Computing dynamically equivalent realizations of biochemical reaction networks with mass conservation

János Rudan[*], Gábor Szederkényi[*,†] and Katalin M. Hangos[†,**]

[*]*Faculty of Information Technology, Pázmány Péter Catholic University, Práter u. 50/a, H-1083 Budapest, Hungary*
[†]*Process Control Research Group, Computer and Automation Research Institute, Hungarian Academy of Sciences. Kende u. 13-17, H-1111 Budapest, Hungary*
[**]*Dept. of Electrical Engineering and Information Systems, Faculty of Information Technology, University of Pannonia, Egyetem u. 10, H-8200 Veszprém, Hungary*

Abstract. An algorithm for the computation of mass conservative dynamically equivalent chemical reaction network structures is proposed in this paper. The algorithm is formulated in an optimization-based framework as a mixed-integer linear programming problem.

INTRODUCTION

Biochemical reaction networks (CRNs) obeying the mass action law form an interesting and important class of nonnegative systems having the property that the nonnegative orthant is invariant for the system dynamics. CRNs are able to produce all important dynamical phenomena like stable, unstable or multiple equilibria, limit cycles or even chaotic behavior [1]. However, their relatively simple algebraic (matrix) structure allows us to use effective computation tools for system analysis and control.

It has been known for long that the dynamical behavior of structurally/parametrically different reaction networks give rise to the same kinetic ODEs. This phenomenon is usually called *macro-equivalence* [2] or *dynamical equivalence*. It was shown in [3, 4, 5, 6] that important properties like the number of reactions, (weak) reversibility, detailed or complex balance and deficiency are realization-specific properties. Moreover, the computation of such CRN structures can be handled in the framework of mixed integer linear programming. Linear conjugacy [7] can be considered as a kind of extension of dynamical equivalence allowing diagonal state transformations between the CRNs.

During the mathematical generalization of mass-action CRN systems the mass conservation assumptions which are present in classical chemical theory were usually omitted. Despite of this, mass-conservative CRNs formulate an important system class describing physically plausible chemical reaction sets [8].

The purpose of this paper is to show that mass-conservation is also a realization property for CRNs, and to propose a computational procedure for computing dynamically equivalent mass-conservative structures.

DEFINITIONS AND SYSTEM DESCRIPTION

Through this paper, we consider CRNs which obey the mass-action law. CRNs are characterized based on [9] where a CRN containing *n species* and *m complexes* is uniquely described by its Kirchoff matrix $A_k \in \mathbb{R}^{m \times m}$ and complex composition matrix $Y \in \mathbb{Z}_+^{n \times m}$. Complexes $C_1, \ldots, C_m$ are formulated from the species (or chemical substances) $X_1, \ldots, X_n$ as $C_j = \sum_{i=1}^{n} \alpha_{ij} X_i$, $j = 1, \ldots, m$, where α_{ij} are called as *stoichiometric coefficients*. The jth column of Y describes the composition of the complex C_j, namely $Y_{ij} = \alpha_{ij}$.

The set of *reactions* is formulated from ordered pairs of complexes (C_i, C_j) where C_i is transformed to C_j in the CRN. This relation will be denoted as $C_i \rightarrow C_j$. For each reaction a *reaction rate coefficient* is assigned, which is a nonnegative value denoted by k_{ij}. If $k_{ij} = 0$, the reaction $C_i \rightarrow C_j$ is not present in the CRN.

