Computers and Chemical Engineering 106 (2017) 286–296

Contents lists available at ScienceDirect

Computers and Chemical Engineering

journal homepage: www.elsevier.com/locate/compchemeng

Steady-state optimization of biochemical systems by bi-level programming

Gongxian Xu*, Yang Li

Department of Mathematics, Bohai University, Jinzhou 121013, China

ARTICLE INFO

Article history:
Received 4 August 2016
Received in revised form 13 June 2017
Accepted 14 June 2017
Available online 16 June 2017

Keywords:
Optimization
Bi-level programming
Algorithm
Generalized mass action
Biochemical systems

ABSTRACT

A new method is proposed for the steady-state optimization of biochemical systems described by Generalized Mass Action (GMA) models. In this method, a bi-level programming with a two-layer nested structure is established. In this bi-level problem, the upper-level objective is to maximize a flux or a metabolite concentration, and the lower-level objective is to minimize the total sum of metabolite concentrations of biochemical systems. The biological significance of the presented bi-level programming problem is to maximize the production rate or concentration of the desired product under a minimal metabolic cost to the biochemical system. To efficiently solve the above NP-hard, non-convex and nonlinear bi-level programming problem, we reformulate it into a single-level optimization problem by using appropriate transformation strategies. The proposed framework is applied to four case studies and has shown the tractability and effectiveness of the method. A comparison of our proposed method and other methods is also given.

1. Introduction

Mathematical optimization plays a key role in the establishment of rational strategies for the yield improvement of biochemical systems. In the last decades, much research has been directed towards the development of model-based optimization strategies for biochemical systems (Voit, 1992; Torres et al., 1996, 1997; Hatzimanikatis et al., 1996; Marín-Sanguino and Torres, 2000, 2003; Torres and Voit, 2002; Chang and Sahinidis, 2005; Marín-Sanguino et al., 2007; Polisetty et al., 2008; Xu et al., 2008, 2015; Pozo et al., 2010, 2011, 2015; Sorribas et al., 2010; Vera et al., 2010; Xu, 2012, 2013; Zomorrodi et al., 2012; Hsu and Wang, 2013; Xu and Wang, 2014). For example, Hatzimanikatis et al. (1996) used mixed-integer linear programming to analyze and design the metabolic reaction networks of biochemical systems. Pozo et al. (2010) proposed an outer-approximation algorithm to optimize a biochemical system. Zomorrodi et al. (2012) reviewed a number of optimization-based frameworks developed towards dealing with some challenges in the analysis and engineering for metabolic networks of biochemical systems. Hsu and Wang (2013) developed an approach named fuzzy equal metabolic adjustment to formulate an optimal enzyme target design problem for drug discovery. Pozo et al. (2015) addressed the global optimization of hybrid kinetic/FBA models via an outer-approximation method.

In the modeling framework called Biochemical Systems Theory (BST) (Savageau 1969a,b, 1970, 1976; Savageau et al., 1987a,b; Voit, 2000), some methods have been proposed to address the steady-state optimization of biochemical systems (Voit, 1992; Torres et al., 1996, 1997; Marín-Sanguino and Torres, 2000, 2003; Marín-Sanguino et al., 2007; Xu et al., 2008; Vera et al., 2010; Xu, 2013; Xu and Wang, 2014). For example, Voit (1992) and Torres et al. (1996, 1997) proposed an approach called Indirect Optimization Method (IOM) to solve the steady-state optimization problem of biochemical systems. The advantage of the IOM approach is that one can use a linear programming technique to obtain an optimal steady-state of a biochemical system through the S-system representation and logarithmic transformation. Marín-Sanguino and Torres (2000) used an iterative version of the IOM approach to decrease the difference between the S-system and IOM solutions. Xu et al. (2008) proposed a modified iterative IOM algorithm to attain the globally consistent S-system and IOM solutions by iteratively changing the reference steady-state of the biochemical system under consideration. Marín-Sanguino and Torres (2003) presented a method called GMA-IOM for steady-state optimization of biochemical systems by GMA model representations and linear programming. Marín-Sanguino et al. (2007) and Vera et al. (2010) used controlled error and penalty treatment methods to solve the geometric programming problem generated from biochemical systems described by

* Corresponding author.
E-mail addresses: gxxu@bhu.edu.cn, gongxianxu@aliyun.com (G. Xu).

http://dx.doi.org/10.1016/j.compchemeng.2017.06.019
0098-1354/

the GMA model. These two approaches possibly cannot yield the globally optimal solution of geometric programming with a GMA system (Xu, 2013). To address this issue, an iterative method for the steady-state optimization of biochemical systems was proposed by Xu (2013). Xu and Wang (2014) presented an improved version to enhance the computational efficiency of the Xu (2013) method.

