OPTIMIZING THE DILUTION PROTOCOL IN CONTINUOUS CELL CULTURES

OPTIMIZACIÓN DEL PROTOCOLO DE DILUCIÓN EN CULTIVOS CELULARES CONTINUOS

B. A. Pérez-Fernández[a,b†], R. Mulet[a,c]

a) Group of Complex Systems and Statistical Physics, Physics Faculty, University of Havana, San Lázaro and L, 10400 La Habana, Cuba
b) Department of Applied Physics, Physics Faculty, University of Havana, San Lázaro and L, 10400 La Habana, Cuba; barbara.perezf@fisica.uh.cu[†]
c) Department of Theoretical Physics, Physics Faculty, University of Havana, San Lázaro and L, 10400 La Habana, Cuba
† autor para la correspondencia

Due to its potential impact on the biotech industry, continuous cell cultures have attracted great interest in recent years. However, this type of culture remains expensive and, in some aspects, inefficient. They are often optimized based on the empirical experience of the personnel involved in their development and maintenance, but this is time-consuming and leads to material and nutrients waste that could be reduced with a deeper understanding of the process. We propose an optimization method for continuous cell culture applying techniques from the theory of Optimum Control. As a proof of concept, we apply this method in a simple metabolic network that consumes glucose and secretes lactate. We predict the most appropriate behavior for the dilution rate that maximizes the cell density achieved in the steady state of the system.

Los cultivos celulares continuos han alcanzado un potencial impacto en la industria biotecnológica, lo que ha despertado un gran interés en la comunidad científica en los últimos años. Sin embargo, este tipo de cultivo sigue siendo caro y, en algunos aspectos, ineficiente. Comúnmente, el proceso de producción es optimizado en base a la experiencia empírica del personal involucrado en su desarrollo y mantenimiento, lo cual requiere mucho tiempo y conduce a desperdicios de materiales y nutrientes que podrían reducirse con una comprensión más profunda del proceso. Nosotros proponemos un método de optimización para cultivo celular continuo utilizando Teoría de Control. Como prueba de concepto, aplicamos este método a una red metabólica simple que consume glucosa y secreta lactato. Predecimos el comportamiento de la tasa de dilución que maximiza la densidad celular alcanzada en el estado estacionario del sistema.

I. INTRODUCTION

Mammalian cell cultures are at the core of the productions of many therapeutic proteins and viral vaccines [1]. This has made it critical for pharmaceutical companies to improve their stability and efficiency. All the more reason when it comes to complex production systems such as continuous processes. In this case, fresh medium is continually added, while the culture liquid containing the leftover nutrients, metabolic end products, and toxic by-products is continuously removed to maintain the culture volume.

This type of large-scale production mode has attracted great interest in recent years. [2]. However, despite enormous advances in understanding and manipulation, continuous bioprocessing remains time consuming, expensive, and in some respects even an inefficient process [3].

In principle, we can distinguish four different ways to improve this process [4]: the selection of the strain, the formulation of a culture medium, the selection of critical parameters and the definition of the optimal protocol.

The strain is usually selected considering its resistance to changes in the environmental conditions of the culture and the possibility of manipulating them to obtain the products of interest to the industry [5]. The formulation of the culture media is even more delicate. To lower costs, the industry would like a medium that is as cheap as possible, but at the same time it should be rich enough in nutrients to guarantee maximum growth [6] or protein production in the culture. On the other hand, setting parameters such as temperature, *pH* and oxygen concentration, is a very complex process. This generally requires prior knowledge of the values used for similar strains, or alternatively it is necessary to establish it after numerous experiments, implying a waste of time and resources [7].

But setting the values of specific parametes is not the only problem, the biggest difficulty is finding an efficient time-dependent protocol for each parameter. For example, in a perfusion system, controlling the dilution rate at the exponential stage of cell growth significantly influences the growth rate and the cell density achieved at steady state [8,9].

In this work, we will concentrate our efforts in this last problem: what is the proper protocol of dilution rate that maximizes the cell density of the system at the steady state? As a proof of concept, to answer this question, we will study a toy model of cellular metabolism exploiting results from the Theory of Optimum Control.

Previously, a similar theoretical approach was applied to this problem, using empirical kinetic models for the reactions [10]. For example, the Monod equations that correlate cell specific growth rate with extracellular substrate concentration by two kinetic parameters [11].

Indeed, the Monod equation is one of the best known processes that describe kinetic reactions in the cell. It has a simple formulation and has been extensively validated [12, 13]. However, several studies have shown that growth kinetic constants are actually a function of culture history and composition of the cell community [14, 15]. On the other, for genomic-scale cell lines, it is almost impossible to know all the kinetic parameters that describe the cell growth rate. In particular, for complex lines, such as mammalian cells, it is difficult to know the real form of the kinetic functions.

To avoid these difficulties, we will substitute the use of kinetic equations by the standard assumptions of Flux Balance Analysis (FBA). FBA allows treatment the cell metabolism in steady state, does not require the use of kinetic parameters, and has been shown to be a good approximation to estimate the cell growth rate [16].

The rest of the work is organized as follows: In the next section, we will present the standard equations that describe the dynamics of the chemostat and a toy model that describe the metabolism of the cell. There, we will show what is the behavior of said system under standard protocols, also referring to the relevant literature where the reader can seek a more in-depth and detailed analysis of the model. We proceed by providing a brief review of the Theory of Optimum Control and in particular of the technique used in our work. Then, we show and discuss the optimal protocol to be used in our model. Finally, we present the conclusions of our work and suggest possible extensions to our approach.

II. THE MODEL

This section is divided into two parts. We first describe the equations that define the behavior of the perfusion system. Essentially, how the number of cells and the concentration of metabolites in the culture change depending on the external parameters of the problem. Next, we present the Toy model mimicking the metabolism of cells. We will remark the connection between these two seemingly disconnected systems of equations.