The matrix A_k is a column-conservation matrix (i.e. the sum of the values in each column is zero) containing the reaction rate coefficients in its off-diagonal elements

$$[A_k]_{ij} = \begin{cases} -\sum_{l=1,l\neq i}^{m} k_{il}, & \text{if} \quad i = j \\ k_{ji}, & \text{if} \quad i \neq j. \end{cases} \tag{1}$$

In this work a computationally advantageous factorization of the kinetic ODEs [9] is used which describes the time-evolution of species concentration as follows:

$$\dot{x} = M \cdot \Psi(x) = Y \cdot A_k \cdot \Psi(x), \tag{2}$$

where $x = [X] \in \mathbb{R}^n$ is the concentration vector of the species X_i, $i = 1, \ldots, n$, M is a given constant matrix, and $\Psi(x)$ is a monomial-type vector mapping defined by

$$\psi_j(x) = \prod_{i=1}^{n} x_i^{Y_{ij}}, \; j = 1, \ldots, m. \tag{3}$$

Let us consider two CRNs with the same M, having the same complex set described by Y but different Kirchoff matrices, namely $A_k^{(1)}$ and $A_k^{(2)}$. Then these systems are said to be dynamically equivalent if and only if the following equation holds:

$$Y \cdot A_k^{(1)} = Y \cdot A_k^{(2)}. \tag{4}$$

A weighted directed graph $\mathscr{D} = (\mathscr{V}, \mathscr{E})$ can be assigned to each CRN, where the finite nonempty set $\mathscr{V}$ denotes the vertices and the finite set $\mathscr{E}$ consists of ordered pairs of distinct vertices representing the edges. The vertices are the complexes from the reaction network: $\mathscr{V} = \{C_1, \ldots, C_m\}$. The directed edges represent the reactions, i.e. $(C_i, C_j) \in \mathscr{E}$ if reaction $C_i \rightarrow C_j$ occurs in the network. The reaction rates k_{ij} appear as nonnegative weights on the edges in the directed graph.

In [8] (stoichiometric) *mass conservation* is formulated as follows. Let us define g_v as the scaled molecular weight of the species X_v with strictly positive value. If reaction $C_i \rightarrow C_j$ is present in the network, the following can be written:

$$\sum_{v=1}^{n} \alpha_{vi} g_v = \sum_{v=1}^{n} \alpha_{vj} g_v = c_s \tag{5}$$

where $c_s > 0$ is a constant column-sum. Let us define vector $g \in \mathbb{R}_+^n$ as a row vector formulated from the scaled molecular weights. Now eq. (5) can be rewritten as $g \cdot Y(\cdot, i) = g \cdot Y(\cdot, j) = c_s$ where $Y(\cdot, i)$ refers to the ith column of matrix Y. Finally, it can be said that a reaction is mass conservative if the following holds:

$$g \cdot \rho^{(i,j)} = 0 \tag{6}$$

where $\rho^{(i,j)} = Y(\cdot, i) - Y(\cdot, j)$ is the *reaction vector* and g is strictly positive. In [8] the *mass conservative reaction set* is defined as a set of the reactions having the above property. It should be noted that a given CRN is called mass conservative if all of its reactions are in the mass conservative reaction set and a common g can be determined.

In the sequel, we will define an algorithm to compute dynamically equivalent realizations which are mass conservative in the above defined sense.

COMPUTING DYNAMICALLY EQUIVALENT, MASS-CONSERVATIVE REALIZATIONS USING OPTIMIZATION

In [3] an optimization-based framework is proposed to search for possible CRN structures with prescribed properties. The computation of the realizations leads to a mixed integer linear programming (MILP) problem. In this work, we introduce additional constraints into the optimization problem to ensure that the resulting realizations, besides other requirements, also fulfil the mass-conservation property.

Recall from [3] that the linear constraints corresponding to the mass action dynamics and the Kirchhoff property can be formulated as follows:

$$Y \cdot A_k = M \tag{7}$$

$$\textstyle\sum_{i=1}^{m} [A_k]_{ij} = 0, \; j = 1, \ldots, m \tag{8}$$

where the decision variables are the off-diagonal elements of A_k and the following bounds are introduced: $[A_k]_{ij} \geq 0$ for $i, j = 1, \ldots, m$, $i \neq j$ and $[A_k]_{ii} \leq 0$ for $i = 1, \ldots, m$.

