IDENTIFICATION AND ROBUSTNESS ANALYSIS OF
NONLINEAR HYBRID DYNAMICAL SYSTEM OF GENETIC
REGULATION IN CONTINUOUS CULTURE

QI YANG
School of Mathematical Science, Dalian University of Technology
School of Energy and Engineering, Dalian University of Technology
Dalian, Liaoning 116024, China

LEI WANG* AND ENMIN FENG
School of Mathematical Sciences, Dalian University of Technology
Dalian, 116024, China

HONGCHAO YIN
School of Energy and Engineering, Dalian University of Technology
Dalian, Liaoning 116024, China

ZHIHONG XIU
School of Life Science and Biotechnology, Dalian University of Technology
Dalian, Liaoning 116024, China

(Communicated by Changzhi Wu)

Abstract. In this paper, we present a framework to infer the possible transmembrane transport of intracellular substances. Considering four key enzymes, a modified fourteen-dimensional nonlinear hybrid dynamic system is established to describe the microbial continuous culture with enzyme-catalytic and genetic regulation. A novel quantitative definition of biological robustness is proposed to characterize the system’s resilience when system parameters were perturbed. It not only considers the expectation of system output data after parameter disturbance but also considers the influence of the variance of these data. In this way, the definition can be used as an objective function of the system identification model due to the lack of data on the concentration of intracellular substances. Then, we design a parallel computing method to solve the system identification model. Numerical results indicate that the most likely transmembrane mode of transport is active transport coupling with passive diffusion for glycerol and 1,3-propanediol.

2010 Mathematics Subject Classification. Primary: 90B30, 65C20; Secondary: 35B20.

Key words and phrases. Pathway identification, nonlinear hybrid dynamical system, robustness analysis, parameter identification, continuous culture, parallel optimization.

This work was supported by the National Natural Science Foundation of China (Grant Nos. 11771008, 11171050 and 11371164), the National Science Foundation for the Youth of China (Grant Nos. 11301051, 11301081 and 11401073), the Provincial Natural Science Foundation of Fujian (Grant Nos. 2014J05001), the Fundamental Research Funds for Central Universities in China (Grant DUT15LK25) and the China Scholarship Council (CSC, No. 201506060021). The authors acknowledge the Supercomputer Center of Dalian University of Technology for providing computing resources.

* Corresponding author. E-mail address: wanglei@dlut.edu.cn.
1. Introduction. The bioconversion of glycerol to 1,3-propanediol (1,3-PD) has a wide range of potential applications on a large commercial scale, especially as a monomer for polyesters, polyethers and polyurethanes, due to its lower cost, higher production rate and no pollution [2, 35]. There are three typical cultures for microbial fermentation of glycerol, including batch, continuous and fed-batch cultures. Over the past years, a number of models have been proposed to describe the bio-dissimilation process of glycerol to 1,3-PD by *Klebsiella pneumoniae* (*K. pneumoniae*).

In recent years, some nonlinear dynamical systems have been proposed to describe the microbial fermentation process mentioned above. A well-cited suitable dynamic model was first proposed by Zeng et al. [43]. In this model, only the concentrations of microorganisms, glycerol, 1,3-PD, and two other major by-products were considered. Then, Xiu et al. [36, 37] made further improvements based on this model to analyze some qualitative and quantitative characteristics of the modified model. Based on this improved mathematical model, some scholars have studied the parameter identification problem, stability of equilibria, phenomena of oscillation and hysteresis for the continuous fermentation [5, 6, 14, 16, 17, 18, 21, 22, 39]. However, the above literatures didn’t consider the effect of changes in the concentration of intracellular substances on the microbial fermentation reaction. With the deepening of research, such as the understanding of intracellular transport of substances across the membrane can be very helpful for the research in genetic engineering and other related aspects, the impact of intracellular substances cannot be ignored. Therefore, in 2008, Sun et al. [27] proposed a novel mathematical model, in which the concentration changes of both extracellular substances and intracellular substances were taken into consideration. Thereafter, a large number of further investigations have been carried out, including sensitivity analysis [31], pathway identification problems [30, 38, 44], stochastic optimal control [33] and robust parameter identification [42].

In 2012, Sun et al. [28] proposed a fourteen-dimensional nonlinear dynamic system concerning the repression of the *dha* regulon and two key enzymes by 3-hydroxypropionaldehyde (3-HPA) to describe the continuous culture, which is more accurate and complex than eight-dimensional model. On the basis of the fourteen-dimensional model, Yuan et al. [41] establish a nonlinear hybrid dynamical system with uncertain system parameters and pathway parameters for describing the process of batch culture. Gao et al. [7] propose a novel global sensitivity analysis method to reduce the parameters in the parameter identification model. Considering the three possible transport mechanisms of 1,3-PD across cell membrane. The transmembrane transport of intracellular glycerol was not considered in the above literature, Guo et al. [8] establish a nonlinear hybrid dynamical system about continuous culture, as this would impose a huge computational burden on the subsequent analysis.

Since it is difficult to accurately measure the concentration of intracellular substances, it is inevitable to determine the most reasonable glycerol metabolism mechanism by means of numerical analysis. Hiroaki [12, 13] proposed a popular concept of biological robustness to judge the reliability of numerical solution for the concentrations of intracellular substances. Biological robustness is a basic characterization of biological systems, which has also been accepted by researchers in the field [19, 24, 25, 34]. Many researchers have tried to give a quantitative definition of biological robustness and achieved some desired results [32, 40]. The core idea
of quantitative definition is to quantify the changes in system performance by perturbing system parameters. In this paper, we establish a parameter identification model with a novel quantitative definition of biological robustness. And through this identification model to infer the most likely glycerol transmembrane transport. The novel quantitative definition not only considers the expectation of system output data after parameter disturbance but also considers the influence of the variance of these data. In this way, the definition can be used as an objective function of the system identification model due to the lack of data on the concentration of intracellular substances, the relative error of extracellular substances and the approximate stability of dynamical systems as constraint conditions. Finally, we design a parallel computing method to solve the system identification model. Then, we obtain the optimal pathway and parameters of the corresponding system.

This paper is organized as follows. In Section 2, a nonlinear dynamical system of enzyme-catalytic and genetic regulation of continuous culture is presented. In Section 3, some important properties of the system are discussed. In Section 4, a new quantitative description of biological robustness is presented and with this presented biological robustness as performance index, a system identification model is established. Section 5 constructs a parallel particles warm optimization algorithm for the system identification model. In Section 6, a numerical result is given to infer the parameter identification model. Conclusions are presented at the end of the paper.

2. Nonlinear hybrid dynamical system with genetic regulation. The biochemical process under consideration is the dissociation of glycerol in Klebsiella pneumoniae, which is a disproportionation process that occurs through coupled oxidation and reduction pathways, as shown in Fig.1. During glycerol metabolism, glycerol is transformed from the extracellular environment to the intracellular environment and then converted to 3-hydroxypropionaldehyde (3-HPA) by glycerol dehydratase (GDHt). The important intermediate 3-HPA was further reduced to 1,3-PD by enzyme 1,3-propanediol oxidoreductase (PDOR). Finally, the by-products of the target product 1,3-PD and coupled oxidation are transformed from the intracellular environment to the extracellular environment. Based on the above process, we propose a fourteen-dimensional nonlinear dynamic system with two key reductases to describe continuous culture [28].

Since the transport mechanisms of glycerol and 1,3-PD across cell membrane in the reductive pathway are still unclear, we shall regard the metabolic pathway with a combination of possible transport mechanisms of glycerol and 1,3-PD as a possible metabolic system, which is convenient to be represented by a simple metabolic network. Taking all possible transport mechanisms of glycerol and 1,3-PD into consideration, we can obtain nine possible combinations of transports mechanisms, which correspond to nine possible simple metabolic networks (or nine possible metabolic systems), denoted by NHDS($l_1$), NHDS($l_2$), ···, NHDS($l_9$), see Table 1.

According to the experiment process, we make the following assumptions:

(H1) The concentrations of reactants are uniform in reactor, time delay and nonuniform space distribution are ignored.

(H2) During the process of continuous culture, the substrate added to the reactor only includes glycerol and the fermentation broth is exported by the dilution rate $D$.