II.1. The perfusion system

A cell culture produced in continuous mode can be represented schematically as in figure 1. The fundamental dynamic equations that describe this system are [17]:

$$\frac{dX}{dt} = (\mu - \phi D)X \qquad X(0) = X_0 \tag{1}$$

$$\frac{ds_i}{dt} = -u_i X - (s_i - c_i)D \qquad s_i(0) = s_{i0} \tag{2}$$

where X denotes the cell density in the bioreactor (units: gDW/L), μ the effective cell growth rate (units: 1/h), and the perfusion coefficient, ϕ (unit-less), which characterizes the fraction of cells that escape from the culture through a
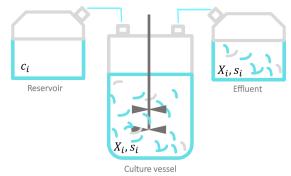
cell-retention device.

Figura 1. Cells are growing in a bioreactor that is continuously fed fresh medium at a constant flux. While an equivalent flux transports metabolic end products and toxic by-products out of the tank, keeping the culture volume constant. Notation: substrate concentrations in the medium (c_i), cell density and metabolite concentrations in the culture (X, s_i), dilution rate ($D = F/V$, where F is the influx/outflux and V the culture volume).

Equation (1) describes a process in which the number of cells grows at a rate of μ (growth rate), and is removed from the culture at a rate of ϕD. On the other hand, equation (2) describes a process in which metabolite i is removed/injected at a rate $(s_i - c_i)D$ while u_i denotes the specific uptake of metabolite i (units: mM) from the cells. The variable s_i denotes the concentration of metabolite i in the culture (units: mM), while c_i is the concentration of metabolite i in the media formulation. In what follows, i will take on two values that describe glucose consumption and lactate production.

Here two things must be taken into account, first that this set of equations depends on the actual metabolism of the cells through u_i and μ and that the external parameters that control the culture are the dilution rate, D (units: 1/h) and the perfusion coefficient ϕ.

II.2. A Toy Metabolic Network

We study a simple metabolic network first introduced by A. Vazquez et al. [18] to explain the Warburg effect and later exploited by Fernández de Cossio and collaborators [19] to study the effects of toxic by-products of the metabolism, the heterogeneity in a chemostat [20] and the bistability of a Warburg-like phenotype due to microenviornmental cooperation [21]. A diagram of the network is shown in figure 2.

To make the discussion a bit more realistic, we assume that the nutrient consumed is glucose, c_g, and the waste product is lactate, s_l, a well-recognized toxic waste in mammalian cell cultures. The cell can consume c_g from the medium at a rate $u_g > 0$, and secrete lactate at a rate $u_l < 0$. The stoichiometric relationships for this model are summarized in equations (3) and (4) from the conservation of metabolite P and the mass balance constraint of the cell:

$$u_g + u_l - r = 0 \tag{3}$$

$$N_F u_g + N_R r - e - yZ = 0 \tag{4}$$

where N_F and N_R define the contribution of these reactions to the growth rate of biomass of the cell Z. The parameter

e denotes an energetic maintenance demand, modeled as a constant drain of ATP [17,22]. The term *y* represents the units of metabolite consumed, in this case it is glucose, required per unit of biomass produced [23,24].
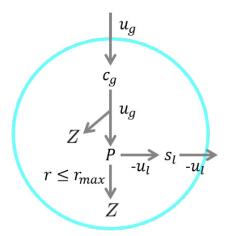

Figura 2. A primary nutrient c_g is consumed by the cell at a rate $u_g \geq 0$. This is processed and the intermediary P is obtained, generating N_F units of energy for each unit of c_g consumed. P can be excreted as a waste product s_l at a rate $(-u_l \leq 0)$ (fermentation), or it can be completely oxidized at a rate of $r \geq 0$ (respiration) generating N_R additional energy units ($N_R \gg N_F$). Respiration is valid for $r \leq r_{max}$.

We also introduce the role of the enzymatic costs [17, 19]. Since this cost is higher in the respiration reaction than during fermentation, we limit, for simplicity, only the flux *r*:

$$r \leq r_m = 0{,}45 \text{ mmol/gDW/h} \tag{5}$$

Now, since metabolite concentrations in the culture cannot be negative ($s_i \geq 0$), Equation(2) in steady state implies that $u_i \leq c_i D/X$. In our setting:

$$u_g \leq \text{Min}(V, c_g D/X) \tag{6}$$

where *V* is a constant maximum uptake rate determined by molecular details of the transport process. This bound (Equation (6)) connects the variables in the chemostat with the actual metabolism of the cell.

Moreover, due to the enzymatic cost of respiration shown in the Equations (3) and (5), the flux *r* takes the form

$$r \leq \text{Min}(u_g, r_m) \tag{7}$$

Therefore, substituting the expressions (6) and (7) in (3), we obtain:

$$u_l = r - u_g = \text{Min}(u_g, r_m) - u_g \tag{8}$$
$$= \text{Min}(\text{Min}(V, c_g D/X), r_m) - \text{Min}(V, c_g D/X) \tag{9}$$

Notice that he study of cell death has been an important area in biomedical research for several years. Many industrially important cell lines have been shown to die by apoptosis (see, for example, the original work of Al-Rubeai et al. [25]). This phenomenon has been observed mainly at the end of the batch culture, and it is due to the deprivation of nutrients such as glucose, glutamine, growth factors and oxygen or due to the presence of toxic metabolites such as ammonia and lactate [26–28].

Nevertheless, in batch feeding, perfusion, and chemostat cultures, the situation is more complex, as nutrients are constantly being replenished. In these type of cultures, cells can undergo apoptosis in response to various environmental stressors, such as high osmolality, pH fluctuations, oxygen gradients, and deprivation due to insufficient mixing at high cell densities [26, 29–33]. In addition, exposure to increasing hydrodynamic forces from liquid surface bubble burst, gas bubbling, liquid flow or gas entrainment and energy dissipation from the impeller stream can induce apoptosis [31, 34]. Consequently, it is very difficult to obtain a mathematical expression that regulates all these factors involved in continuous or semi-continuous cultures. With more reason in a metabolic model as simple as ours. In particular, in our work we do not take into account bioreactor damage, osmolality, pH fluctuations or oxygen gradients; therefore, we do not take into account cell death due to apoptosis.

