Kaleta, Dynamic optimization identifies optimal programmes for pathway regulation in prokaryotes. Nat. Commun. **4**(2243), 1–9 (2013)
8. G. de Hijas-Liste, E. Klipp, E. Balsa-Canto, J. Banga, Global dynamic optimization approach to predict activation in metabolic pathways. BMC Syst. Biol. **8**(1), 1–15 (2014)
9. A. Zaslaver, A. Mayo, R. Rosenberg, P. Bashkin, H. Sberro, M. Tsalyuk, M. Surette, U. Alon, Just-in-time transcription program in metabolic pathways. Nat. Genet. **36**(5), 486–491 (2004)
10. J. Kuntz, D. Oyarzun, G.B. Stan, Model reduction of genetic-metabolic networks via time scale separation, in *A Systems Theoretic Approach to Systems and Synthetic Biology I: Models and System Characterizations*, ed. by V.V. Kulkarni, G.-B. Stan, K. Raman (Springer, Berlin, 2014), pp. 181–210
11. M. Abramowitz, I.A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1972)
12. A. Chiang, *Elements of Dynamic Optimization* (Waveland Press, Long Grove, 2000)