HAL open science

Modelling Interactions of a Microbial Community through an Optimal Tracking Control Approach

Pablo Ugalde-Salas, Jérôme Harmand, Elie Desmond-Le Quéméner

To cite this version:

Pablo Ugalde-Salas, Jérôme Harmand, Elie Desmond-Le Quéméner. Modelling Interactions of a Microbial Community through an Optimal Tracking Control Approach. 2020 European Control Conference (ECC), May 2020, Sant Petersburg, Russia. pp.587-592, 10.23919/ECC51009.2020.9143828 . hal-02954751

HAL Id: hal-02954751

https://hal.inrae.fr/hal-02954751v1

Submitted on 1 Oct 2020

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

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

Distributed under a Creative Commons Attribution 4.0 International License

Modelling Interactions of a Microbial Community through an Optimal Tracking Control Approach.

Pablo Ugalde-Salas[1*], Jérôme Harmand[1], and Elie Desmond-Le Quéméner[1]

***Abstract*— Genetic sequencing measurements of the existing species in a reactor are incorporated into a classic chemostat model through a non linear optimal tracking approach. The framework is applied to data coming from a nitrification process, resulting in a dynamical system with a high dimensional state. The observed ecological dynamics from the bacterial species are able to explain the substrates dynamics.**

I. INTRODUCTION

In the last decades genetic sequencing techniques have become an affordable measurement in bioprocesses. However using this measurements for the prediction and control of the bioreactors performance (in terms of proportions of chemicals consumed and produced) lacks a modelling framework [1].

It has been theorised that interactions between microorganisms could help explain the difference in performance of reactors operating under very similar environmental conditions [2]. Some types of interactions have already been accounted in chemostat theory, such as inhibition in the presence of other species, or density dependence. Nevertheless this represent just a fraction of the possible ecological interactions[3] that could take place in a microbial community.

The central idea behind the method presented in this work is that if one can accurately follow the dynamic of the species present in an ecosystem for a certain time horizon then one can recover the underlying interactions that might be driving the system.

II. MODEL DEFINITION AND PROBLEM STATEMENT

The following conventions are used: (1) Let m be an integer then $[m] := \{1,\dots,n\}$, (2) For integers m_1 and m_2 and $a \in \mathbb{R}$, $a_{m_1\times m_2}$ represents a matrix of m_1 rows and m_2 columns with a in every entry, (3) let m be an integer then I_m is the identity matrix of size m, (4) given a vector $v = (v_1,\dots,v_n) \in \mathbb{R}^n$ the function $\mathbf{diag}(v)$ returns a square matrix M with entries $M_{ii} = v_i$, $\forall i \in [n]$, and $M_{ij} = 0$, $\forall i,j \in [n]$, $i \neq j$, and (5) Let $a_1, a_2 \in \mathbb{R}$ then $a_1 \vee a_2 := \max\{a_1,a_2\}$ and $a_1 \wedge a_2 := \min\{a_1,a_2\}$.

A. Stoichiometric Equations

Consider the situation where n different Operative Taxonomic Units (OTU) are present in a system. A group of OTU ($G_1 \subset [n]$) consumes a substrate s_1 for growth and as a result s_2 is produced, while another group of OTU ($G_2 \subset [n]$) consumes s_2 for growth and produces s_3. G_1 and G_2 are called functional groups. This situation is often referred to as a cascade bioreaction. The number of organisms in G_1 and G_2 will be denoted n_1 and n_2, respectively. It will be assumed that sets G_1 and G_2 satisfy: $G_1 \cap G_2 = \emptyset$ and $G_1 \cup G_2 = [n]$.

The situation is schematically described in bioreactions (R G1) and (R G2). The terms y_i are known as yields, they represent the number of moles of biomass produced per mole of substrate consumed. The term M_x represents a molecule of biomass and several expressions can be found in the literature (e.g. $CH_{1.613}O_{0.557}N_{0.158}$ [4]). $s = (s_1,s_2,s_3)$ and $x = (x_1,\dots,x_n)$ denote the vectors of concentrations of substrates and of OTU present in the system, respectively. Each reaction has its own process rate (also known as growth function or kinetics) $\mu_i(s,x)$. Note that it depends on x to show the influence of other OTU on the growth function. A classic treatment of general reactions and kinetics in biotechnology can be found in [5].