Let $x(t) := (x_1(t), x_2(t), \ldots, x_{14}(t))^T \in \mathbb{R}^{14}, t \in [t_0, t_f]$ ( $t_0 = 0$ is the given initial time, $t_f$ is the given terminal time ) be the state vector, where $x_i(t)(i = 1, 2, \ldots, 14)$
Fig. 1. Main pathway of 1,3-propanediol biosynthesis in *K. pneumoniae*. [28]. Abbreviations: GDH–glycerol dehydrogenase; GDHt–glycerol dehydratase; DHAK I–dihydroxyacetone kinases (ATP dependent); DHAK II–dihydroxyacetone kinases (PEP dependent); PDOR–1,3-propanediol oxydoreductase; HOR–hypothetical oxydoreductase; PDH–pyruvate dehydrogenase; PFL–pyruvate format lyase; DHA–dihydroxyacetone; DHAP–dihydroxyacetone phosphate; PEP–phosphoenolpyruvate; Pyr–pyruvate; HAc–acetate; EtOH–ethanol; 3-HPA–3-hydroxypropionaldehyde; 1,3-PD–1,3-propanediol; RP–regulatory protein.

respectively denote the concentration of biomass, extracellular glycerol, extracellular 1,3-PD, acetic, ethanol, intracellular glycerol, intracellular 3-HPA, intracellular 1,3-PD, mR, R, mGDHt, GDHt, mPDOR and PDOR at time $t$, respectively. To simplify notations, $I_n := \{1, 2, \cdots, n\}$ and $x_i := x_i(t)$. Let $N$ be the total number of experiments carried out under different dilution rates and initial glycerol concentrations $(D_j, C_{s0,j})$, $j \in I_N$, and $p_{j,k} := (p_{1,j,k}, p_{2,j,k}, \cdots, p_{31,j,k})^T$, $k \in I_9$ be the kinetic parameter vector to be determined.

Under the assumptions (H1) and (H2), the continuous fermentation process of glycerol by *K. pneumoniae* of the $k$th metabolic system in the $j$th experiment, denoted by $\text{NHDS}(j, l_k)$, can be described as follows:
\[
\dot{x}_1 = (\mu - D_j)x_1, \\
\dot{x}_2 = D_j(C_{s_{0,j}} - x_2) - q_2x_1, \\
\dot{x}_i = q_ix_1 - D_jx_i, \quad i = 3, 4, 5, \\
\dot{x}_6 = \frac{1}{p_{j,k}^{11}}\left[\frac{18}{k_jp_{j,k}}x_2 + \frac{1}{p_{j,k}^{10}}(x_2 - x_6)N_{R+}(x_2 - x_6) - q_2^0\right] - \mu x_6, \\
\dot{x}_7 = \frac{p_{j,k}^{11}x_6x_{12}}{K_{mG}(1 + \frac{x_6}{p_{j,k}^{10}})} - \frac{p_{j,k}^{13}x_7x_{14}}{K_{mP} + x_7(1 + \frac{x_7}{p_{j,k}^{18}})} - \frac{p_{j,k}^{15}(p_{j,k}^{16})^2p_{j,k}^{17}x_7}{x_7 + p_{j,k}^{18}}, \\
\dot{x}_8 = \frac{p_{j,k}^{13}x_7x_{14}}{K_{mP} + x_7(1 + \frac{x_7}{p_{j,k}^{18}})} - \frac{1}{k_jp_{j,k}^{19}}x_8 - \frac{1}{p_{j,k}^{20}x_7} - \frac{1}{p_{j,k}^{21}}p_{j,k}^{01}(x_8 - x_3)N_{R+}(x_8 - x_3) - \mu x_8, \\
\dot{x}_9 = \frac{x_7 + p_{j,k}^{18}}{x_7 + p_{j,k}^{18} + p_{j,k}^{22}} - (p_{j,k}^{23} + \mu)x_9, \\
\dot{x}_10 = \frac{x_7 + p_{j,k}^{18}}{x_7 + p_{j,k}^{18} + p_{j,k}^{22}} - (p_{j,k}^{23} + \mu)x_{10}, \\
\dot{x}_11 = \frac{x_7 + p_{j,k}^{18}}{x_7 + p_{j,k}^{18} + p_{j,k}^{22}} - (p_{j,k}^{23} + \mu)x_{11}, \\
\dot{x}_12 = \frac{x_7 + p_{j,k}^{18}}{x_7 + p_{j,k}^{18} + p_{j,k}^{22}} - (p_{j,k}^{23} + \mu)x_{12}, \\
\dot{x}_13 = \frac{x_7 + p_{j,k}^{18}}{x_7 + p_{j,k}^{18} + p_{j,k}^{22}} - (p_{j,k}^{23} + \mu)x_{13}, \\
\dot{x}_14 = \frac{x_7 + p_{j,k}^{18}}{x_7 + p_{j,k}^{18} + p_{j,k}^{22}} - (p_{j,k}^{23} + \mu)x_{14},
\]

where \(D_j\) and \(C_{s_{0,j}}\) are respectively the dilution rate and the glycerol concentration in the feed medium in the \(j\)th experiment. The specific cellular growth rate \(\mu\) in Eq.(1) can be given as [43]:

\[
\mu = \mu_m \frac{x_2}{x_2 + K_s}(1 - \frac{x_2}{x_2})(1 - \frac{x_3}{x_3})(1 - \frac{x_4}{x_4})(1 - \frac{x_5}{x_5})(1 - \frac{x_7}{x_7}).
\]

For \(K.\ pneumoniae\) grown under anaerobic conditions at 37°C and pH value 7.0, the maximum specific growth rate \(\mu_m\) and Monod constant \(K_s\) in Eq.(13) take the values of 0.67h\(^{-1}\) and 0.28mmol/L, where, \(K_{mG}\) and \(K_{mP}\) denote Michaelis-Menten constants of enzymes GDHt and enzymes PDOR, which are 0.53mmol/L and 0.14mmol/L respectively.

The transport system of glycerol and 1,3-3PD are considered, according to Monod equation and Fick diffusion law, the specific consumption rate of glycerol and specific product formation of 1,3-PD can be expressed by

\[
q_2 = l_k^{11}p_{j,k}^{10}x_2 + l_k^{12}p_{j,k}^{11}x_2 - (x_2 - x_6)N_{R+}(x_2 - x_5),
\]

\[
q_3 = l_k^{24}p_{j,k}^{10}x_8 + l_k^{25}p_{j,k}^{11}x_8 - (x_8 - x_3)N_{R+}(x_8 - x_3),
\]

where \(l_k^i \in \{0,1\}, \ i \in I_4, \ k \in I_9, \ l_k^1 = 0 (l_k^2 = 0)\) and \(l_k^3 = 0 (l_k^4 = 0)\) indicates glycerol and 1,3-PD pass the membrane the passive diffusion (resp. active transport) doesn’t exist, whereas \(l_k^1 = 1 (l_k^2 = 1)\) and \(l_k^3 = 1 (l_k^4 = 1)\) represents the
presence of active transport of glycerol and 1,3-PD (resp. passive diffusion), respectively. According to the actual fermentation process, we have \( \max\{l^1_k, l^2_k\} = 1 \) and \( \max\{l^3_k, l^4_k\} = 1 \). Let \( l_k := (l^1_k, l^2_k, l^3_k, l^4_k)^T \in S_9 := \{l_1, l_2, \cdots, l_9\} \), \( k \in I_9 \), then \( l_k \) corresponds to the \( k \)th metabolic system \( NHDS(l_k) \) in the complex metabolic network \( NHDS \). For details, see Table 1.

| \( NHDS(l_k) \rightarrow l_k \) | Glycerol | 1,3-PD |
|-------------------------------|---------|--------|
| \( NHDS(l_1) \rightarrow l_1 \) | P       | P      |
| \( NHDS(l_2) \rightarrow l_2 \) | P       | A      |
| \( NHDS(l_3) \rightarrow l_3 \) | P       | AP     |
| \( NHDS(l_4) \rightarrow l_4 \) | A       | P      |
| \( NHDS(l_5) \rightarrow l_5 \) | A       | A      |
| \( NHDS(l_6) \rightarrow l_6 \) | A       | AP     |
| \( NHDS(l_7) \rightarrow l_7 \) | AP      | P      |
| \( NHDS(l_8) \rightarrow l_8 \) | AP      | A      |
| \( NHDS(l_9) \rightarrow l_9 \) | AP      | AP     |

In addition, the indicator function defined as

\[
N_{R+}(z) = \begin{cases} 
1, & z \geq 0, \\
0, & z < 0.
\end{cases}
\]

The uptake of extracellular glycerol is considered as “black box” model, its specific consumption rate expressed by [27]:

\[
q_2^0 = m_2 + \mu Y_2 + \Delta q_2 \frac{x_2}{x_2 + K_2^*}. \tag{16}
\]

The specific formation rate of extracellular acetate and ethanol can be expressed respectively as follows [27]:

\[
q_4 = m_4 + \mu Y_4 + \Delta q_4 \frac{x_2}{x_2 + K_4^*}, \tag{17}
\]

\[
q_5 = m_5 + \mu Y_5. \tag{18}
\]

The parameters in Eqs.(16)-(18) were estimated by the previous work [37], as shown in Table 2.

| \( i \) | \( m_i \) | \( Y_i \) | \( \Delta q_i \) | \( K_i^* \) |
|-------|--------|--------|--------|--------|
| 2     | 2.20   | 0.0082 | 28.58  | 11.43  |
| 4     | 0.97   | 33.07  | 5.74   | 85.71  |
| 5     | 5.26   | 11.66  | -      | -      |
To simplify notation, we denote $p_{j,k}$ as $p$, the nonlinear hybrid dynamical system NHDS($j,l_k$), can be described by the combination of the following two systems:

$$
\begin{align*}
\dot{x}(t) &= f(t; x, p, j, l_k), & t \in [t_0, t_b], \\
x(t_0) &= x_0, \\
\end{align*}
$$
(19)

$$
\begin{align*}
\dot{x}(t) &= f(t; x, p, j, l_k), & t \in [t_b, t_f], \\
x(t_b) &= x_b, \\
\end{align*}
$$
(20)

where $f(t; x, p, j, l_k) := (f_1(t; x, p, j, l_k), f_2(t; x, p, j, l_k), \ldots, f_{14}(t; x, p, j, l_k))^T$ is the rate of reactions with the parameter vector $p := p_{j,k} := (p_{j,k}^1, p_{j,k}^2, \ldots, p_{j,k}^{31})^T$.