However, more generally, we can consider cell death due to the accumulation of toxic by-products, which affects the growth rate. Lactate and ammonia have been shown to inhibit and destroy cells [35–37]. Specifically, the effects of lactate, a metabolite used in our model as s_l, are incorporated in the form of growth inhibition and increased death rate as follows:

$$\mu = Z - \tau s_l \tag{10}$$

where μ is the effective cell growth, Z represents the biomass production and τ is the death rate.

In short, the internal metabolism of the cell is defined by stoichiometric equations (3) and (4), subject to the bounds (6), (7) and (9). The effective growth rate is given by (10).

In practice, we used: $N_F = 2$, $N_R = 38$ as the characteristic ATP yields of glycolysis and respiration respectively [38], $e = 1{,}0625$ mmol/gDW/h typical of mammalian cells [22], $y = 348$ mmol/gDW is within the range of duplication rates in mammalian cells [36, 39], $V = 0{,}5$ mmol/gDW/h is the maximum uptake of glucose measured for HeLa cells [40] and the toxicity of waste is $\tau = 0{,}0022$ h^{-1}mM^{-1} obtained from linearizing the death rate dependence on lactate in a mammalian cell culture reported by S. Dhir et al. [36].

II.3. Stability of the System

To determine the stability of a system of coupled differential equations, on must analyze the Jacobian eigenvalues of the system. If the real parts are all negative the state is stable, but if at least one eigenvalue has a positive real part, the state is unstable.

Our system of equations (1) and (2) was previously presented by Cossio et al. in the report [17]. In that work, the authors carried out a detailed and very general study of the stability of the system of equations. For the particular set of parameters used in our work, the system is always stable.

II.4. The General Picture

To get a clear picture of what to expect from the model presented above, we summarize some general results here. A relevant question is the following: Given a dilution rate D set by the experimenter, how do cell density and metabolite concentration evolve over time?
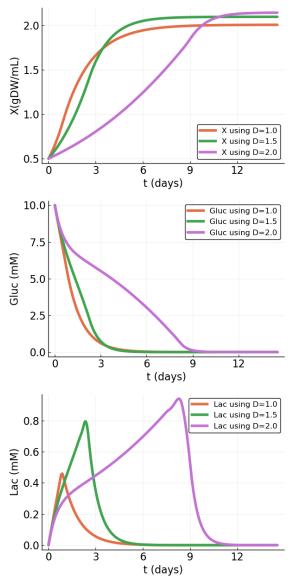

Figura 3. Behavior of cell density and concentrations of nutrients, glucose and lactate, in the culture as a function of time for three different dilution rates ($1{,}0$ days^{-1}, $1{,}5$ days^{-1} and $2{,}0$ days^{-1}). Initial condition, $X_0 = 0{,}5$ gDW/mL, and $c_{g_0} = 10$ mM.

The answer to this question can be obtained by means of the numerical solution of the dynamic system (Equations (1)-(2)) with the addition of the FBA hypothesis for the metabolism of the cell, for different values of D. As an example, we show in figure 3, for three different values of D, how the cell density and the concentrations of the glucose and lactate metabolites vary as a function of time.

As can be seen in the figure 3, the dilution rate affects both, the cell density at the steady state and the time required to reach that steady state. Of Course, it also influences the rate at which metabolites are delivered to the culture, and thus changes the

concentrations of metabolites in the vessel, including toxic ones like lactate.

To summarize the previous image, in figure 4 we present, as a function of D, the concentration of cells reached at the steady state. Notice, that X first increases, essentially, because increasing D increases the rate at which nutrients are delivered in the system and this facilitates the growth of the culture. However, if D is too large, ($D > 2$ d^{-1}) the removal of the cells from the system ocurrs so rapidly that the culture is essentially washed out and X is drastically decreased. In this regime, cells just don't growth fast enough.
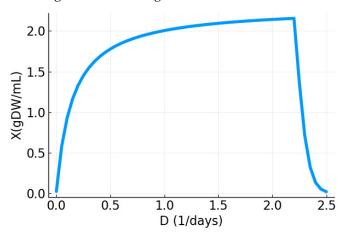

Figura 4. Expected cell density at a stable state for different constant flux of dilution rates.

In the rest of the work, our task is to suggest a suitable protocol for the experimenters in which, controlling and changing D the number of cells at steady state is maximized as quickly as possible. But, before presenting this protocol, we are going to introduce below some concepts and techniques of the Optimal Control Theory.

III. OPTIMAL CONTROL PROBLEM

In Optimal Control Theory [41], you usually solve problems in which you need to maximize a utility function. In this case, it is the cell density X. This maximization occurs through the use of one or more variables from the problem, usually call control parameters. In our case, the control parameter will be the dilution rate D. The difficulty arises because the function to be maximized is constrained to following specific dynamic equations [41], as in our case equations (1) and (2).

Let us, put this in more explicit form. We want to find the optimal control parameter $u(t) = D(t)$ that maximizes $X(t_f)$, where t_f is an unknown final time of the experiment. Similar to the Lagrangian Mechanics formalism [42], this is of course equivalent to minimizing the following objective function:

$$F = -X(t_f) \tag{11}$$

subject to the equations (1) and (2), that for simplicity we rewrite in vectorial form

$$\dot{\mathbf{y}} = \mathbf{g}[\mathbf{y}(t), u(t)] \quad \mathbf{y}(0) = \mathbf{y}_0 \tag{12}$$

The above problem is equivalent to minimizing the augmented functional

$$J = \int_0^{t_f} [H - \lambda^T \dot{\mathbf{y}}]dt = \int_0^{t_f} [F + \lambda^T(-\dot{\mathbf{y}} + \mathbf{g})]dt \tag{13}$$

subject to $\mathbf{y}(0) = \mathbf{y_0}$; where $H = F + \lambda^T \mathbf{g}$ is the Hamiltonian, λ the Lagrange multipliers and $H - \lambda^T\dot{\mathbf{y}}$ in the language of Theoretical Mechanics is the Lagrangian L of the problem.