$$s_1 \xrightarrow{\mu_i(s,x)} s_2 + y_i M_x \quad \forall i \in G_1 \tag{R G1}$$

$$s_2 \xrightarrow{\mu_i(s,x)} s_3 + y_i M_x \quad \forall i \in G_2 \tag{R G2}$$

B. Mass Balance Equations

The system where reactions (R G1) and (R G2) take place is a container of fixed volume V containing a liquid medium with the same input and output flow $Q = Q(t)$, which can vary in time. The dilution rate is then defined as $D = D(t) = \frac{Q}{V}$. The input flow contains a time dependent concentration $s_{in} = s_{in}(t)$ of the same chemical composition than s_1, for readability the time dependence notation is dropped. For chemostat theory detailed modelling and mathematical analysis one can refer to [6]. The dynamics of the described situation are seen in (1).

$$\begin{aligned}
\dot{x_i} &= (\mu_i(s,x) - D)x_i \quad \forall i \in G_1 \\
\dot{x_i} &= (\mu_i(s,x) - D)x_i \quad \forall i \in G_2 \\
\dot{s_1} &= (s_{in} - s_1)D - \sum_{i\in G_1} \frac{1}{y_i}\mu_i(s_1,x)x_i \\
\dot{s_2} &= -s_2 D + \sum_{i\in G_1} \frac{1}{y_i}\mu_i(s,x)x_i - \sum_{i\in G_2} \frac{1}{y_i}\mu_i(s,x)x_i \\
\dot{s_3} &= -s_3 D + \sum_{i\in G_2} \frac{1}{y_i}\mu_i(s,x)x_i
\end{aligned} \tag{1}$$

[1] LBE, INRAE, Univ. Montpellier, 11100, Narbonne, France.
[*] Corresponding author : pablo.ugalde-salas@inrae.fr

C. Kinetics and Interactions

Given a vector $(v_1,\dots,v_n)^\top$ the interaction function I is noted as:

$$\begin{aligned} I:\mathbb{R}^n \to \quad & [0,1]^n \\ v \to \quad & \begin{pmatrix} I_1(v) \\ I_2(v) \\ \vdots \\ I_n(v) \end{pmatrix} \end{aligned} \tag{2}$$

The growth function of each OTU will be assumed to be a product of substrate dependent function and the interaction function. Let $f_i(s)$ be a bounded, positive, and continuous function of s (e.g. Monod, Haldane). The growth equation of OTU i becomes:

$$\mu_i(s,x) = f_i(s)I_i(x) \tag{3}$$

The notation $f(s) := (f_1(s),\dots,f_n(s))^\top$ will be used.

Since the growth of a single strain in chemostat experiments is limited by the substrate concentration , when no interactions are present one should recover expression $f_i(s)$. Therefore if there are no interactions then $I_i(x)=1$. From this reasoning, note that $\lim_{x\to 0} I_i(x) = 1$ since if there is minimal presence of OTU, interactions can not exist. Furthermore it is assumed that $I_i(\cdot)$ is a continuously differentiable function on x. For making explicit all of the former:

The interaction function I previously defined satisfies:

1) $I\left((0,\dots,0)^\top\right) = (1,\dots,1)^\top$
2) Let Ω be an open subset of $\mathbb{R}^n$ then $I \in C^1(\Omega)$.

By forcing the positivity of the interaction function then the growth functions are still positive and the system remains well posed [6], meaning that if system (1) is provided with non-negative initial conditions, the trajectories remain non-negative.

The problem is then to reconstruct the interaction function from the observations of the microbial community composition in time.

D. Optimal Control Setting for Unravelling the Interaction Function

Suppose that the functions $f_i(s)$, and the yields y_i are well-known. By using experimental measurements of x, represented by $z(t)$, the objective is to reconstruct function $I(x)$. For doing so, the terms $I_i(x)$ are replaced by a control $u_i(t)$ $(u_i(t) = I_i(x(t)))$ and a state feedback control law is obtained by solving a nonlinear optimal tracking problem.