Denote the time intervals of batch and continuous periods by $[t_0, t_b]$ and $[t_b, t_f]$, respectively, where $t_0 < t_b < t_f < \infty$. The system (19) describes the batch fermentation process, in which $x_0$ is the initial state, $x_b$ is the terminal state, and the dilution rates to be zero. The system (20) describes the continuous fermentation process starting from the terminal state of the system (19) with prescribed dilution rate $D_j \neq 0$.

According to the practical culture process, the concentrations of biomass, substrate, intermediates and products will not exceed their critical concentrations. Let $x_{si}$ and $x^*_i$ be the lower and upper bounds of the state vector $x_i$, so we consider the properties of the system on:

$$W_{ad} := \prod_{i=1}^{14} [x_{si}, x^*_i] \subseteq \mathbb{R}^{14},$$

$$x_{si} = (0.001, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)^T,$$

$$x^*_i = (10, 2039, 939.5, 1026, 360.9, 2039, 300, 939.5, M, M, M, M, M)^T,$$

where $M$ is a large enough positive number.

Let $D_c := [0.08, 0.5] \times [110.96, 1883] \subseteq \mathbb{R}^2$ be the admissible set of the dilution rate and the initial glycerol concentration in the feed medium.

Let $p_s$ and $p^*$ be the lower and upper bounds of the parameter vector $p_{j,k}$, denote $P_{ad}$ be the allowable set of vector $p_{j,k}$, which can be given by [28]:

$$P_{ad} := [p_s, p^*] = \prod_{i=1}^{31} [p_i^0 - \Delta p_i^0, p_i^0 + \Delta p_i^0] \subseteq \mathbb{R}^{31},$$

(22)

where $|\Delta p_i^0| := 0.5p_i^0$,

$$p^0 := (p_{1,0}^0, p_{2,0}^0, \ldots, p_{31,0}^0)^T = (80.26, 1.90, 419.17, 55.06, 3.71, 39.69, 0.15, 54.29, 1.53, 1053.3, 44.94, 45.14, 25.15, 2.24, 1.41, 14.05, 1.8 \times 10^{-5}, 1264.44, 5.56, 17.62, 27.73, 17.14, 0.16, 0.766, 319.92, 1.59, 21.49, 8.77, 11.85, 26.22, 14.04)^T.$$

3. **Properties of the nonlinear hybrid dynamical system.** The existence, uniqueness and continuous differentiability of the solution to the system NHDS($j,l_k$) with respect to parameters are explored in this section. In addition, we denote Euclidean norm by $|| \cdot ||$ in this paper.

According to the factual fermentation mechanism, we have the following assumptions:

(H3) The set $W_{ad}$, $D_c$ and $P_{ad}$ are all nonempty, bounded and closed convex sets.
(H4) The absolute differences of concentrations between extracellular and intracellular glycerol concentration and that of 1, 3-PD concentration are uniformly bounded, i.e., \( \exists M_1 > 0, M_2 > 0 \), such that
\[
| x_2(t) - x_0(t) | < M_1, \forall t \in [t_0, t_f]; \quad | x_3(t) - x_0(t) | < M_2, \forall t \in [t_0, t_f]. \tag{23}
\]

**Property 1.** The functions \( f(t; x, p, j, l_k) : [t_0, t_f] \times W_{ad} \times P_{ad} \times I_N \times S_I \to \mathbb{R}^4 \), defined by Eqs. (1)-(12) satisfy that:

(a) \( f(t; x, p, j, l_k) \) are Lipschitz in \( x \) on \( W_{ad} \);

(b) \( f(t; x, p, j, l_k) \) satisfies the linear growth condition, i.e., \( \exists a, b > 0 \), such that
\[
\| f(t; x, p, j, l_k) \| \leq a \| x \| + b, \quad \forall x \in W_{ad}. \tag{24}
\]

**Proof.** The conclusions of the property can be easily obtained by using the methods similar to the proof of Property 1 in [7]. \( \square \)

Under the assumptions (H3) and (H4) and Property 1, we can obtain the following properties.

**Property 2.** For given \( (j, l_k) \in I_N \times S_I \) and \( p \in P_{ad} \), there exists a unique solution to the system \( \text{NHDS}(j, l_k) \), denoted by \( x(\cdot; x_0, p, j, l_k) \). Furthermore, \( x(\cdot; x_0, p, j, l_k) \) is continuous in \( p \) on \( P_{ad} \).

For given \( (j, l_k) \in I_N \times S_I \), \( p \in P_{ad} \), \( x_0 \in W_{ad} \), we define the following three sets:

\[
S_0(j, l_k) := \left\{ x(\cdot; x_0, p, j, l_k) \left| x(\cdot; x_0, p, j, l_k) \right. \text{is the solution of system } \text{NHDS}(j, l_k) \right\}.
\]

\[
S_2(j, l_k) := \left\{ x(\cdot; x_0, p, j, l_k) \left| x(\cdot; x_0, p, j, l_k) \in S_0(j, l_k) \right. \text{and } x(\cdot; x_0, p, j, l_k) \in W_{ad} \right\}.
\]

\[
D_p(j, l_k) := \left\{ p \in P_{ad} \left| x(\cdot; x_0, p, j, l_k) \in S_2(j, l_k) \right. \right\}.
\]

**Property 3.** Under the assumptions (H3) and (H4), given \( (j, l_k) \in I_N \times S_I \), if the sets \( S_0(j, l_k) \), \( S_2(j, l_k) \) and \( D_p(j, l_k) \) are all nonempty, then \( S_0(j, l_k), S_2(j, l_k) \) are compact in \( C([t_0, t_f], \mathbb{R}^4) \) and \( D_p(j, l_k) \) is compact in \( \mathbb{R}^{31} \).

4. Biological robustness analysis and parameter identification model. The purpose of this section is to establish a reasonable criterion to infer the most impossible metabolic mechanism. In continuous fermentation experiments, although the steady-state data of extracellular substances can be measured, the intracellular metabolic mechanisms are not completely clear, and it is difficult to measure the concentration of the intracellular substances and the accuracy of measurement.

To deal with this problem, a novel quantitative definition of biological robustness is proposed to characterize the system’s resilience when system parameters were perturbed. It not only considers the expectation of system output data after parameter disturbance but also considers the influence of the variance of these data. Thus, this quantitative definition can be used as part of the objective function in the system identification model to characterize the lack of data for intracellular concentrations. Then we discuss the existence of the optimal solution to the identification problem. Some basic definitions are given as follows.

**Definition 4.1.** (Approximately Stable State [31]) For given \( (j, l_k) \in I_N \times S_I \), a state vector \( x(t_s; x_0, p, j, l_k) \in W_{ad} \) is said to be an approximately stable solution within a precision \( \eta > 0 \), if there exists \( p \in D_p(j, l_k) \) and \( t_s \in [t_0, t_f] \) such that \( x(t; x_0, p, j, l_k) \) is a solution of system \( \text{NHDS}(j, l_k) \) satisfying
(The biological robustness of the system) For given $(j, l_k) \in I_N \times S_I$ and $p \in D_p(j, l_k)$, let $x_i(t_s; x_0, p, j, l_k), i \in I_3$ be the $i$th component of approximately stable state by the system NHDS$(j, l_k)$ and $y_i(j) \in \mathbb{R}^3, i \in I_3$ be the $j$th experimental extracellular concentration (i.e., biomass, glycerol and 1,3-PD) measured at the steady stage. Then, the relative error of extracellular substances is defined as:

$$SSD(p, j, l_k) := \frac{3 \sum_{i=1}^{3} \left| x_i(t_s; x_0, p, j, l_k) - y_i(j) \right|}{\sum_{i=1}^{3} |y_i(j)|}.$$  

(26)

The set of feasible parameter vector is defined as follows:

$$D_{ps}(j, l_k) := \{ p \in P_{ad} \mid SSD(p, j, l_k) \leq \varepsilon_0 \},$$  

(27)

where the $\varepsilon_0$ is a given small constant, $D_{ps}(j, l_k) \subseteq D_p(j, l_k)$ is nonempty and compact in $\mathbb{R}^{31}$.