It is easy to show that the necessary conditions for the minimum of F are: [41]

$$\begin{cases} \dot{\mathbf{y}} = \mathbf{g}, \quad \mathbf{y}(0) = \mathbf{y_0} & \text{State equations} \\ \dot{\lambda} = -\mathbf{H_y}, \quad \lambda(\mathbf{t_f}) = \mathbf{0} & \text{Costate equations} \\ H_u = \mathbf{0} \\ H(t_f) = 0 \end{cases}$$

III.1. Improvements in u and t_f to reduce J

To implement the improvement strategy numerically [41], we split the time interval $[0, t_f]$ into N subintervals of equal length and use numerical integration (for example, composite Simpson's 1/3 Rule for even number of subintervals) to calculate δJ as well as J. Thus, J is rendered into a function dependent on the vector of optimization parameters

$$p \equiv [u[t_0] \quad u[t_1] \quad \dots \quad u(t_N) \quad t_f]^T \tag{14}$$

where $t_0, t_1, \dots, t_N$ form the time-grid of $(N + 1)$ equispaced grid points in the time interval $[0, tf]$.

However, there is a complication with this approach. An improvement in t_f changes the time grid, thereby requiring the estimation of controls and states on the new time grid for the next round of improvements. We avoid this situation by linearly transforming the independent variable t in the variable interval $[0, t_f]$ to a new independent variable σ in the fixed interval $[0, 1]$.

Besides, this improvement is ideal in our problem because we don't know the value of the final time t_f. We need to minimize, F, but both t_f and $X(t_f)$ are free. We introduce a new independent variable $\sigma \in [0, 1]$ through the linear relation

$$\sigma = at + b \tag{15}$$

where a and b are some unknown constants.

Thus,

$$\sigma = \frac{t}{t_f} \quad \frac{d\sigma}{dt} = \frac{1}{t_f} \quad and \quad \frac{dy_i}{dt} = \frac{dy_i}{d\sigma}\frac{d\sigma}{dt} = \frac{dy_i}{d\sigma}\frac{1}{t_f} \tag{16}$$

Substituting, respectively, the initial and final values of σ and t in the above equation, we obtain

$$0 = b \quad and \quad 1 = at_f + b \quad \Rightarrow \quad a = \frac{1}{t_f} \tag{17}$$

Based on the these relations, the objective of the optimal control problem is to find the control u and the final time t_f that minimize the functional (Eq.(13))

$$J = \int_0^{t_f} L(\mathbf{y}(t), u(t))\frac{d\sigma}{d\sigma}dt = \int_0^1 t_f L(\mathbf{y}(t), u(t))d\sigma \tag{18}$$

We compute the integral in the above equation numerically over the fixed σ-interval $[0, 1]$ using composite Simpson's 1/3 Rule (see the Appendix). The interval is split into N subintervals of equal length using $(N + 1)$ equi-spaced grid points.

$$\sigma_0 = 0, \quad \sigma_1, \quad \sigma_2, \quad \dots, \quad \sigma_N = 1 \tag{19}$$

We initialize the control parameter u with N arbitrary values and $t_f = 1{,}0$. Then, we integrate state equations ($\dot{\mathbf{y}} = \mathbf{g}$) using the initial conditions and the control function values, and save the values of state variables at each grid point. After, we evaluate the objective functional (18) using the values of control functions and state variables and we save the objective functional value. Finally, we integrate costate equations ($\dot{\lambda} = -\mathbf{H_y}$) using the final conditions, the control function values, and the saved values of the state variables. It is also necessary to save the values of costate variables at each grid point.

Now, we improve u and t_f using the gradient of the objective functional as follows:

$$u_i^{k+1} = u_i^k - \epsilon H_{u,i}^k, \quad i = 0, 1, \dots, N \tag{20}$$

$$t_f^{k+1} = t_f^k - \epsilon J_{t_f}^k \tag{21}$$

where $\epsilon \equiv \epsilon_0 / \parallel \Delta J \parallel$, and ϵ_0 is some positive real number.

This minimization procedure is known as the Gradient Algorithm. It affords a simple and effective way to solve a wide range of optimal control problems [41].

III.2. Penalty function method

On the top of the previous discussion, we have to consider also that in some cases control parameters most fulfill specific restrictions. For example, in our case that:

$$D \ge D_0 \tag{22}$$

A constraint that can be generally expressed as

$$f(\mathbf{y}, u) = -D + D_0 \le 0$$

To solve this additional restriction problem, we use the *Penalty Function Method* [41]: This is a simple method for solving an optimal control problem with inequality constraints. As the name suggests, the method penalizes the objective functional in proportion to the violation of the constraints, which are not enforced directly. A constrained problem is then solved

using successive applications of an optimization method with increasing penalties for these constraint violations. This strategy gradually leads to the solution, which satisfies the inequality constraints.

The problem is equivalent to minimizing the augmented functional

$$J = \int_0^{t_f} [F + \lambda^T(-\dot{\mathbf{y}} + \mathbf{g}) + \mu_\lambda^T f]dt = \int_0^{t_f} (L - \lambda^T \dot{\mathbf{y}})dt \qquad (23)$$

where, in general, μ_λ are time dependent multipliers associated with the algebraic constraints f and the Lagrangian is

$$L = F + \lambda^T \mathbf{g} + \mu_\lambda^T f = H + \mu_\lambda^T f$$

The necessary conditions for the minimum of J are

$$\dot{\mathbf{y}} = \mathbf{g} \qquad \mathbf{y}(0) = \mathbf{y_0} \qquad \dot{\lambda} = -L_\mathbf{y} \qquad \lambda(t_f) = 0$$
$$f \leq 0 \qquad L_u = \mathbf{0} \qquad L(t_f) = 0 \qquad \mu_\lambda \geq \mathbf{0} \qquad \mu_\lambda^T L_{\mu_\lambda} = \mathbf{0}$$