Consider the observable system (4).

$$\begin{aligned} \dot{x_i} &= (f_i(s)u_i(t) - D)x_i \quad \forall i \in G_1 \\ \dot{x_i} &= (f_i(s)u_i(t) - D)x_i \quad \forall i \in G_2 \\ \dot{s_1} &= (s_{in} - s_1)D - \sum_{i\in G_1} \frac{1}{y_i} f_i(s)u_i(t)x_i \\ \dot{s_2} &= -s_2 D + \sum_{i\in G_1} \frac{1}{y_i} f_i(s)u_i(t) - \sum_{i\in G_2} \frac{1}{y_i} f_i(s)u_i(t)x_i \\ \dot{s_3} &= -s_3 D + \sum_{i\in G_2} \frac{1}{y_i} f_i(s)u_i(t)x_i \\ y(t) &= x(t) \end{aligned} \tag{4}$$

Let Q, and R be positive definite matrices, and let $\|\cdot\|_Q$ and $\|\cdot\|_R$ be the weighted norms by the corresponding matrices. The optimal tracking problem is defined as:

$$\begin{aligned} \min \quad & \int_0^T \|y-z\|_Q dt + \int_0^T \|(u - 1_{n\times 1})\|_R dt \\ s.t. \quad & (x, s_1, s_2, s_3) \text{ solution of (4)} \\ & u_i(t) \in [0,1] \end{aligned} \tag{5}$$

The control $u(t)$ is intended to drive the system to be near a desired output $z(t)$, which in this context are the measurements of the concentrations of OTU. The term $\|(u - 1_{n\times 1})\|_R$, was added for two reasons:

First because the interest is testing the idea that interactions could be driving the system. Therefore adding a penalization in the objective function for each control to remain near 1 can be seen as an attempt to explain data without any interaction.

Second, to force a regularized control. Otherwise note that u is linear in the state equations (4), therefore if the integral cost does not have a non-linear expression of u the optimal control will be of a bang-bang type with possibly singular arcs. Since the objective is to find a differentiable expression of $I(x)$ the addition of the regularization term is deemed necessary.

The problem of approximating the solution of the system to a desired reference (z in this case) is called the optimal tracking problem. For solving such a problem the approach developed by Cimen *et al.* [7] was adapted to our problem, using [8]. The method proposed involves the resolution of Approximating Sequences of Ricatti Equations.

The method consists on iteratively calculating trajectories of system (4) with a certain control and feeding a non-autonomous Ricatti differential equation with the resulting trajectory. Then a control law that uses the Ricatti equation is proposed and a new trajectory is calculated. The iteration is stopped when a convergence in the output or in the control is observed.

First the tracking problem does not consider a constrained control, so one should consider a reformulation as proposed by the same author in other works [8]. Nevertheless, the methods in [7] were directly used, and an explicit constraint in the synthesis of the control was added for the control to remain within the desired bounds (see equation (21)).

Even though this is probably suboptimal it worked for our purposes.

Second in their problem they used an approximation of their original dynamics for calculating the trajectories and subsequently proved that such linearisation converged to the original dynamics. In the case here presented the linearisation of the dynamics was not necessary.

E. Tracking Problem Reformulation

For applying the methods developed in [7]. Define the system state $X=(x,s)$, and the change of variables $v_i=u_i-1$ with $v=(v_1,\dots,v_n)$. System (4) may be rewritten then as:

$$\begin{aligned}
\dot{x_i} &= (f_i(s)(1+v_i(t))-D)x_i \quad \forall i\in G_1\\
\dot{x_i} &= (f_i(s)(1+v_i(t))-D)x_i \quad \forall i\in G_2\\
\dot{s_1} &= s_1\left(\frac{s_{in}}{s_1}-1\right)D-\sum_{i\in G_1}\frac{1}{y_i}f_i(s)(1+v_i(t))x_i\\
\dot{s_2} &= -s_2D+\sum_{i\in G_1}\frac{1}{y_i}f_i(s)(1+v_i(t))\\
&\quad -\sum_{i\in G_2}\frac{1}{y_i}f_i(s)(1+v_i(t))x_i\\
\dot{s_3} &= -s_3D+\sum_{i\in G_2}\frac{1}{y_i}f_i(s)(1+v_i(t))x_i\\
y(t) &= x(t)
\end{aligned} \tag{6}$$