For given $\tilde{p} := (\tilde{p}_1, \tilde{p}_2, \cdots, \tilde{p}_{31})^T \in D_{ps}(j, l_k)$ and $p := (p_1, p_2, \cdots, p_{31})^T \in D_{ps}(j, l_k)$. We define the rectangular field of dimension normalization, denoted by $B(p, \delta)$, can be expressed as follows:

$$B(p, \delta) := \left\{ \tilde{p} \in D_{ps}(j, l_k) \mid \| p - \tilde{p} \| \leq \delta \right\},$$

(28)

where $\delta > 0$ is sufficiently small to measure the magnitude of the parameter disturbances. Obviously, $B(p, \delta)$ is compact in $\mathbb{R}^{31}$.

Definition 4.3. (Relative deviation with respect to parameters perturbation) Given $(j, l_k) \in I_N \times S_I$, assume that the system reaches an approximately stable state at $t_s$ and $t_s^*_k$ with the kinetic parameter vector taking values $p', p'' \in B(p, \delta)$, respectively. Let $t_s := \max\{t_s^*, t_s^*_k\}$ denote the corresponding approximately stable solutions by $x_i(t_s; p', j, l_k)$ and $x_i(t_s; p'', j, l_k)$ ($i = 6, \cdots, 14$). The relative deviation of intracellular substances with regard to parameter vector $p'$ against the disturbance $p''$ is defined as

$$RSD_i(p', p'', j, l_k) := \frac{\left| x_i(t_s; p', j, l_k) - x_i(t_s; p'', j, l_k) \right|}{x_i(t_s; p', j, l_k)}, \quad i = 6, \cdots, 14.$$  

(29)

For given parameter vector $p \in D_{ps}(j, l_k)$, the expectation and variance of the relative deviation $RSD_i(p', p'', j, l_k), i = 6, \cdots, 14$, caused by $p'$ uniformly distributed in field of $B(p, \delta)$ are respectively defined by

$$E_i(t_s; RSD_i, p, j, l_k) := \int_{B(p, \delta)} \frac{1}{B(p, \delta)} RSD_i(p', p'', j, l_k) d(p' - p''),$$

(30)

$$D_i(t_s; RSD_i, p, j, l_k) := E_i(t_s; RSD_i^2, p, j, l_k) - [E_i(t_s; RSD_i, p, j, l_k)]^2.$$  

(31)

Definition 4.4. (The biological robustness of the system) For given $(j, l_k) \in I_N \times S_I$ the robustness of system NHDS$(j, l_k)$ about parameter vector $p_{j,k} \in D_{ps}(j, l_k)$
is defined as

\[ J(p_{j,k}, j, l_k) := \frac{1}{9} \sum_{i=6}^{14} \left[ w_1 E_i(t_s; RSD_i, p_{j,k}, j, l_k) + w_2 D_i(t_s; RSD_i, p_{j,k}, j, l_k) \right], \quad (32) \]

where \( w_1, w_2 \) are the weight coefficients for the balance of the magnitude.

The identification problem of finding the corresponding parameters such that the computational values of extracellular steady states approximate the experimental results as possible as they can and that the system is the most robust, can be formulated as:

\[
\text{IPM} : \quad \min J(p_{j,k}, l_k) := \frac{1}{N} \sum_{j=1}^{N} J(p_{j,k}, j, l_k) \quad \text{(33)}
\]

\[
s.t. \quad x(t; x_0, p_{j,k}, j, l_k) \in W_{ad}, \; \forall t \in [t_0, t_f], \quad (34)
\]

\[
\| f(t; x, p_{j,k}, j, l_k) \| \leq \varepsilon, \; \forall t \in [t_s, t_f], \quad (35)
\]

\[
\text{SSD}(p_{j,k}, j, l_k) \leq \varepsilon_0, \quad (36)
\]

\[
p_{j,k} \in P_{ad}, \quad (37)
\]

\[
l_k \in S_t, \quad (38)
\]

where Eq. (34) is said to be continuous state inequality constraints \([29]\), which must be satisfied at every point in the time horizon \([t_0, t_f]\). Eq. (35) is the state vector \( x(t; x_0, p_{j,k}, j, l_k) \in S_{x}(j, l_k) \) to be an approximately stable solution, Eq. (36) is a quality constraint for ensuring an acceptable level of the system NHDS\((j, l_k)\) better fits given experimental data.

**Theorem 4.5.** For given \((j, l_k) \in I_N \times S_t\), assumed that feasible region of IPM is nonempty, there exists an optimal solution of IPM, that is \( \exists p^*_{j,k} \in D_{ps}(j, l_k) \) such that

\[ J(p^*_{j,k}, l_k) \leq J(p_{j,k}, l_k), \; \forall p \in D_{ps}(j, l_k). \quad (39) \]

**Proof.** By the Properties 2, 3 and definition of the \( J(p_{j,k}, l_k) \), we know that \( J(p_{j,k}, l_k) \) is continuous at \( p_{j,k} \in D_{ps}(j, l_k) \). Since \( D_{ps}(j, l_k) \) is compact in \( \mathbb{R}^{31} \), there exists an optimal solution of IPM. \( \square \)

In the identification problem IPM, the objective function \( J(p_{j,k}, l_k) \) is continuous with respect to the parameter vector \( p_{j,k} \in \mathbb{R}^{31} \) \((j \in I_N, \; \kappa \in I_9)\) and discrete with respect to the path parameter \( l_k \) \((i \in I_4, \; k \in I_9)\). We consider 9 possible paths and \( N \) experiments, so in IPM we have \( 31 \times 9 \times N \) continuous variables and \( 4 \times 9 \times N \) discrete variables. Therefore, the burden of calculation for the definition of biological robustness will increase significantly as the number of experiments increases. Through observation, we find that system parameters \( p_{j,k} \) and path parameters \( l_k \) are independent of each other in the objective function \( J(p_{j,k}, l_k) \), the original problem thence can be decomposed into two sub-problems to optimize the path parameters and system parameters.

Firstly, given \( l_k \in S_t \), the identification subproblem corresponding to system NHDS\((j, l_k)\) denoted by IPM\(_1\)(\(l_k\)), can be formulated as:

\[
\text{IPM}\(_1\)(l_k) : \quad J_1(p^*_{j,k}, l_k) := \min J(p_{j,k}, l_k)
\]

s.t. \( x(t; x_0, j, l_k) \in W_{ad}, \; \forall t \in [t_0, t_f], \)

\[
p_{j,k} \in D_{ps}(j, l_k). \quad (32)
\]
Since \( l_k \in S_t \) is given, there is only continuous parameter vector \( p_{j,k} \in D_{ps}(j, l_k) \) in IPM(\( l_k \)).

Secondly, the identification subproblem concerning \( l_k \in S_t \), denoted by IPM, can be formulated as:

\[
\text{IPM}_2: \quad J_2(p_{j,k}^*, l_k^*) := \min_{l_k \in S_t} J(p_{j,k}^*, l_k),
\]

where \( p_{j,k}^* \) is the optimal parameter vectors of IPM(\( l_k \)). Let \( l_k^* := \text{argmin}\{J(p_{j,k}^*, l_k) \mid l_k \in S_t)\} \). Then, the optimal robustness value and the optimal discrete parameter vectors of IPM are \( J(p_{j,k}^*, l_k^*) \) and \( l_k^* \).

### 5. Numerical solutions

The identification problem IPM given by last section can be treated as a optimal parameter selection problem. Similar with the reference [23], any kind of gradient-based optimization methods can be used to solve the problem IPM. However, those gradient-based algorithm are designed to find local optimal solutions. In order to solve the IPM problem globally, we combined a gradient-based and particle swarm optimization method to design a global search algorithm.