To handle the inequalities, the penalty function method is slightly modified as follows. At any time, the multipliers corresponding to the inequalities are prescribed as

$$\mu_\lambda = \mathbf{W}h \qquad (24)$$

where $\mathbf{W}$ is an $l \times l$ diagonal weighting matrix with all positive diagonal elements and h is

$$h = \begin{cases} 0 & \text{if } f(t) \leq 0 \\ \\ 1 & \text{if } f(t) > 0 \end{cases}$$

Thus, the augmented objective functional

$$J = \int_0^{t_f} [F + \lambda^T(-\dot{\mathbf{y}} + \mathbf{g}) + h\mathbf{W}f]dt$$

has the penalty function $h\mathbf{W}f$, which is positive and enlarges J whenever any inequality is violated. The computational algorithm to find the minimum is in the Appendix.

IV. RESULTS AND DISCUSSION

IV.0.1. Solution of Optimal Control Problem

First, we describe how the above formalism translates into our specific problem. As we already mentioned, the equivalent problem in the fixed σ-interval $[0,1]$ is to minimize

$$F = -X(t_f) \qquad (25)$$

The necessary conditions for the minimum in our problem are

1. $\dot{\mathbf{y}} = \mathbf{g}$, $y(0) = y_0$

   The state equations are

   $$\frac{dX}{d\sigma} = t_f(\mu - \phi D)X; \qquad X(0) = X_0$$
   $$\frac{ds_g}{d\sigma} = t_f(-u_g X - (s_g - c_g)D); s_g(0) = s_{g0}$$
   $$\frac{ds_l}{d\sigma} = t_f(-u_l X - s_l D); \qquad s_l(0) = 0$$

   The Hamiltonian is then given by

   $$H = t_f[-X + \lambda_x(\mu - \phi D)X + \cdots$$
   $$+ \lambda_g(-u_g X - (s_g - c_g)D) + \lambda_l(-u_l X - s_l D)]$$

   The Lagrangian is

   $$L(\sigma) = H + [f(y_i, u_i)]^T \mathbf{W} f(y_i, u_i);$$
   $$i = 0, 1, \ldots, N$$

   where

   $$f(\mathbf{y}, u) = -D + D_0 \leq 0$$

   and $\mathbf{W} = \alpha^r \mathbf{1}$. Then

   $$L(\sigma) = t_f[-X + \lambda_x(\mu - \phi D)X + \cdots$$
   $$+ \lambda_g(-u_g X - (s_g - c_g)D) + \cdots$$
   $$+ \lambda_l(-u_l X - s_l D)] + \alpha^r(-D + D_0)^2$$

2. $\frac{d\lambda}{d\sigma} = -L_y, \lambda(1) = 0$

3. $L_u = 0$

In the Appendix, there is a detailed description of these expressions and their numerical implementations.

The results of our numerical simulations are presented in figure 5, where we compare the output of our optimization with the ones obtained using fixed dilution rates (see also figure 3).

We can see that using an appropriate dilution rate protocol it is possible to achieve a higher cell density than that obtained using a constant dilution rate value, and also to reach this value more quickly. Note that through this protocol also the concentration of lactate, a toxic by-product of metabolism, is significantly reduced in the culture. We want to point out that predicting the optimal protocol to follow based on experience is not trivial and requires a high expenditure of time and resources.

As can be seen in figure 5, the graph indicating the optimal dilution protocol is very noisy. This noise is an effect of finite size due to the finite number N of partitions used to describe the interval $\sigma = [0, 1]$. This can be solved by increasing N, but
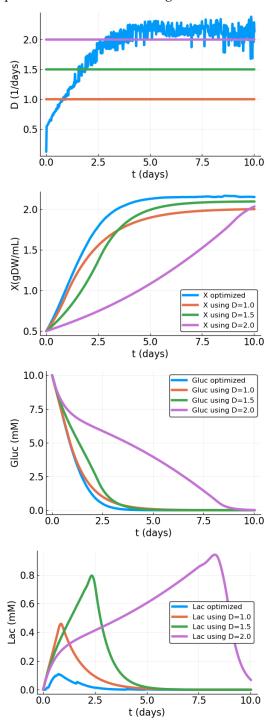
the computational cost would be higher.

Figura 5. Dilution rate, cell density, and concentrations of glucose and lactate in the culture in time. The blue line shows the results of the Optimal Control Problem, and the others lines are shown above in figure 4. In all graphs, we use as initial conditions $X_0 = 0,5$ gDW/mL, $c_{g_0} = 10$ mM; and $\phi = 0,5$. In the optimization, we use the follow parameters: $N = 500$, $\epsilon = 0,1$, and 30×10^6 iterations.

Therefore, to provide experimenters with a $D(t)$ curve without this noise, we performed a simple interpolation to describe the protocol for D. We use a linear fit for the first part of the protocol, and a constant for the second. In practice, $D(t) = 0,60$ days^{-1} $+ 0,02$ days^{-2}t during the first 3 days of the process and $D = 2,15$ days^{-1} for the rest of the time (red dashed line in figure 6). To test the efficiency of the protocol, we compareed

the results of using our interpolated function with two similar protocols for the dilution rate D. The results are summarized in figure 6. We show that if the dilution increases too slowly, the steady state will be reached later in time. On the contrary, if it is increased too quickly, it will take much longer to reach the steady state, but in the process it will produce much more toxic by-products.
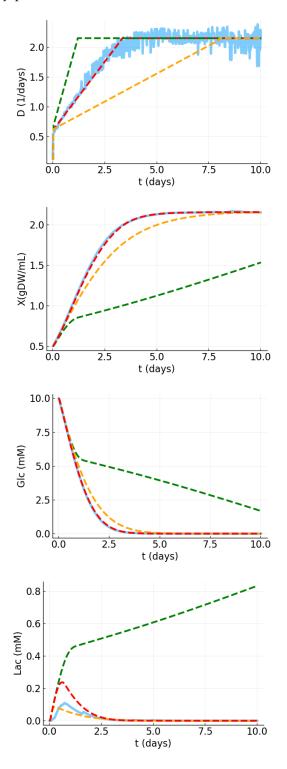

Figura 6. Comparison of cell density and glucose and lactate concentrations using two dilution rates somewhat different from the optimal obtained (blue line).