Considering eq. (6) in order to ensure the mass-conservation property, a reaction between complex i and j can be present if and only if $g \cdot \rho^{(i,j)} = 0$ where the elements of g are in the optimization vector. This is formulated for every (i, j) pair $i, j = 1, \ldots, m$, $i \neq j$ in the following way:

$$g \cdot \rho^{(i,j)} = 0 \;\; \Rightarrow \;\; [A_k]_{ij} + [A_k]_{ji} \geq 0 \tag{9}$$

where $\Rightarrow$ stands for the logical implication operation. To keep track that which complexes are able to react with each other (while preserving the mass-conservation) we introduced an auxiliary binary valued decision variable δ_l, $l = 1, \ldots, m \cdot (m-1)/2$ for each constraint in the following way:

$$g \cdot \rho^{(i,j)} = 0 \;\; \Rightarrow \;\; \delta_l = 0, \quad g \cdot \rho^{(i,j)} \neq 0 \;\; \Rightarrow \;\; \delta_l = 1 \tag{10}$$
$$\delta_l = 0 \;\; \Rightarrow \;\; [A_k]_{ij} + [A_k]_{ji} \geq 0 \tag{11}$$
$$\delta_l = 1 \;\; \Rightarrow \;\; [A_k]_{ij} + [A_k]_{ji} = 0 \tag{12}$$

where eq. (11)-(12) control the reaction graph structure: if reaction $C_i \to C_j$ is forbidden then $[A_k]_{ij}$ is forced to be 0, otherwise it has a non-negative value. After proper reformulation, these constraints can be incorporated into the MILP problem [10].

The proper bounds for the elements of g is introduced to ensure that g is a strictly positive vector. Also, if the problem in question dictates it, it is possible to exclude isomers (i.e. species with the same scaled molecular weights) by adding the following constraint: $g(i) \neq g(j)$, $i, j = 1, \ldots, n$, $i \neq j$.

Finally, a CRN which fulfills all the constraints is obtained by evaluating the following objective function with respect to the above listed constraints:

$$F([A_k]_{ij}) = \sum_{i,j,\ i \neq j} [A_k]_{ij} \tag{13}$$

where the optimization vector consists of the elements of matrix A_k and vector g as real-valued variables, and $m \cdot (m-1)/2$ binary valued variables obtained from (9)-(12). It should be noted that other objective functions can lead to other realizations having further special properties, e.g. minimal/maximal number of reactions etc.

The size of the emerging MILP problem is mainly determined by the number of appearing auxiliary binary variables which has the magnitude $\mathcal{O}(m^2)$.

EXAMPLE

Let us consider the following kinetic ODEs:

$$\dot{x}_1 = -2x_1^2 x_2 + 4x_2^2 - 2x_1^4$$
$$\dot{x}_2 = x_1^2 x_2 - 2x_2^2 + x_1^4 \tag{14}$$

The so-called 'canonical' reaction graph of eq. (14) can be formulated with the help of the algorithm described in [11]. The resulting reaction network is depicted in Fig. 1(a) where the complex composition matrix is the following:

$$Y = \begin{bmatrix} 2 & 1 & 0 & 1 & 4 & 3 & 2 & 0 & 4 \\ 1 & 1 & 2 & 2 & 0 & 0 & 2 & 1 & 1 \end{bmatrix} \tag{15}$$

The non-zero off-diagonal elements of the matrix A_k describing the reaction graph are the following: $A_k(2,1) = 2$, $A_k(7,1) = 1$, $A_k(4,3) = 4$, $A_k(8,3) = 2$, $A_k(6,5) = 2$, $A_k(9,5) = 1$. Based on eq. (2) the algebraic structure of the system can be characterized by the following matrix:

$$Y \cdot A_k = \begin{bmatrix} -2 & 0 & 4 & 0 & -2 & 0 & 0 & 0 & 0 \\ 1 & 0 & -2 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

It can be seen that this realization is not mass conservative because in each linkage class at least one semi-positive reaction vector can be found.