Define

$$k_{G_1}=\left(\frac{1}{y_1},\dots,\frac{1}{y_{nA}}\right)^\top\in\mathbb{R}^{n_1} \tag{7}$$

$$k_{G_2}=\left(\frac{1}{y_{n_1+1}},\dots,\frac{1}{y_n}\right)^\top\in\mathbb{R}^{n_2} \tag{8}$$

$$A(X)=\begin{bmatrix}A_{11}(X) & A_{12}(X)\\ A_{21}(X) & A_{22}(X)\end{bmatrix} \tag{9}$$

$$B(X)=\begin{bmatrix}B_1(X)\\ B_2(X)\end{bmatrix} \tag{10}$$

$$A_{11}(X)=\mathbf{diag}(f(s)-D_{n\times 1}) \tag{11}$$

$$A_{21}(X)=\begin{bmatrix} f(s)^\top\begin{bmatrix}-\mathbf{diag}(k_{G_1}) & 0_{n_1\times n_2}\end{bmatrix}^\top\\ f(s)^\top\begin{bmatrix}\mathbf{diag}(k_{G_1}) & -\mathbf{diag}(k_{G_2})\end{bmatrix}^\top\\ f(s)^\top\begin{bmatrix}0_{n_2\times n_1} & \mathbf{diag}(k_{G_2})\end{bmatrix}^\top\end{bmatrix} \tag{12}$$

$$A_{12}(X)=0_{n\times 3} \tag{13}$$

$$A_{22}(X)=\begin{bmatrix}\left(\frac{s_{in}}{s_1}-1\right)D & 0 & 0\\ 0 & -D & 0\\ 0 & 0 & -D\end{bmatrix} \tag{14}$$

$$B_1(X)=\mathbf{diag}(f(s))\mathbf{diag}(x) \tag{15}$$

$$B_2(X)=\begin{bmatrix} f(s)^\top\mathbf{diag}(x)\begin{bmatrix}-\mathbf{diag}(k_{G_1}) & 0_{n_1\times n_2}\end{bmatrix}^\top\\ f(s)^\top\mathbf{diag}(x)\begin{bmatrix}\mathbf{diag}(k_{G_1}) & -\mathbf{diag}(k_{G_2})\end{bmatrix}^\top\\ f(s)^\top\mathbf{diag}(x)\begin{bmatrix}0_{n_2\times n_1} & \mathbf{diag}(k_{G_2})\end{bmatrix}^\top\end{bmatrix} \tag{16}$$

$$C(X)=\begin{bmatrix}I_n & 0_{n\times 3}\end{bmatrix} \tag{17}$$

Then system (4) becomes:

$$\begin{aligned}\dot{X}&=A(X)X+B(X)v\\ y&=C(X)X\end{aligned} \tag{18}$$

Let $z(t)\in\mathbb{R}^n$ be the vector containing the experimental measurements of the OTU concentrations in time. The cost functional is now rewritten as (19).

$$\begin{aligned}J(v) &= \int_{t_0}^{t_f}(z(t)-X(t))^\top Q(z(t)-X(t))\,dt\\ &\quad+\int_{t_0}^{t_f}v(t)^\top Rv(t)dt\end{aligned} \tag{19}$$

Define the dynamic sequences for $i\in\mathbb{N}$, $X^{[i]}$ as :

$$\begin{aligned}\dot{X}^{[i]}&=A(X^{[i]})X^{[i]}+B(X^{[i]})v^{[i]}\\ y^{[i]}&=X^{[i]}\\ X^{[i]}(t_0)&=X_0\end{aligned} \tag{20}$$

And for $i=0$ let $X^{[0]}(t)=X_0$. The control law is given by

$$\begin{gathered}v^{[i]}(t)_j=\\ 0\wedge\left(-1\vee\left(-R^{-1}B^\top\left(X^{[i-1]}(t)\right)\left(P^{[i]}(t)X^{[i]}(t)-s_f^{[i]}(t)\right)\right)_j\right)\\ \forall j\in[n]\end{gathered} \tag{21}$$

Where $P^{[i]}(t)\in\mathscr{M}_{n+3\times n+3}(\mathbb{R})$ and $s_f^{[i]}(t)\in\mathbb{R}^{n+3}$ are the solution to the systems of differential equations (22) and (23), respectively.

$$\begin{aligned}\dot{P}^{[i]}&=-C^T\left(X^{[i-1]}(t)\right)QC\left(X^{[i-1]}(t)\right)\\ &\quad-P^{[i]}A\left(X^{[i-1]}(t)\right)-A^\top\left(X^{[i-1]}(t)\right)P^{[i]}\\ &\quad+P^{[i]}B\left(X^{[i-1]}(t)\right)R^{-1}B^\top\left(X^{[i-1]}(t)\right)P^{[i]}\\ P^{[i]}(t_f)&=0\end{aligned} \tag{22}$$

$$\begin{aligned}\dot{s_f^{[i]}}&=-C^\top\left(X^{[i-1]}(t)\right)Qz(t)-\Big[A\left(X^{[i-1]}(t)\right)\\ &\quad-B\left(X^{[i-1]}(t)\right)R^{-1}B^\top\left(X^{[i-1]}(t)\right)P^{[i]}(t)\Big]^\top s_f^{[i]}\\ s_f^{[i]}(t_f)&=0\end{aligned} \tag{23}$$