V. CONCLUSIONS

In this work, we have modeled a continuous cell culture using a set of non-linear equations. Our model describes the dynamics of the cell density as a combination of intracellular properties (cell growth) and external parameters (perfusion coefficient and dilution) in the chemostat. It also describes the dynamics of metabolite concentrations in the culture as a balance between cell consumption/secretion and the dilution rate.

As a proof of concept, we study a simple metabolic network and formulated the cell growth rate through the FBA algorithm. In this way, we avoid the need to use kinetic parameters to describe metabolic reactions. These parameters are usually unknown, or variable over time.

For the studied model, exploiting Optimal Control Theory techniques, including the dynamic penalty method, we were able to predict an optimal protocol for the dilution rate as a function of time. The use of this protocol maximizes the cell density at the steady state, minimizes the time in which this steady state is reached, and also minimizes the secretion of lactate in the culture.

Although these results can no longer be translated already into experimental systems, they support the feasibility of the approach for more complex and realistic problems. We are sure that the extension of this technique to the study of Genome Scale Metabolic Networks could be of great relevance in the biotechnological industry. Work is being done in this direction.

VI. APPENDICES

VI.1. Computational Algorithm

The objective in this problem is to find the control u that minimizes

$$\int_0^{t_f} F[\mathbf{y}(t), u(t)]dt$$

subject

$$\dot{\mathbf{y}} = \mathbf{g}[\mathbf{y}(t), u(t)], \quad \mathbf{y}(0) = \mathbf{y_0}$$

and l algebraic equality constraints

$$f(\mathbf{y}, u) = 0$$

The equivalent problem is to find the minimum of the augmented functional

$$J = \int_0^{t_f} [F + \lambda^T(-\dot{\mathbf{y}} + \mathbf{g}) + \mu_\lambda^T f]dt = \int_0^{t_f} (L - \lambda^T \dot{\mathbf{y}})dt \tag{26}$$

where, in general, μ_λ are time dependent multipliers associated with the algebraic constraints f.

The necessary conditions for the minimum of J are

$$\dot{\mathbf{y}} = \mathbf{g} \qquad \mathbf{y}(0) = \mathbf{y_0} \qquad \dot{\lambda} = -L_{\mathbf{y}} \qquad \lambda(t_f) = 0$$
$$f \le 0 \qquad L_u = \mathbf{0} \qquad L(t_f) = 0$$

At any time, the multipliers corresponding to the inequalities are prescribed as

$$\mu_\lambda = \mathbf{W}h \tag{27}$$

where $\mathbf{W}$ is an $l \times l$ diagonal weighting matrix with all positive diagonal elements and h is

$$h = \begin{cases} 0 & \text{if } f(t) \le 0 \\ 1 & \text{if } f(t) > 0 \end{cases}$$

To determine u and t_f at the minimum [41],

1. Initialize the outer counter $r = 0$. Choose a real number $\alpha > 1$.

2. Set the $l \times l$ diagonal weighting matrix $\mathbf{W} = \alpha^r \mathbf{1}$, where $\mathbf{1}$ is the $l \times l$ identity matrix.

3. Set the iteration counter $k = 0$. Assume $t_k = 1{,}0$ and obtain the fixed σ-grid of $(N+1)$ equi-spaced grid points

$$\sigma_0 = 0, \quad \sigma_1, \quad \sigma_2, \quad \dots, \quad \sigma_N = 1 \tag{28}$$

At each grid point, assume control function values as follows:

$$u_i^k \equiv u^k(\sigma_i); \quad i = 0, 1, \dots, N \tag{29}$$

4. Integrate state equations forward from $\sigma = 0$ to 1 using the initial conditions and the control function values u_i^k, $i = 0, 1, \dots, N$. Save the values of state variables at the grid points.

5. Evaluate the objective functional using the controls u_i^k and the state variables $\mathbf{y_i^k}$. Using composite Simpson's 1/3 Rule, the objective functional value is given by

$$J = \frac{1}{3N} \left(A_0 + 4 \sum_{1,3,5,\dots}^{N} A_i + 2 \sum_{2,4,6,\dots}^{N} A_i + A_N \right) \tag{30}$$

where $A_i \equiv t_f F(y_i, u_i) + [f(y_i, u_i)]^T \mathbf{W} f(y_i, u_i)$; $i = 0, 1, \dots, N$.

6. Check the improvement in F for $k > 0$. Given a tolerable error $\epsilon_1 > 0$, if

$$| F^k - F^{k-1} | < \epsilon_1 \tag{31}$$

7. Integrate costate equations backward from $\sigma = 1$ to 0 using the final conditions, the controls u_i^k, and the state variables $\mathbf{y_i^k}$. Save the values of costate variables at the grid points.

8. Evaluate the gradient of J k by calculating the partial derivatives

$$H_{u,i}^k \equiv H_u^k(\sigma_i) = H_u(\mathbf{y}_i^k, u_i^k, \lambda_i^k); \quad i = 0, 1, \ldots, N \tag{32}$$

where

$$H_i^k \equiv t_f^k[F(y_i^k, u_i^k, \lambda_i^k) + \lambda_i^T g(y_i^k, u_i^k, \lambda_i^k)]; \quad i = 0, 1, 2, \ldots, N \tag{33}$$

and

$$J_{t_f}^k = \int_0^1 \frac{H^k}{t_f^k} d\sigma = \frac{1}{3Nt_f^k}\left(H_0^k + 4\sum_{1,3,5,\ldots}^N H_i^k + 2\sum_{2,4,6,\ldots}^N H_i^k + H_N^k \right) \tag{34}$$

Check the magnitude of the gradient. Given a small positive real number ϵ_2, if the norm of the gradient

$$\sqrt{\sum_{i=0}^N [H_u(y_i^k, u_i^k, \lambda_i^k)]^2 + [J_{t_f}^k]^2} < \epsilon_2 \tag{35}$$

then take these results.