#### 5.1. Constraint transcription

When we intend to solve the problem IPM, our main difficulty lies in solving the subproblem IPM(\( l_k \)). In nature, given \( l_k \in S_t \), IPM(\( l_k \)) is a constrained optimization problem. However, it is difficult to cope with the continuous state inequality constraints. To surmount these difficulties, let

\[
g_i(t; p_{j,k}, j, l_k) := x_i^* - x_i(t; x_0, p_{j,k}, j, l_k), \]

\[
g_{14+i}(t; p_{j,k}, j, l_k) := x_i(t; x_0, p_{j,k}, j, l_k) - x_{s+i}, \quad i \in I_{14}.
\]

The condition \( x(t; x_0, p, j, l_k) \in W_{ad} \) is equivalently transcribed into

\[
G(p_{j,k}, j, l_k) = 0,
\]

where \( G(p_{j,k}, j, l_k) := \sum_{i=1}^{28} r_i \min_{t_0} \{0, g_i(t; p_{j,k}, j, l_k)\} dt \). However, \( G(p_{j,k}, j, l_k) \) is non-smooth and the non-differentiable points are at \( g_i(t; p_{j,k}, j, l_k) = 0 \), \( i \in I_{28} \). Consequently, standard optimization routines would have difficulties with this type of equality constrains. The following smoothing technique introduced in [15] is to replace \( \min\{0, g_i(t; p_{j,k}, j, l_k)\} \) with \( \varphi_{\varepsilon,i}(t; p_{j,k}, j, l_k) \), where

\[
\varphi_{\varepsilon,i}(t; p_{j,k}, j, l_k) := \begin{cases} 
  g_i(t; p_{j,k}, j, l_k), & \text{if } g_i(t; p_{j,k}, j, l_k) < -\varepsilon, \\
  (g_i(t; p_{j,k}, j, l_k) - \varepsilon) / 4\varepsilon, & \text{if } -\varepsilon \leq g_i(t; p_{j,k}, j, l_k) \leq \varepsilon, \\
  0, & \text{if } g_i(t; p_{j,k}, j, l_k) > \varepsilon,
\end{cases}
\]

where \( \varepsilon > 0 \) is an adjustable parameter controlling the accuracy of the approximation. Then, the equality constraint (40) now can be approximated by

\[
\tilde{G}_{\varepsilon,\gamma}(p_{j,k}, j, l_k) := \gamma + \sum_{i=1}^{28} \int_{t_0}^{t_f} \varphi_{\varepsilon,i}(t; p_{j,k}, j, l_k) dt \geq 0,
\]
where $\gamma > 0$ is an adjustable parameter. Hence, with this approximation scheme, the concentration bounds in problem IPM and subproblem IPM_1(l_k) are approximated by the single canonical constraint (41). This constraint is a standard constraint and can be readily handled using the computational algorithm described later in Section 5.2.

Then, IPM and IPM_1(l_k) can be approximated by the following approximate problems called IPM$_{\varepsilon,\gamma}$ and IPM$_{1,\varepsilon,\gamma}(l_k)$:

**IPM$_{\varepsilon,\gamma}$:**
\[
\begin{align*}
\text{min } & \quad \mathcal{J}(p_{j,k}, l_k) := \frac{1}{N} \sum_{j=1}^{N} J(p_{j,k}, j, l_k) \\
\text{s.t. } & \quad \tilde{G}_{\varepsilon,\gamma}(p_{j,k}, j, l_k) := \gamma + \sum_{i=1}^{28} \int_{t_0}^{t_f} \varphi_{\varepsilon,i}(t; p_{j,k}, j, l_k) dt \geq 0, \\
& \quad \|f(t; x, p_{j,k}, j, l_k)\| \leq \varepsilon, \quad \forall t \in [t_s, t_f], \\
& \quad \text{SSD}(p_{j,k}, j, l_k) \leq \varepsilon_0, \\
& \quad p_{j,k} \in P_{ad}, \\
& \quad l_k \in S_1.
\end{align*}
\]

**IPM$_{1,\varepsilon,\gamma}(l_k)$:**
\[
\begin{align*}
\text{min } & \quad \mathcal{J}_1(p_{j,k}^*, l_k) := \min \mathcal{J}(p_{j,k}, l_k) \\
\text{s.t. } & \quad \tilde{G}_{\varepsilon,\gamma}(p_{j,k}, j, l_k) := \gamma + \sum_{i=1}^{28} \int_{t_0}^{t_f} \varphi_{\varepsilon,i}(t; p_{j,k}, j, l_k) dt \geq 0, \\
& \quad p_{j,k} \in D_{ps}(j, l_k).
\end{align*}
\]

**Theorem 5.1.** Given $(j, l_k) \in I_N \times S_1$, for each $\varepsilon > 0$, there exists a $\gamma(\varepsilon) > 0$ such that for all $\gamma$, $0 < \gamma < \gamma(\varepsilon)$, any feasible solution of the model IPM$_{1,\varepsilon,\gamma}(l_k)$, and IPM$_{\varepsilon,\gamma}$, is also a feasible solution of the model IPM$_1(l_k)$, and IPM.

**Remark 1.** If the parameter $p_{j,k}$ is not feasible, we can move the parameter towards the feasible region in the direction of the gradients of constraint $G_{\varepsilon,\gamma}(p_{j,k}, j, l_k)$ with respect to parameter $p_{j,k}$. Then, the gradient of the constraint $G_{\varepsilon,\gamma}(p_{j,k}, j, l_k)$ will be needed and the gradient information is given by the next theorem.

**Theorem 5.2.** Given $l_k \in S_1$, for each $\varepsilon > 0$, $\gamma > 0$, the derivatives of the constraint functionals $G_{\varepsilon,\gamma}(p_{j,k}, j, l_k)$ with respect to the $i$th $(i \in I_3)$ component of the parameter vector are

\[
\frac{\partial \tilde{G}_{\varepsilon,\gamma}(p_{j,k}, j, l_k)}{\partial p_{j,k}^i} := \int_{t_0}^{t_f} \frac{\partial H(t; x, p_{j,k}, l_k, \lambda(t))}{\partial p_{j,k}^i} dt, \quad (42)
\]

where

\[
H(t; x, p_{j,k}, l_k, \lambda(t)) := \sum_{i=1}^{28} \varphi_{\varepsilon,i}(t; p_{j,k}, j, l_k) + \lambda^T(t)f(t; x, p_{j,k}, j, l_k), \quad (43)
\]

and

\[
\lambda(t) := (\lambda_1(t), \lambda_2(t), \ldots, \lambda_{14}(t))^T, \quad (44)
\]

is the solution of the costate system

\[
\dot{\lambda}(t) = -\frac{\partial H(t; x, p_{j,k}, l_k, \lambda(t))}{\partial x}, \quad (45)
\]
with the boundary conditions

\[ \lambda(t_f) = (0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0) \]

Proof. The proof can be completed using the method of Chapter 3 in [4].

5.2. Optimization algorithms. Particle Swarm Optimization (PSO) is an evolutionary computational method which is based on swarm intelligence. PSO was proposed by Kennedy and Eberhart inspired by the research of the artificial life in [11]. Since then, it has been widely used to solve a broad range of optimization problems. PSO is based on a simple concept, similar to evolutionary optimization methods, PSO is a derivative-free, population-based global search algorithm. Therefore, the computation time is short and it requires few memories. It was originally developed for nonlinear optimization problems with continuous variables and discrete variables [1, 3, 10, 26].

To solve the parameters identification, we design a parallel algorithms as follows, is similar to [20].

Parallel Algorithms 1:

Parallel algorithm of solving identification problem IPM, is as follows:

- \( p_{j,k} := (p^i_{j,k}, p^2_{j,k}, \cdots, p_{j,k}^3) \) be the kinetic parameter vector to be determined.
- \( M_\rho \) is the total number of particles in the swarm.
- \( A_{sup} \) is a sufficient large number.
- \( V_{min} \) and \( V_{max} \) are vectors containing the minimum and maximum particle velocities.
- \( \theta \) is the iteration index, \( \Theta_{max} \) is the minimum and maximum number of iterations.
- \( \rho_{best} \) is the personal best position of the \( w \)th particle.
- \( \rho_{best} \) is the best position of whole population.
- \( c_1 \) and \( c_2 \) are the cognitive and social scaling parameters.
- \( r_1 \) and \( r_2 \) are the random numbers with the uniform distribution on \([0, 1]\).
- \( w \) is the inertia weight, \( w_{min} \) and \( w_{max} \) are the minimum and maximum inertia weights.
- \( \tau_1 \) and \( \tau_2 \) are the convergence tolerances.

Step 1: Allocate \( n_\rho \) processes for the algorithm. Let the population size be \( M_\rho \). Then the number of particles is \( n_w = M_\rho/n_\rho \). Initialized data on the root process as follows:

1.1. Read known data such as \( \mu_m, K_s, K_mG, K_mP, W_{ad}, P_{ad}; m_i, Y_i, \Delta q_i, K_s', i = 2, \cdots, 5 \).

1.2. Set \( x_0, t_b, t_f, \varepsilon, \sigma, \delta, M, M_\rho, V_{min}, V_{max}, c_1, c_2, w_{min}, w_{max}, \tau_1, \tau_2, \Theta_{max} \).

1.3. Read experimental data as \( D_j, C_{s0,j}, y_j = (y_{1,j}, \cdots, y_{5,j})^T, j \in I_N \).

2. Broadcast data of root process to all processes.

3. Set \( J(P_{best,j}, l_k) := J(p_{best,w,j}, l_k) := A_{sup} \), and \( j := 1 \).

4. Execute the following procedure on each process, and denote the ID of process as \( \rho \).

4.1. Set \( \theta := 1 \), according to the uniform distribution, randomly generate the positions of \( n_w \) particles from the region \( P_{ad} \). Denote the position and velocity of particle by \( p_{j,k}^{\rho,w,\theta} \in P_{ad} \) and \( p_{j,k}^{\rho,w,\theta} \in [V_{min}, V_{max}] \), \( w = 1, 2, \cdots, n_w \).