In the study of the steady-state optimization of a biochemical system, one should not only maximize the production rate or concentration of the final product but also optimize its metabolic cost. This is because an abnormally high concentration of intermediate metabolite would cause the biochemical system non-viable, with the burden on the cellular metabolism being too high for the cell to survive (Voit, 2000; Torres and Voit, 2002). Several researchers have used the criteria of minimizing the intermediate metabolite concentrations or their sum to reduce the metabolic cost of biochemical systems (Schuster and Heinrich, 1991; Alvarez-Vasquez et al., 2000; Torres and Voit, 2002). Moreover, a biochemical system has a self-optimization feature for metabolic cost to manage and limit the cell's biosynthetic investments in metabolic material and energy (Torres and Voit, 2002). Therefore, it will be more biologically significant if we can consider and resolve the problem of maximizing the production rate or concentration of desired product under the condition of minimal metabolic cost. For this purpose, we propose a bi-level programming framework for the steady-state optimization of biochemical systems. Bi-level programming is a special form of optimization problems where an optimization problem is embedded within another one (Vicente and Calamai, 1994; Bard, 1998; Dempe, 2002; Dempe et al., 2006; Colson et al., 2007; Domínguez and Pistikopoulos, 2010). The outer optimization task is commonly referred to as the upper-level optimization task, and the inner optimization task is commonly referred to as the lower-level optimization task. A general formulation of bi-level programming problems is written as (Dempe, 2002; Sahin and Ciric, 1998; Colson et al., 2007):

$$\min_{z^U, z^L} \quad F^U\left(z^U, z^L\right)$$

$$\text{s.t.} \quad G^U\left(z^U, z^L\right) \le 0$$

$$H^U(z^U, z^L) = 0$$

$$\min_{z^L} \quad F^L(z^U, z^L)$$

$$\text{s.t.} \quad G^L(z^U, z^L) \le 0$$

$$H^L(z^U, z^L) = 0$$

where $z^U \in R^{n1}$ and $z^L \in R^{n2}$ are the decision variables. The functions $F^U(z^U, z^L) \in R$, $G^U(z^U, z^L) \in R^{m1}$ and $H^U(z^U, z^L) \in R^{p1}$ are the objective function, inequality and equality constraints of the upper-level problem, respectively. The functions $F^L(z^U, z^L) \in R$, $G^L(z^U, z^L) \in R^{m2}$ and $H^L(z^U, z^L) \in R^{p2}$ are the objective function, inequality and equality constraints of the lower-level problem, respectively. From the mathematical point of view, bi-level programming problems are complicated optimization problems due to the following three aspects (Dempe, 2002): (1) they are intrinsically NP-hard; (2) their nested structure has inherent difficulties even with respect to the notion of a solution; (3) for many methods regularity conditions cannot be satisfied at any feasible point. These features of bi-level programming problems make them very difficult to globally solve. In recent years, considerable research has sought to address the solution strategies of bi-level programming problems (Domínguez and Pistikopoulos, 2010; Bosco and Etoa, 2011; Allende and Still, 2013; Kuo et al., 2015; Adasme and Lisser, 2016; Paulavičius et al., 2016; Sinha et al., 2017) and their applications to some processes (Ryu et al., 2004; Cecchini et al., 2013; Yeh et al., 2015; Robbins and Lunday, 2016; Saranwong and Likasiri, 2016, 2017).

In our proposed bi-level programming framework, we propose a bi-level programming problem whose upper-level and lower-level objectives are, respectively, to maximize a flux or a particular metabolite concentration and to minimize the total sum of metabolite concentrations of biochemical systems. A solver is proposed to efficiently solve the presented NP-hard and non-convex bi-level programming problem by transforming it into a simple single-level optimization problem. We illustrate the capabilities of the proposed framework through the steady-state optimization of four biochemical systems, comparing our results with those obtained by other single-level methods.

This paper is organized as follows. Section 2 presents the bi-level programming problem for the steady-state optimization of biochemical systems described by GMA representations. Section 3 develops a method to solve the proposed bi-level programming problem. In Section 4, four case studies are presented. Finally, brief conclusions are given in Section 5.

2. Bi-level programming problem for the steady-state optimization of biochemical systems

2.1. The GMA formalism

A biochemical system can be modeled as the following mathematical form:

$$\frac{dX_i}{dt} = \sum_{j=1}^{p} \mu_{ij} V_j, \quad i = 1, 2, \cdots, n \tag{1}$$

where X_i are the metabolite concentrations, V_j denote the reaction rates, and the parameters $\mu_{ij} \in R$ are the stoichiometric coefficients of the metabolite concentrations X_i in the reactions V_j. The reaction rates V_j can be expressed as the following mathematical functions:

$$V_j = U_j(X), \quad j = 1, 2, \cdots, p \tag{2}$$

where $X = (X_1, X_2, \cdots, X_{n+m})^T \in R^{n+m}$ and X_k $(k = n+1, n+2, \cdots, n+m)$ are the m external metabolites (enzyme activities, fixed extracellular concentrations, kinetic parameters).

In the GMA version of a biochemical system, the rate expression functions $U_j(X)$ in Eq. (2) adopt the following power-law form (Savageau 1969a,b; Voit, 2000):

$$U_j(X) = \gamma_j \prod_{k=1}^{n+m} X_k^{g_{jk}}, \quad j = 1, 2, \cdots, p \tag{3}$$

In this representation, the model parameters $\gamma_j > 0$ are the rate constants for the reaction rates V_j, and $g_{jk} \in R$ are the kinetic orders that reflect the direct effects of system variables X_k on reaction rates V_j. The model parameters γ_j and g_{jk} can be defined respectively as (Savageau 1969a,b, 1976; Voit, 2000):

$$\gamma_j = \left(V_j\right)_0 \prod_{k=1}^{n+m} (X_k)_0^{-g_{jk}} \tag{4}$$

$$g_{jk} = \left(\frac{\partial V_j}{\partial X_k} \frac{X_k}{V_j}\right)_0 \tag{5}$$

where the subscript 0 indicates that the results are calculated at the steady-state of metabolite concentrations X_i $(i = 1, 2, \cdots, n)$.