9. Improve control functions and the final time by calculating

$$u_i^{k+1} = u_i^k - \epsilon H_{u,i}^k, \quad i = 0, 1, \ldots, N \tag{36}$$

$$t_f^{k+1} = t_f^k - \epsilon J_{t_f}^k \tag{37}$$

where $\epsilon \equiv \epsilon_0 / \parallel \Delta J \parallel$, ϵ_0 is some positive real number and

$$\nabla J \equiv [L_u(\sigma_0) \quad L_u(\sigma_1) \quad \ldots \quad L_u(\sigma_N) \quad J_{t_f}]^T \tag{38}$$

where

$$L(\sigma_i) = t_f[F(y_i, u_i) + \lambda_i^T g(y_i, u_i)] + [f(y_i, u_i)]^T \mathbf{W} f(y_i, u_i); \quad i = 0, 1, \ldots, N \tag{39}$$

10. When the gradient algorithm converges in, say, k iterations, save the state, the control, and the final time,

$$\mathbf{y}_i^k, \quad u_i^k; \quad i = 0, 1, \ldots, N; \quad \text{and} \quad t_f^k$$

11. Increment the counter r by one and assign

$$\mathbf{y}_i^k \to \mathbf{y}_i^r, \quad u_i^k \to u_i^r; \quad \mathbf{h}_i^k \to \mathbf{h}_i^r; \quad i = 0, 1, \ldots, N; \quad \text{and} \quad t_f^k \to t_f^r$$

12. Given a positive real number ϵ_3 close to zero, if the error

$$E = \sum_{i=0}^N [\mathbf{h}_i^r]^T f(y_i^r, u_i^r) > \epsilon_3 \tag{40}$$

then some constraints are violated. Therefore, repeat the algorithm.

Otherwise, the constraints $f \leq 0$ are satisfied, and the values

$$\mathbf{y}_i^r, \quad u_i^r; \quad i = 0, 1, \ldots, N; \quad \text{and} \quad t_f^r$$

correspond to the minimum.

VI.2. Solution of Optimal Control Problem

VI.2.1. Toy model

Dynamic Problem

The Lagrangian is given by

$$L(\sigma) = t_f[-X + \lambda_x(\mu - \phi D)X + \lambda_g(-u_g X - (s_g - c_g)D) + \lambda_l(-u_l X - s_l D)] + \alpha^r[(-D + D_0)^2 + \phi^2 + (\phi - 1)^2] \tag{41}$$

The following conditions are necessary:

- $\frac{d\lambda}{d\sigma} = -L_y, \lambda(1) = 0$

  The costate equations are

  $$\frac{d\lambda_x}{d\sigma} = -L_X = t_f[1 - \lambda_x(\mu - \phi D + \frac{\partial \mu}{\partial X}X) + \lambda_g(u_g + \frac{\partial u_g}{\partial X}X) + \lambda_l(u_l + \frac{\partial u_l}{\partial X}X)], \lambda_x(1) = 0$$

  $$\frac{d\lambda_g}{d\sigma} = -L_{s_g} = t_f \lambda_g D, \lambda_g(1) = 0$$

  $$\frac{d\lambda_l}{d\sigma} = -L_{s_l} = t_f(-\lambda_x X \frac{\partial \mu}{\partial s_l} + \lambda_l D), \lambda_l(1) = 0$$

  where

  $$\frac{\partial \mu}{\partial X} = \begin{cases} -\frac{1}{y}\frac{c_g D}{X^2}(N_F + N_R) & \text{if } c_g D/X < r_m \\ -\frac{1}{y}\frac{c_g D}{X^2}N_F & \text{if } r_m \leq c_g D/X \leq V \\ 0 & \text{if } V < c_g D/X \end{cases}$$

  $$\frac{\partial \mu}{\partial s_l} = -\tau$$

  and

  $$\frac{\partial u_g}{\partial X} = \begin{cases} -\frac{c_g D}{X^2} & \text{if } V \geq c_g D/X \\ 0 & \text{if } V < c_g D/X \end{cases}$$

$$\frac{\partial u_l}{\partial X} = \begin{cases} 0 & \text{if } c_g D/X < r_m \\ \frac{c_g D}{X^2} & \text{if } r_m \leq c_g D/X \leq V \\ 0 & \text{if } V < c_g D/X \end{cases}$$

- $L_u = 0$

$$\frac{\partial L}{\partial D} = t_f \left\{ X \left[\lambda_x \left(\frac{\partial \mu}{\partial D} - \phi \right) - \lambda_g \frac{\partial u_g}{\partial D} - \lambda_l \frac{\partial u_l}{\partial D} \right] \right\} \tag{42}$$
$$- t_f [\lambda_g (s_g - c_g) - \lambda_l s_g - 2\alpha^r(-D + D_0)] = 0$$

$$\frac{\partial L}{\partial \phi} = t_f [-DX\lambda_x] + \alpha^r(4\phi - 2) = 0 \tag{43}$$

where

$$\frac{\partial \mu}{\partial D} = \begin{cases} \frac{1}{y}\frac{c_g}{X}(N_F + N_R) & \text{if } c_g D/X < r_m \\ \frac{1}{y}\frac{c_g}{X}N_F & \text{if } r_m \leq c_g D/X \leq V \\ 0 & \text{if } V < c_g D/X \end{cases}$$

and

$$\frac{\partial u_g}{\partial D} = \begin{cases} \frac{c_g}{X} & \text{if } V \geq c_g D/X \\ 0 & \text{if } V < c_g D/X \end{cases}$$

$$\frac{\partial u_l}{\partial D} = \begin{cases} 0 & \text{if } c_g D/X < r_m \\ -c_g/X & \text{if } r_m \leq c_g D/X \leq V \\ 0 & \text{if } V < c_g D/X \end{cases}$$

REFERENCES

[1] T. S. Le, M. McCann, S. M. Azarin and W.-S. Hu, Chem. Eng. Prog. **112**, 34 (2016).