4.2. Set \( w := 1 \).

4.3. Put every particle \( p_{j,k}^{\rho,w,\theta} \) in the IPM, and solve the differential
equations defined by Eqs. (1)-(12) by using the improved Euler algorithm. Compute \(G_{\varepsilon, \gamma}(p_{j,k}, j, l_k)\) by (41), denoted by \(G_{\varepsilon, \gamma}(p_{j,k}^w, j, l_k)\), \(SSD(p, j, l_k)\) by (26), denoted by \(SSD(p_{j,k}^w, j, l_k)\) for the \(k\)th metabolic pathway.

Step 4: If \(G_{\varepsilon, \gamma}(p_{j,k}^w, j, l_k) \geq 0\) and \(SSD(p_{j,k}^w, j, l_k) \leq \varepsilon_0\), then compute \(J(p_{j,k}, j, l_k)\) by (32), denoted by \(J(p_{j,k}^w, j, l_k)\). Otherwise, move the parameter \(p_{j,k}^w, i \in I_{31}\) towards the feasible region in the direction of \(\frac{\partial G_{\varepsilon, \gamma}(p_{j,k}, j, l_k)}{\partial p_{j,k}^w}\) by (41), denoted by \(\frac{\partial G_{\varepsilon, \gamma}(p_{j,k}^w, j, l_k)}{\partial p_{j,k}^w}(i), \forall i \in I_{31}\), by Theorem 5.2 with Armijo line searches, and go to Step 4.3.

Step 5: Gather \(J(p_{\text{best}}^w, j, l_k), p_{\text{best}}\) from process \(\rho\) into the root process.

Step 6: Execute the following procedure on each process.

Step 6.1: Set \(\theta := \theta + 1\). Change the velocity and position of each particle then return Step 4.3, update particles according to the following rules:

\[
\nu_{j,k}^{\rho,w,\theta+1}(i) = \varpi(\theta)\nu_{j,k}^{\rho,w,\theta}(i) + c_1r_1(p_{\text{best}}^w(i) - p_{j,k}^{\rho,w,\theta}(i)) + c_2r_2(p_{j,k}^{\text{best}}(i) - p_{j,k}^{\rho,w,\theta}(i)),
\]

where \(\nu_{j,k}^{\rho,w,\theta}(i)\) and \(p_{j,k}^{\rho,w,\theta}(i)\) represent the \(i\)th \((i \in I_{31})\) component of velocity and position of the \(w\)th \((w \in I_{n_w})\) particle of the \(\rho\)th \((\rho \in I_{n_\rho})\) process at the \(\theta\) iteration, respectively. Change \(\varpi(\theta)\) according to the following formula:

\[
\varpi(\theta) = (\varpi_{\max} - \varpi_{\min} - \tau_1)\exp\left(\frac{1}{1 + \tau_2\theta/\Theta_{\max}}\right).
\] (47)

When the parameters are out of their bounds, they are treated as follows:

\[
p_{j,k}^{\rho,w,\theta}(i) = \begin{cases} p_{\text{si}}, & \text{if } p_{j,k}^{\rho,w,\theta}(i) < p_{\text{si}}, \\ p_{\text{si}}^\ast, & \text{if } p_{j,k}^{\rho,w,\theta}(i) < p_{\text{si}}^\ast, \end{cases} \quad \forall i \in I_{31}.
\]

Step 6.2: If \(\theta < \Theta_{\max}\), then set \(\theta := \theta + 1\) and go to Step 4.2, Otherwise go to Step 7.

Step 7: Execute the following procedure on the root process.

Step 7.1: Output the whole population \(p_{\text{best}}^\rho\), and compute \(J(p_{\text{best}}^\rho, j, l_k)\).

Step 7.2: If \(j < N\), then set \(j := j + 1\), and go to Step 4. Otherwise, compute \(J(p_{\text{best}}^\rho, l_k) := \frac{1}{N}\sum_{j=1}^{N} J(p_{\text{best}}^\rho, j, l_k)\), then stop.

Combining Algorithm 1 with Theorem 5.1, we can develop the following algorithm to solve the IPM_{1}(l_k).

Algorithms 2:

Step 1: Choose initial values of \(\varepsilon > 0, \gamma > 0\).

Step 2: Solve Problem IPM_{1,\varepsilon,\gamma}(l_k) using Algorithm 1 to give \(p_{j,k,\varepsilon,\gamma}^\ast\) and \(l_k\).

Step 3: Check feasibility of \(g_i(t; p_{j,k,\varepsilon,\gamma}^\ast, j, l_k)\) for \(\forall t \in [t_0, t_f], \forall i \in I_{28}\).

Step 4: If \(p_{j,k,\varepsilon,\gamma}^\ast\) is feasible, go to Step 5. Otherwise, set \(\gamma := \alpha\gamma\). If \(\gamma > \gamma_t\), where \(\gamma_t\) is a prespecified positive constant, go to Step 2. Otherwise go to Step 6.

Step 5: Set \(\varepsilon := \beta\varepsilon\). If \(\varepsilon > \varepsilon_0\), where \(\varepsilon_0\) is a prespecified positive constant, go to
Step 2. Otherwise go to Step 6.

Step 6: Output $p^s_{j,k,z,\gamma}$ and $l_k$, then stop.

Remark 2. In Algorithm 2, $\varepsilon$ is a parameter controlling the accuracy of the smoothing approximation. $\gamma$ is a parameter controlling the feasibility of the constraint (41). $\bar{\varepsilon}$ and $\bar{\gamma}$ are two predefined parameters ensuring the termination of the algorithm.

Remark 3. It is important for the validity of Algorithm 2 to choose the parameters $\alpha$, $\beta$, $\bar{\varepsilon}$ and $\bar{\gamma}$. Especially, the parameters $\alpha$ and $\beta$ must be chosen less than 1. $\bar{\varepsilon}$ and $\bar{\gamma}$ are two sufficiently small values such that Algorithm 2 can be terminated.

Remark 4. The formula of the $E(t_s; RSD_i, p, j, l_k)$ is expressed as an integral form in Eq.(30), which would be difficult to compute directly. Hence, we need divide with the uniform distribution on the perturbation space $m_0$ subdivisions. The formula of the $E_i(t_s; RSD_i, p, j, l_k)$ can be change as follows:

$$E_i^m(t_s; RSD_i, p, j, l_k) := \sum_{q=1}^{m_0} \frac{RSD_i(p, p^q; j, l_k)}{m_0}, \; i = 6, 7, \cdots, 14. \quad (48)$$

where $p^q(q \in I_{m_0})$ are uniformly distributed in field of $B(p, \delta), \; p \in D_{ps}(j, l_k)$, and $\delta$ is the radius of this field.

To find a appropriate value of the number of subdivisions $m_0$, we propose the following algorithm:

Algorithm 3:

Step 1 : Given $\tau > 0$, choose two positive integers $A < B$, and set $\delta := 0$.

Step 2 : It follows from (48) that we compute the value of $E^A(t_s; RSD_i, p, j, l_k)$ and $E^B(t_s; RSD_i, p, j, l_k)$.

Step 3 : If $\left| \sum_{i=6}^{14} E^A_i(t_s; RSD_i, p, j, l_k) - E^B_i(t_s; RSD_i, p, j, l_k) \right| \leq \tau$, then set $B := A$, $A := 0.5A$, $\delta := 1$ and go to Step 2; otherwise, go to Step 4.

Step 4 : If $\left| \sum_{i=6}^{14} E^A_i(t_s; RSD_i, p, j, l_k) - E^B_i(t_s; RSD_i, p, j, l_k) \right| > \tau$, and $\delta = 0$, then set $A := 2A$, $B := 2B$ and go to Step 2; otherwise, go to Step 5.

Step 5 : If $\delta = 1$, then set $m_0 := 2A$, output $m_0$ and stop.