Note that system (22) and (23) consists of $(n+3)^2+(n+3)$ differential equations. In the following the inspection of system (22) and (23) allows us to reduce the number of differential equations to $2n$. Replacing matrices (9), (10), and (17) in (22) and (23) one obtains:

$$\dot{P}^{[i]}(t) = -\begin{bmatrix} Q & 0_{n\times 3} \\ 0_{3\times n} & 0_{3\times 3} \end{bmatrix} - P^{[i]} A\left(X^{[i-1]}(t)\right) \tag{24}$$
$$-A^{\top}\left(X^{[i-1]}(t)\right) P^{[i]}$$
$$+P^{[i]} B\left(X^{[i-1]}(t)\right) R^{-1} B^{\top}\left(X^{[i-1]}(t)\right) P^{[i]}$$
$$P^{[i]}(t_f) = 0$$

$$\dot{s_f}^{[i]}(t) = -\begin{bmatrix} Qz(t) \\ 0_{3\times 1} \end{bmatrix} - \Big[A\left(X^{[i-1]}(t)\right) \tag{25}$$
$$-B\left(X^{[i-1]}(t)\right) R^{-1} B^{\top}\left(X^{[i-1]}(t)\right) P^{[i]}(t)\Big]^{\top} s_f^{[i]}$$
$$s_f^{[i]}(t_f) = 0$$

The fact that certain entries are autonomous in the dynamic and that the constantly zero function is a solution for them, implies by existence and uniqueness that they should be constantly zero. Then $P^{[i]}$ has $n \times n$ non zero entries and $s^{[i]}$ has n non zero entries, explicitly:

$$P^{[i]}(t) = \begin{bmatrix} \tilde{P}^{[i]}(t) & 0_{n\times 3} \\ 0_{3\times n} & 0_{3\times 3} \end{bmatrix} \tag{26}$$
$$s_f^{[i]}(t) = \begin{bmatrix} \tilde{s}_f^{[i]}(t) \\ 0_{1\times 3} \end{bmatrix} \tag{27}$$

Where $\tilde{P}^{[i]}(t)$ and $\tilde{s_f}^{[i]}(t)$ satisfy differential equations (28) and (29), respectively. Note that matrices depend on the previous state iteration i.e. $A_{11} = A_{11}\left(X^{[i-1]}(t)\right)$ and $B_1 = B_1\left(X^{[i-1]}(t)\right)$.

$$\dot{\tilde{P}}^{[i]}(t) = -Q - \tilde{P}^{[i]} A_{11} - A_{11}^{\top}\tilde{P}^{[i]} + \tilde{P}^{[i]} B_1 R^{-1} B_1^{\top}\tilde{P}^{[i]} \tag{28}$$
$$\tilde{P}^{[i]}(t_f) = 0$$

$$\dot{\tilde{s}}_f^{[i]}(t) = -Qz(t) - \left[A_{11} - B_1 R^{-1} B_1^{\top}\tilde{P}^{[i]}(t)\right]^{\top} \tilde{s}_f^{[i]} \tag{29}$$
$$\tilde{s}_f^{[i]}(t_f) = 0$$

Assume now that matrices Q and R are diagonal matrices. Inspecting the equation (28) one notices that if $i \neq j$, $P_{ij}(t) = 0$ is a solution of all non diagonal entries. And therefore, once again, by existence and uniqueness they are constantly zero. Hence in system (28) only the diagonal entries should be calculated as seen in equation (30) .

$$\dot{\tilde{P}}_{jj}^{[i]}(t) = -Q_{jj} - 2\left(f_j\left(s^{[i-1]}\right) - D\right)\tilde{P}_{jj}^{[i]} \tag{30}$$
$$+R^{-1}\left(\tilde{P}_{jj}^{[i]}\right)^2 f_j\left(s^{[i-1]}\right)^2\left(x_j^{[i-1]}\right)^2$$
$$\tilde{P}_{jj}^{[i]}(t_f) = 0$$

And the control law is given by

$$v^{[i]}(t)_j = \tag{31}$$
$$0 \wedge \left(-1 \vee \left(-R^{-1} B_1^{\top}\left(X^{[i-1]}(t)\right)\left(\tilde{P}^{[i]}(t) x^{[i]}(t) - \tilde{s}_f^{[i]}(t)\right)\right)_j\right)$$
$$\forall j \in [n]$$

Differential equations (30) and (29) were solved using standard backward numerical integration for each iteration. Note that from the former computations the method relies in calculating $2n$ differential equations instead of $(n+3)^2 + (n+3)$ as the straightforward application of the method would have meant.