[2] M. S. Croughan, K. B. Konstantinov and C. Cooney, Biotechnol. Bioeng. **112**, 648 (2015).

[3] A. Kumar, I. A. Udugama, C. L. Gargalo and K. V. Gernaey, Processes **8**, 1641 (2020).

[4] A. P. B. Benyahia, L. Pereira and F. V. Santos da Silva, New and Future Developments in Microbial Biotechnology and Bioengineering, **69**, 89 (2020).

[5] M. Guevara-Martinez, Strain-and bioprocess-design strategies to increase production of (r)-3-hydroxybutyrate by escherichia coli, Ph.D. thesis, KTH Royal Institute of Technology (2019).

[6] B. A. Pérez-Fernández, J. Fernandez-de Cossio-Diaz, T. Boggiano, K. León and R. Mulet, Biotechnol. Bioeng. **118**, 1884 (2021).

[7] P. Biechele, C. Busse, D. Solle, T. Scheper and K. Reardon, Eng. Life Sci. **15**, 469 (2015).

[8] D. Martens, C. De Gooijer, C. Van der Velden-De Groot, E. Beuvery and J. Tramper, Biotechnol Bioeng **41**, 429 (1993).

[9] M. K. F. Wolf, A. Closet, M. Bzowska, J.-M. Bielser, J. Souquet, H. Broly and M. Morbidelli, Biotechnol. J. **14**, 1700722 (2019).

[10] G. Alwan, Int. J. Chem. Process Eng. Res. **5** (2012).

[11] J. Monod, Annu. Rev. Microbiol. **3**, 371 (1949).

[12] J. Alvarez-Ramirez, M. Meraz and E. J. Vernon-Carter, Biochem. Eng. J. **150**, 107305 (2019).

[13] B. Chezeau and C. Vial, Biohydrogen, Biomass, Biofuels, Biochemicals **445**, 483 (Elsevier, 2019), 2nd edition.

[14] D. Orhon, E. U. Cokgor, G. Insel, O. Karahan and T. Katipoglu, Bioresour. Technol. 100, 5678 (2009).

[15] I. Pala-Ozkok, A. Rehman, N. Yagci, E. Ubay-Cokgor, D. Jonas and D. Orhon, Bioresour. Technol. **126**, 274 (2012).

[16] R. U. Ibarra, J. S. Edwards and B. O. Palsson, Nature **420**, 186 (2002).

[17] J. Fernandez-de-Cossío-Díaz, K. Leon and R. Mulet, PLoS Comput. Biol. **13**, e1005835 (2017).

[18] A. Vazquez, J. Liu, Y. Zhou and Z. N. Oltvai, BMC. Syst. Biol. **4**, 58 (2010).

[19] J. Fernandez-de-Cossío-Díaz and A. Vazquez, Sci. Rep. **7**, 13488 (2017).

[20] J. Fernandez-de-Cossío-Díaz and R. Mulet **15**, e1006823 (2019).

[21] J. Fernandez-de-Cossío-Díaz, A. D. Martino, R. Mulet, Sci. Rep. **7**, 3103 (2017).

[22] D. Kilburn, M. Lilly and F. Webb, J. Cell. Sci. **4**, 645 (1969).

[23] A. M. Feist and B. O. Palsson, Curr. Opin. Microbiol. **13**, 344 (2010).

[24] A. M. Feist and B. O. Palsson, Genome Biol. **17**, 110 (2016).

[25] M. Al-Rubeai, D. Mills and A. Emery, Cytotechnology **4**, 13 (1990).

[26] N. Simpson, R. Singh, A. Perani, C. Goldenzon and M. Al-Rubeai, Biotechnol. Bioeng. **59**, 90 (1998).

[27] R. Singh, M. Al.-Rubeai, C. Gregory and A. Emery, Biotechnol. Bioeng. **44**, 720 (1994).

[28] S. Mercille and B. Massie, Biotechnol. Bioeng. **44**, 1140 (1994).

[29] B. Krampe and M. Al-Rubeai, Cytotechnology **62**, 175 (2010).

[30] A. Perani, R. Singh, R. Chauhan and M. Al-Rubeai, Cell. Culture Engineering VI 177, **188** (Springer, 1998).

[31] M. Al-Rubeai, R. Singh, M. Goldman and A. Emery, Biotechnol. Bioeng. **45**, 463 (1995).

[32] R. Singh and M. Al-Rubeai, Apoptosis **167**, 184 (1998).

[33] J. S. Ryu and G. M. Lee, Biotechnol. Bioeng. **62**, 120 (1999).

[34] Y. Chisti, Trends Biotechnol. **18**, 420 (2000).

[35] B. C. Batt and D. S. Kompala, Biotechnol. Bioeng. **34**, 515 (1989).

[36] S. Dhir, K. J. Morrow, R. R. Rhinehart and T. Wiesner, Biotechnol. Bioeng. **67**, 197 (2000).

[37] B. S. Cummings and R. G. Schnellmann, Curr. Protoc. Pharmacol. **25**, 12 (2004).

[38] B. Alberts and A. Johnson, Molecular Biology of the Cell (Garland Science, Taylor and Francis Group, New York, NY, 2015), 6th edition.

[39] M. A. Bree, P. Dhurjati, R. F. Geoghegan and B. Robnett, Biotechnol. Bioeng. **32**, 1067 (1988).

[40] S. Rodríguez-Enríquez, A. Marín-Hernández, J. C. Gallardo-Pérez and R. Moreno-Sánchez, J. Cell. Physiol. **221**, 552 (2009).

[41] S. R. Upreti, Optimal control for chemical engineers (Taylor & Francis, 2013).

[42] H. Goldstein, C. Poole and J. Safko, Classical mechanics (American Association of Physics Teachers, New York, South Carolina, 2002), 3rd edition.