6. Numerical results. In this section, we choose $N = 3$ as experimental data and initial state $x_0 = (0.1, 400, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)^T$. In order to approximate the factual experiments, we take $\sigma = 0.01, \varepsilon_0 = 0.3, \delta = 0.01, t_0 = 5h, t_f = 100h, M = 100$ and let Euler step be $1/7200h$. Furthermore, in the improved PSO algorithm, the number of particles $M_p$ is 60, the maximum iteration $\Theta_{max} = 500$, the $x$th component of maximum allowable velocity of particle is $\nu_{max} = (p_x - p_{ax})/2$, the cognitive and social scaling parameters $c_1 = 2.0, c_2 = 2.0$, the maximum and minimum inertia weight $\varpi_{max} = 0.9, \varpi_{min} = 0.4$, the controlling factors $\tau_1 = 0.20, \tau_2 = 0.7$, sufficiently large number $A_{sup} = 10000$, respectively. In Algorithm 2, $\varepsilon := 0.1, \gamma := 0.01, \alpha := 0.1, \beta := 0.01, \bar{\varepsilon} := 10^{-7}, \bar{\gamma} := 1.0 \times 10^{-7}$. Seek for the appropriate value of the number of subdivisions $m_0$ is using by Algorithm 3. According to Eqs.(30)-(32) and Eq.(48), the results for $m_0 \geq 5000$ show far less variation in the value of the cost function than the results for $m_0 < 5000$. That is, we can see that 5000 is an appropriate value of $m_0$. 
The identification model IPM includes 27 nonlinear dynamical systems, 837 continuous variables and 108 discrete variables. Problem IPM is solved by using Algorithms 1 on Lenovo LingYun Large-scale Clusters Supercomputer (which is located in School of Dalian University of Technology, China). As shown in Table 3, for the metabolic systems NHDS\(l_k\), \(k = 2, 4, 7\), their robustness index are set to be \(+\infty\), because the sets \(D_{ps}(j, l_k)\), \(k = 2, 4, 7\) are empty for all \(j \in I_N\) even if the number of random samples is large enough, which indicates that these metabolic systems (NHDS\(l_2\), NHDS\(l_4\) and NHDS\(l_7\)) are all very poor in representing the actual fermentation process, the optimal performance index of IPM with \(l_9 = (1, 1, 1, 1)^T\) is the smallest, which demonstrates that it is most possible that 1,3-PD and glycerol passes the cell membrane by passive diffusion coupling with active transport. The corresponding optimal dynamical parameter vector \(p^*(l_9)\) with \(l_9 = (1, 1, 1, 1)^T\) is shown in Table 4. The approximate stable state vector \(x^*\) under the optimal dynamical parameter vector \(p^*(l_9)\) is shown in Table 5. The comparison of the three extracellular (biomass, glycerol, 1,3-PD) concentrations between experimental data and computational results is shown in Fig. 2, and Fig. 3 - Fig. 4 show the simulated results of intracellular (glycerol, 3-HPA,1,3-PD, mR, R, mGDHt, DGHt, mPDOR, PDOR) concentrations, respectively.

**Table 3.** The computational values of robustness index for system NHDS corresponding to parameter vector \(l_k\), \(k \in I_g\),

| \(k\) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|------|---|---|---|---|---|---|---|---|---|
| \(J(p^*,j,l_k)\) | 1.748 | \(+\infty\) | 15.267 | \(+\infty\) | 3.6521 | 5.689 | \(+\infty\) | 0.657 | 0.1258 |

**Table 4.** The computational values of the optimal parameters under three experiments.

\[(D, C_{so})\] | \(p^*(l_9)\)
---|---
\((0.08, 435)\) | \((81.7276, 4.97131, 1075.54, 99.9598, 1.89429, 23.9368, 0.362061, 28.1892, 8.95164, 7.79472 \times 10^{-4}, 40.415, 41.1295, 27.8876, 1.12982, 1.54052, 10.0365, 4.9745 \times 10^{-5}, 949.377, 9.47578, 25.4648, 32.6392, 12.7944, 0.1133, 0.520596, 191.37, 1.75482, 20.1505, 9.64496, 14.0072, 0.8037, 13.9279)^T\)
\((0.08, 152)\) | \((89.7339, 2.31116, 680.762, 77.858, 2.64026, 68.5132, 0.296838, 18.2558, 9.80988, 9.56009 \times 10^{-4}, 40.1752, 49.5181, 27.3997, 1.22681, 1.95972, 17.3442, 8.18616 \times 10^{-7}, 998.73, 13.4516, 23.1863, 42.6794, 8.60437, 0.058844, 0.550928, 262.521, 1.69241, 25.8673, 5.83124, 9.62714, 22.5502, 11.171)^T\)
\((0.23, 375.7)\) | \((87.3566, 4.11554, 599.756, 86.7367, 1.47777, 94.5825, 0.269305, 34.3741, 9.01509, 8.30467 \times 10^{-4}, 33.6429, 48.385, 27.4557, 2.66168, 1.97522, 10.6856, 4.51274 \times 10^{-5}, 1012.51, 16.0382, 25.8691, 40.1752, 8.52086, 0.196155, 0.790479, 315.66, 1.75482, 25.8673, 5.83124, 9.62714, 22.5502, 15.35)^T\)
Table 5. The experimental data \( y \) and the numerical results \( x^* \) for three experiments.

| \((D, C_{s0})\) | \(y\) | \(x^*\) |
|-----------------|-------|--------|
| (0.08, 0.435)   | (2.6, 0.22, 136) | (2.66078, 0.199219, 151.787, 56.1728, 204.64, 55.188) |
|                 | (0.189992, 48.8683, 0.0906137, 3.18097, 0.00879915, 0.335116, 0.694374, 0.043648, 0.0648233) |
| (0.08, 152)     | (1.1, 0.09, 53)  | (1.15502, 0.0720033, 53.1286, 24.2607, 88.8432, 28.63) |
|                 | (0.0605281, 47.8053, 0.130965, 5.17053, 0.0108471, 0.405015, 1.77233, 0.073951, 0.148219) |
| (0.23, 375.7)   | (2.86, 0.91, 148.5) | (2.858, 0.893436, 150.942, 83.1966, 98.1893, 65.7) |
|                 | (0.869124, 58.9767, 0.240652, 1.59739, 0.00399712, 0.327636, 0.962971, 0.0785403, 0.171447) |

Figure 2. The simulated results of three extracellular concentrations \( x_1 \) - biomass, \( x_2 \) - glycerol, \( x_3 \) - 1,3-PD) under three different initial conditions, i.e. \((D, C_{s0}) = (0.08, 435)\), \((D, C_{s0}) = (0.08, 152)\) and \((D, C_{s0}) = (0.23, 375.7)\).

7. Discussions and conclusions.

In this paper, we establish a complex metabolic system composed of 9 possible pathways and formulated the corresponding nonlinear hybrid dynamical to describe the continuous culture fermentation of glycerol by \( K.\ pneumoniae \). We then discuss
some basic properties of the metabolic system. A quantitative biological robustness for the concentrations of intracellular substances is defined by penalizing a weighted sum of the expectation and variance of the relative deviation between system outputs before and after parameters are perturbed. In order to determine the parameters, a parameter identification model is presented. The identifiability of the parameter identification model is also proved. Because of the parameter variables of the system concluding continuous and discrete variables, and the parameter identification model does not has analytical solution, we construct a parallel optimization algorithm, based on particle swarm optimization migration (PSO), for solving the identification model. Numerical results indicate that the nonlinear dynamical system is suitable for the process of continuous culture. The results in this work are consistent with previous conjectures in fermentation.

In a future work, our efforts will focus on multistage analysis and optimal control of glycerol input in continuous culture. Additionally, necessary condition of optimality for the system identification model is waiting for publication in another paper.

REFERENCES

[1] G. Ardizzon, G. Cavazzini and G. Pavesi, Adaptive acceleration coefficients for a new search diversification strategy in particle swarm optimization algorithms, Information Sciences, 299 (2015), 337–378.
MAXIMUM PRINCIPLE FOR INVESTMENT/DIVIDEND PROBLEM 597

Figure 4. The simulated results of six intracellular concentrations ($x_9$ - mR, $x_{10}$ - R, $x_{11}$ - mGDHt, $x_{12}$ - DGHt, $x_{13}$ - mPDOR, $x_{14}$ - PDOR) under three different initial conditions, i.e. $(D, C_{SO}) = (0.08, 435)$, $(D, C_{SO}) = (0.08, 152)$ and $(D, C_{SO}) = (0.23, 375.7)$.

[2] H. Biebl, K. Menzel, A. P. Zeng and W. D. Deckwer, Microbial production of 1,3-propanediol, *Applied Microbiology and Biotechnology*, 52 (1999), 289–297.

[3] M. R. Bonyadi and Z. Michalewicz, A locally convergent rotationally invariant particle swarm optimization algorithm, *Swarm Intelligence*, 8 (2014), 159–198.

[4] A. Bryson and Y. C. Ho, *Applied Optimal Control*, Halsted Press, New York, 1975.

[5] G. Cheng, L. Wang, R. Loxton and Q. Lin, Robust optimal control of a microbial batch culture process, *Journal of Optimization Theory and Applications*, 167 (2015), 342–362.

[6] C. X. Gao, Y. Lang, E. M. Feng and Z. L. Xiu, Nonlinear impulsive system of microbial production in fed-batch culture and its optimal control, *Journal of Applied Mathematics and Computing*, 19 (2005), 203–214.