III. Application and Discussion

The choice of matrices Q and R are always of discussion in a quadratic regulator. Q represents the covariance matrix of measurements, while the interpretation of R is not clear. In the simulations here presented $Q = I_n$ and $R = 0.0001 I_n$, found by trial and error in order for the control to fit accurately the species abundance. The method stopped after 110 iterations where each iteration took approximately 600 seconds of computing time in a computer equipped with 8gb of RAM memory and Intel core i3-7100U CPU 2,40 GHz.

The tracking problem was applied to data coming from a nitrification process; the experimental conditions are described in [9]. For exploring the hypothesis of interactions as drivers of bioreactors performance environmental conditions should be kept as constant as possible. Therefore only data from day 183 onwards was used because a change in the operating temperature happened at that point, which is known to have an effect on the kinetics. For choosing which species belong in which functional group, the procedure described in [10] was used: from day 183 to day 230: 17 OTU were identified in the G_1 group (known as Ammonium oxidizing bacteria AOB) and 8 in the G_2 functional group (known as Nitrite oxidizing bacteria NOB). The system state dimension becomes 28.

The knowledge of functions $f_i(s)$ was based on a study of nitrification's kinetic parameters[11]. Particularly given the system's ammonium and nitrate concentration a Monod function (eq (32)) was used for G_1 and G_2.

$$f_i(s) = \bar{\mu}_2 \frac{s_1}{K_1 + s_1} \; \forall i \in G_1 \tag{32}$$
$$f_i(s) = \bar{\mu}_2 \frac{s_2}{K_2 + s_2} \; \forall i \in G_2$$

For the reader to gain understanding of the situation, a simulation of the system using the experiments operating parameters (D and s_{in}) is presented without control (i.e. $v(t) = 0$) in Figure 1. s_3 accumulates all along the trajectory, but when compared to data it is clear that s_3 stops accumulating after a while.

When applying the presented method one obtains the simulation that can be seen in figure 2. The method captures the tendencies in the measured substrates, but not their absolute values. The simulation of each functional G_1 (AOB), and G_2 (NOB) can be seen in figures 3 and 4, respectively. Note that the fit of G_1 seems better than the fit of G_2, which can be explained by the 1 order of magnitude difference of abundance between the members of the group.

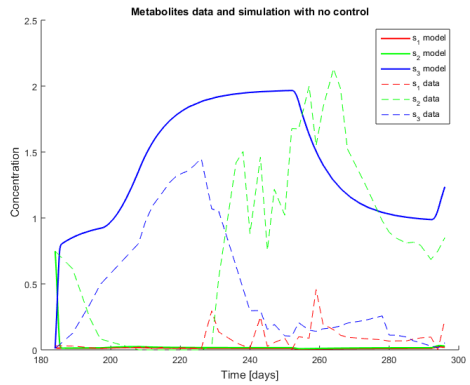

Fig. 1: Simulation when $u = 1$.

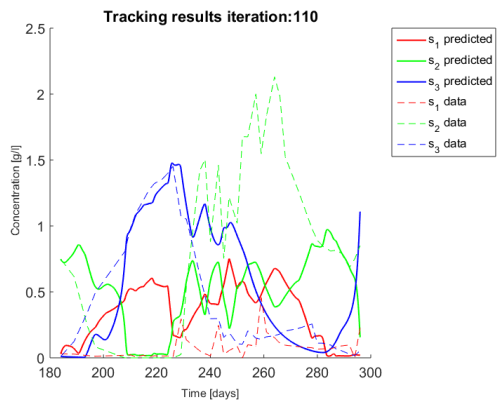

Fig. 2: Simulation applying the control strategy.