By applying the proposed algorithm with the maximization of the cost function, a dynamically equivalent reaction network was found, which fulfills the mass-conservation property. The resulting reaction graph is presented in Fig. 1(b). Isolated complexes (i.e. those complexes which do not react with any other) were omitted. By looking at the reaction vectors in the reaction graph and the obtained vector of scaled molecular weights $g = [1\ 2]$, the presence of the mass-conservation property can be seen. The resulting reaction graph is described by matrix A_k^{mc} in which the non-zero off-diagonal elements are the following: $A_k^{mc}(3,1) = 1$, $A_k^{mc}(1,3) = 2$, $A_k^{mc}(1,5) = 1$. For this system the complex composition matrix Y is still the one from eq. (15). It can be seen that $Y \cdot A_k = Y \cdot A_k^{mc}$. Therefore the two systems are dynamically equivalent based on eq. (4).
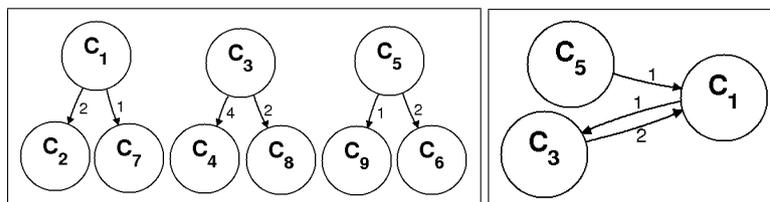
FIGURE 1. **(a)** Canonical realization of eq. (14). Elements of A_k appear as edge weights. **(b)** A mass conservative realization of eq. (14). Elements of A_k^{mc} appear as edge weights. In both subfigures, node numbers refer to the columns of Y from eq. (15).

SUMMARY AND FUTURE WORK

A new MILP-based algorithm was proposed to compute alternative, dynamically equivalent and mass conservative realizations of CRNs. This method gives us the possibility to analyze the phenomenon of mass conservation in reaction networks having different structural and dynamical properties. A large set of network properties can be investigated with the help of the presented method, too, e.g. linear conjugate CRNs with mass conservation, structural/parametrical uniqueness of mass conservative networks, effect of mass conservation on the deficiency etc.

ACKNOWLEDGMENTS

REFERENCES

1. P. Érdi, and J. Tóth, *Mathematical Models of Chemical Reactions. Theory and Applications of Deterministic and Stochastic Models*, Manchester University Press, Princeton University Press, Manchester, Princeton, 1989.
2. F. Horn, and R. Jackson, *Archive for Rational Mechanics and Analysis* **47**, 81–116 (1972).
3. G. Szederkényi, *Journal of Mathematical Chemistry* **47**, 551–568 (2010).
4. G. Szederkényi, and K. M. Hangos, *Journal of Mathematical Chemistry* **49**, 1163–1179 (2011).
5. G. Szederkényi, K. M. Hangos, and T. Péni, *MATCH Commun. Math. Comput. Chem.* **65**, 309–332 (2011).
6. G. Szederkényi, K. M. Hangos, and Z. Tuza, *MATCH Commun. Math. Comput. Chem.* **67**, 193–212 (2012).
7. M. D. Johnston, D. Siegel, and G. Szederkényi, *Journal of Mathematical Chemistry* **50**, 274–288 (2012).
8. K. M. Hangos, and G. Szederkényi, "The effect of conservation on the dynamics of chemical reaction networks," in *IFAC Workshop on Thermodynamic Foundations of Mathematical Systems Theory, TFMST'13*, 2013.
9. M. Feinberg, *Lectures on chemical reaction networks*, Notes of lectures given at the Mathematics Research Center, University of Wisconsin, 1979.
10. S. P. Bradley, A. C. Hax, and T. L. Magnanti, *Applied Mathematical Programming*, Addison-Wesley, 1977.
11. V. Hárs, and J. Tóth, "On the inverse problem of reaction kinetics," in *Qualitative Theory of Differential Equations*, edited by M. Farkas, and L. Hatvani, North-Holland, Amsterdam, 1981, vol. 30 of *Coll. Math. Soc. J. Bolyai*, pp. 363–379.