[7] K. K. Gao, X. Zhang, E. M. Feng and Z. L. Xiu, Sensitivity analysis and parameter identification of nonlinear hybrid systems for glycerol transport mechanisms in continuous culture, *Journal of Theoretical Biology*, 347 (2014), 137–143.
[8] Y. J. Guo, E. M. Feng, L. Wang and Z. L. Xiu, Complex metabolic network of 1,3-propanediol transport mechanisms and its system identification via biological robustness, Bioprocess and Biosystems Engineering, 37 (2014), 677–686.
[9] S. Honda, T. Toraya and S. Fukui, In situ reactivation of glycerol-inactivated coenzyme B12-dependent enzymes, glycerol dehydratase and diol dehydratase, Journal of Bacteriology, 143 (1980), 1458–1465.
[10] C. Karakuzu, Fuzzy controller training using particle swarm optimization for nonlinear system control, ISA Transactions, 48 (2009), Page 245.
[11] J. Kennedy and R. C. Eberhart, Particle swarm optimization, in: Proceedings of the IEEE International Conference on Neural Networks, IEEE Press, Australia, (1995), 1942–1948.
[12] H. Kitano, Biological robustness, Nature Reviews Genetics, 5 (2004), 826–837.
[13] H. Kitano, Towards a theory of biological robustness, Molecular Systems Biology, 3 (2007), 137.
[14] H. S. Lian, E. M. Feng, X. F. Li, J. X. Ye and Z. L. Xiu, Oscillatory behavior in microbial continuous culture with discrete time delay, Nonlinear Analysis: Real World Applications, 10 (2009), 2749–2757.
[15] L. S. Jennings and K. L. Teo, A computational algorithm for functional inequality constrained optimization problems, Automatica, 26 (1990), 371–375.
[16] C. Y. Liu, Optimal control of a switched autonomous system with time delay arising in fed-batch processes, IMA Journal of Applied Mathematics, 80 (2015), 569–584.
[17] C. Y. Liu, Z. H. Gong, K. L. Teo, R. Loxton and E. M. Feng, Bi-objective dynamic optimization of a nonlinear time-delay system in microbial batch process, Optimization Letters, 12 (2018), 1249–1264.
[18] C. Y. Liu, Z. Gong, K. L. Teo, J. Sun and L. Caccetta, Robust multi-objective optimal switching control arising in 1, 3-propanediol microbial fed-batch process, Nonlinear Analysis: Hybrid Systems, 25 (2017), 1–20.
[19] C. Y. Liu, Z. H. Gong, K. L. Teo, Robust parameter estimation for nonlinear multistage time-delay systems with noisy measurement data, Applied Mathematical Modelling, 53 (2018), 353–368.
[20] L. Liu, W. X. Liu and D. A. Cartes, Particle swarm optimization based parameter identification applied to permanent magnet synchronous motors, Engineering Applications of Artificial Intelligence, 21 (2008), 1092–1100.
[21] X. H. Li, E. M. Feng and Z. L. Xiu, Stability and optimal control of microorganisms in continuous culture, Journal of Applied Mathematics and Computing, 22 (2006), 425–434.
[22] X. H. Li, J. J. Li, E. M. Feng and Z. L. Xiu, Discrete optimal control model and bound error for microbial continuous fermentation, Nonlinear Analysis: Real World Applications, 11 (2010), 131–138.
[23] Q. Lin, R. Loxton and K. L. Teo, The control parameterization method for nonlinear optimal control: A survey, Journal of Industrial and Management Optimization, 10 (2014), 275–309.
[24] M. Morohashi, A. E. Winn, M. T. Borisuk, H. Bolouri, J. Doyle and H. Kitano, Robustness as a measure of plausibility in models of biochemical networks, Journal of Theoretical Biology, 216 (2002), 19–30.
[25] M. Perc and M. Marhl, Sensitivity and flexibility of regular and chaotic calcium oscillations, Biophysical Chemistry, 104 (2003), 509–522.
[26] J. F. Schutte, J. A. Reinbolt, E. J. Fregly, R. T. Haftka and A. D. George, Parallel global optimization with the particle swarm algorithm, International Journal for Numerical Methods in Engineering, 61 (2004), 2296–2315.
[27] Y. Q. Sun, W. T. Qi, H. Teng, Z. L. Xiu and A. P. Zeng, Mathematical modeling of glycerol fermentation by Klebsiella pneumoniae: Concerning enzyme-catalytic reductive pathway and transport of glycerol and 1,3-propanediol across cell membrane, Biochemical Engineering Journal, 38 (2008), 22–32.
[28] Y. Q. Sun, J. X. Ye, X. J. Mu, H. Teng, E. M. Feng, A. P. Zeng and Z. L. Xiu, Nonlinear mathematical simulation and analysis of dha regulon for glycerol metabolism in Klebsiella pneumoniae, Chinese Journal of Chemical Engineering, 20 (2012), 958–970.
[29] K. L. Teo, C. J. Goh and K. H. Wong, A Unified Computational Approach to Optimal Control Problems, Long Scientific Technical, Essex, 1991.
[30] J. Wang, J. X. Ye, E. M. Feng, H. C. Yin and B. Tan, Complex metabolic network of glycerol fermentation by Klebsiella pneumoniae and its system identification via biological robustness, Nonlinear Analysis: Hybrid Systems, 5 (2011), 102–112.
[31] J. Wang, J. X. Ye, H. C. Yin, E. M. Feng and L. Wang, Sensitivity analysis and identification of kinetic parameters in batch fermentation of glycerol, Journal of Computational and Applied Mathematics, 236 (2012), 2268–2276.

[32] L. Wang, G. M. Cheng, E. M. Feng, T. Su and Z. L. Xiu, Analysis and application of biological robustness as performance index in microbial fermentation, Applied Mathematical Modelling, 39 (2015), 2048–2055.

[33] L. Wang, J. L. Yuan, C. Wu and X. Wang, Practical algorithm for stochastic optimal control problem about microbial fermentation in batch culture, Optimization Letters, (2017), 1–15.

[34] J. M. Whitacre, Biological robustness: Paradigms, mechanisms, and system principles, Frontiers in Genetics, 3 (2012), 1–15.

[35] U. Witt, R. J. Müller, J. Augusta, H. Widdecke and W. D. Deckwer, Synthesis, properties and biodegradability of polyesters based on 1,3-propanediol, Macromolecular Chemistry and Physics, 195 (1994), 793–802.

[36] Z. L. Xiu, A. P. Zeng and W. D. Deckwer, Multiplicity and stability analysis of microorganisms in continuous culture: Effects of metabolic overflow and growth inhibition, Biotechnology and Bioengineering, 57 (1998), 251–261.

[37] Z. L. Xiu, A. P. Zeng and A. J. Li, Mathematical modeling of kinetics and research on multiplicity of glycerol bioconversion to 1,3-propanediol, Journal of Dalian University of Technology, 40 (2000), 428–433.

[38] H. H. Yan, X. Zhang, J. X. Ye and E. M. Feng, Identification and robustness analysis of nonlinear hybrid dynamical system concerning glycerol transport mechanism, Computers and Chemical Engineering, 40 (2012), 171–180.

[39] J. X. Ye, E. M. Feng, H. S. Lian and Z. L. Xiu, Existence of equilibrium points and stability of the nonlinear dynamical system in microbial continuous cultures, Applied Mathematics and Computation, 207 (2009), 307–318.

[40] J. X. Ye, A. Li and J. G. Zhai, A measure of concentration robustness in a biochemical reaction network and its application on system identification, Applied Mathematical Modelling, 58 (2018), 270–280.

[41] J. L. Yuan, X. Zhang, X. Zhu, E. M. Feng, H. C. Yin and Z. L. Xiu, Pathway identification using parallel optimization for a nonlinear hybrid system in batch culture, Nonlinear Analysis: Hybrid Systems, 15 (2015), 112–131.

[42] J. L. Yuan, Y. D. Zhang, J. X. Ye, J. Xie and K. L. Teo, Robust parameter identification using parallel global optimization for a batch nonlinear enzyme-catalytic time-delayed process presenting metabolic discontinuities, Applied Mathematical Modelling, 46 (2017), 554–571.

[43] A. P. Zeng and W. D. Deckwer, A kinetic model for substrate and energy consumption of microbial growth under substrate-sufficient condition, Biotechnology Progress, 11 (1995), 71–79.

[44] J. G. Zhai, J. X. Ye, L. Wang, E. M. Feng, H. C. Yin and Z. L. Xiu, Pathway identification using parallel optimization for a complex metabolic system in microbial continuous culture, Nonlinear Analysis: Real World Applications, 12 (2011), 2730–2741.

Received March 2018; 1st revision June 2018; 2nd revision July 2018.

E-mail address: yangqi0028@163.com
E-mail address: wanglei@dlut.edu.cn
E-mail address: emfeng@dlut.edu.cn
E-mail address: hcyin@dlut.edu.cn
E-mail address: zhlixiu@dlut.edu.cn