One can observe that in figure 2 the predicted s_3 around day 240 is much higher than the measurement. If one looks at figure 4 it suggests attributing this extra production of s_3 around day 240 to the growth of OTU 24. What would have happened if OTU 24 had been classified in the other functional group? One would have seen a decrease in s_1, an increase in s_2, and a decrease in s_3 which could have better fitted the measurements. Thus suggesting that OTU 24 was misclassified from the method presented in [10].

In figures 5 and 6 it can be seen the obtained controls for each functional group. Both controls saturate at certain points, but in general terms they vary throughout the dynamic.

Recalling that $u(t)$ is supposed to represent $I(x(t))$ (the interactions terms), one could try to fit different forms of $I(\cdot)$ for example $I(x) = 1 + Ax$, where A is a real valued square matrix. In figure 7 the simulation of the system after the identification of matrix A is presented, suggesting that an affine interaction function is not able to capture the behaviour of the system.

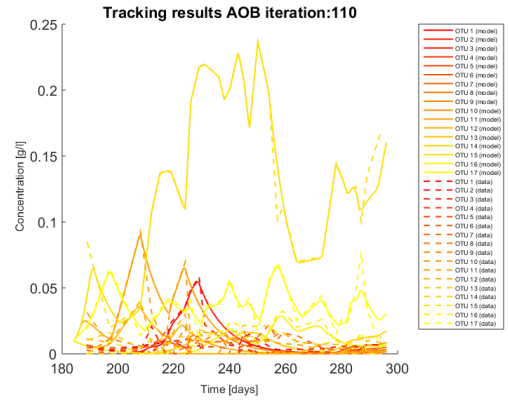

Fig. 3: Abundance of G_1 (AOB) simulation applying the control strategy.

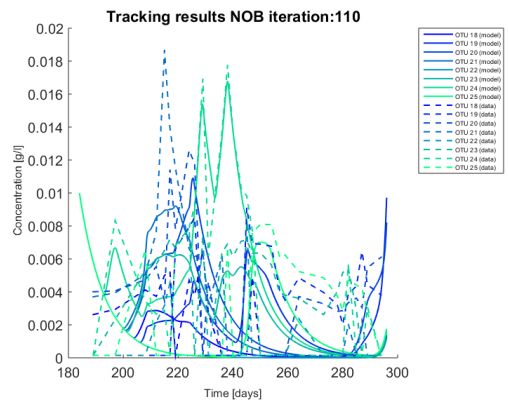

Fig. 4: Abundance of G_2 (NOB) simulation applying the control strategy.

IV. Conclusion and Perspectives

Considering the dimension of the system analysed, the approach here presented seems satisfactory for the integration of genetic sequencing in the context of bioreactors, particularly the Ricatti equations grow linearly in the number of species, rendering the method scalable for a bigger number of species. However some issues should be taken care of: 1) test the method on synthetic data for a class of interaction functions 2) the optimality of the control should be improved, considering is just a truncated control from a state feedback 3) A clear rational on the choice of matrices Q and R. Finally the hypothesis of misclassification presented for OTU 24 proposes as well a retroactive way to reclassify an OTU if the substrates predictions go astray.

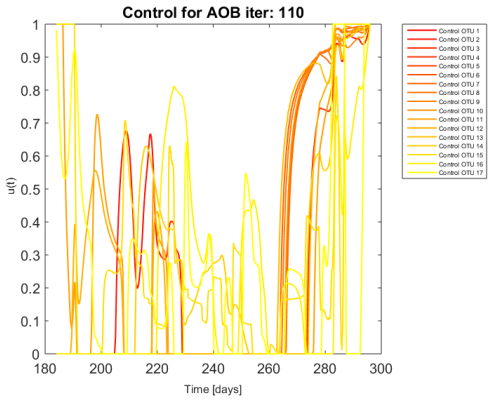

Fig. 5: Control obtained for OTU in G_1 (AOB).

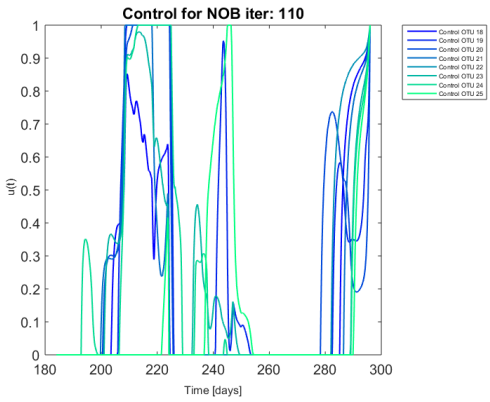

Fig. 6: Control obtained for OTU in G_2 (NOB).

ACKNOWLEDGMENT

This work was supported by project Thermomic ANR-16-CE04-0003.

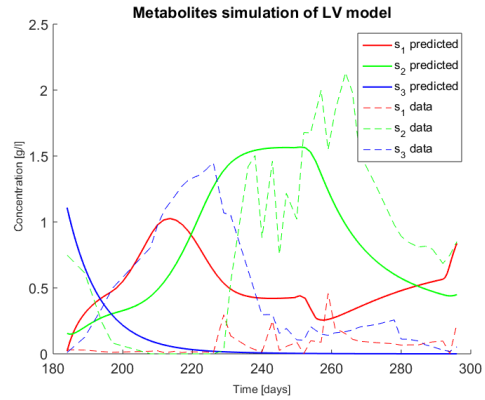

Fig. 7: Simulation obtained for the interaction function $I(x) = 1 + Ax$, where A was fitted with the obtained controls from figures 5 and 6.

REFERENCES

[1] S. Widder, R. J. Allen, T. Pfeiffer, T. P. Curtis, C. Wiuf, W. T. Sloan, O. X. Cordero, S. P. Brown, B. Momeni, W. Shou, H. Kettle, H. J. Flint, A. F. Haas, B. Laroche, J. U. Kreft, P. B. Rainey, S. Freilich, S. Schuster, K. Milferstedt, J. R. Van Der Meer, T. Grobkopf, J. Huisman, A. Free, C. Picioreanu, C. Quince, I. Klapper, S. Labarthe, B. F. Smets, H. Wang, O. S. Soyer, S. D. Allison, J. Chong, M. C. Lagomarsino, O. A. Croze, J. Hamelin, J. Harmand, R. Hoyle, T. T. Hwa, Q. Jin, D. R. Johnson, V. de Lorenzo, M. Mobilia, B. Murphy, F. Peaudecerf, J. I. Prosser, R. A. Quinn, M. Ralser, A. G. Smith, J. P. Steyer, N. Swainston, C. E. Tarnita, E. Trably, P. B. Warren, and P. Wilmes, "Challenges in microbial ecology: Building predictive understanding of community function and dynamics," 2016.

[2] M. J. Wade, J. Harmand, B. Benyahia, T. Bouchez, S. Chaillou, B. Cloez, J.-J. Godon, B. M. Boudjemaa, A. Rapaport, T. Sari *et al.*, "Perspectives in mathematical modelling for microbial ecology," *Ecological Modelling*, vol. 321, pp. 64–74, 2016.

[3] C. Zuñiga, L. Zaramela, and K. Zengler, "Elucidation of complexity and prediction of interactions in microbial communities," 2017.

[4] E. H. Battley, R. L. Putnam, and J. Boerio-Goates, "Heat capacity measurements from 10 to 300 K and derived thermodynamic functions of lyophilized cells of Saccharomyces cerevisiae including the absolute entropy and the entropy of formation at 298.15 K," *Thermochimica Acta*, vol. 298, no. 1-2, pp. 37–46, sep 1997. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0040603197001081

[5] M. A. Savageau, " Energetics and Kinetics in Biotechnology. J. A. Roels ," *The Quarterly Review of Biology*, 2004.

[6] C. Lobry, A. Rapaport, T. Sari *et al.*, *The Chemostat: Mathematical Theory of Microorganism Cultures*. John Wiley & Sons, 2017.

[7] T. Çimen and S. P. Banks, "Nonlinear optimal tracking control with application to super-tankers for autopilot design," *Automatica*, 2004.

[8] T. Çimen, "State-Dependent Riccati Equation (SDRE) control: A survey," in *IFAC Proceedings Volumes (IFAC-PapersOnline)*, 2008.

[9] M. Dumont, A. Rapaport, J. Harmand, B. Benyahia, and J.-J. Godon, "Observers for microbial ecology-how including molecular data into bioprocess modeling?" in *Control and Automation, 2008 16th Mediterranean Conference on*. IEEE, 2008, pp. 1381–1386.

[10] P. Ugalde-Salas, J. Harmand, and E. D.-L. Quemener, "Asymptotic Observers and Integer Programming for Functional Classification of a Microbial Community in a Chemostat," 2019.

[11] U. Wiesmann, *Biological nitrogen removal from wastewater*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1994, pp. 113–154. [Online]. Available: https://doi.org/10.1007/BFb0008736