Model reduction of detailed-balanced reaction networks by clustering linkage classes

Rao, Shodhan; Jayawardhana, Bayu; van der Schaft, Abraham

Published in:
Proc. of the 6th IFAC Conference on Foundations of Systems Biology in Engineering

Publication date:
2016

Model reduction of detailed-balanced reaction networks by clustering linkage classes

Shodhan Rao * **Bayu Jayawardhana** ** **Arjan van der Schaft** ***

* *Ghent University Global Campus, 187 Songdo-dong, Yeonsu-Gu, Incheon, South Korea (e-mail: shodhan.rao@ghent.ac.kr).*
** *ENTEG, Faculty of Mathematics and Natural Sciences, University of Groningen, The Netherlands (e-mail: b.jayawardhana@rug.nl)*
*** *Johann Bernoulli Institute of Mathematics and Natural Sciences, University of Groningen, The Netherlands, (e-mail: a.j.van.der.schaft@rug.nl)*

Abstract: We propose a model reduction method that involves sequential application of clustering of linkage classes and Kron reduction. This approach is specifically useful for chemical reaction networks with each linkage class having less number of reactions. In case of detailed balanced chemical reaction networks that are governed by general enzyme kinetics, we show that our procedure ensures that the space of equilibria corresponding to the original model is a subspace of the space of equilibria of the reduced model.

Keywords: Biochemical reaction networks, Model reduction, Graph theory, Network dynamics.

1. INTRODUCTION

An important topic in systems biology is the study of large-scale models of biochemical reaction networks that describe the dynamics of metabolism, signaling, gene expression, etc. in living cells. Biochemical reaction networks usually involve many enzyme-catalyzed processes governed by nonlinear enzyme-kinetic rate laws. Owing to intricate stoichiometry and regulatory mechanisms, the enzyme-catalyzed processes of a network are highly interdependent. Consequently, the complexity of the models describing their dynamics is daunting, even if we make simplifying assumptions of uniform spatial distribution and constant system temperature and pressure. A deterministic model of such a reaction network may contain high-dimensional sets (typically of the order of 100) of coupled polynomial or rational ordinary differential equations, which sometimes require huge computational effort to analyze.

There is a growing need of mathematical tools for analyzing complex models of reaction networks and for obtaining biological insights from the analysis. The computational effort required to analyze a mathematical model of biochemical reaction network is considerably reduced if we can obtain a reduced model that mimics the behaviour of the original model satisfactorily, but contains less differential equations and parameters as compared to the original model.

Model reduction of biochemical reaction networks has been a very active area of research in the past few years and a lot of complementary approaches have been devised for the purpose. The reader is referred to [14] for a detailed exposition of some well known model reduction procedures. Here we list only a few of the known methods. The most commonly used techniques are the singular perturbation method, the time-scale separation technique [6,12,15,17,20], the rapid-equilibrium approximation, also known as the quasi-equilibrium approximation (see [18]) and the quasi steady-state approximation (QSSA) (see for e.g., [19]). Model reduction can also be carried out by reducing the number of parameters (e.g. [2,10–12]) or the number of reactions (e.g. [1,3,5,13]). Some of these model reduction methods are based on a priori experimental information and/or biological knowledge and hence are only locally effective.

The authors of the current paper have previously proposed a model reduction method [16] that overcomes most of the limitations of the other well known model reduction approaches. The method proposed in [16] proceeds by a simple stepwise reduction in the number of 'complexes', which are the combinations of species on the left and right-hand sides of the various reactions in the network. The method is based on the reduction of the underlying weighted Laplacian (see [4] for a definition) describing the structure of the graph of complexes which is a graph with vertices corresponding to the complexes and edges corresponding to the reactions in the network. The effect of this stepwise reduction is monitored by an error integral, which quantifies how much the behaviour of the reduced model deviates from the original. The procedure ensures that the set of complexes that are deleted is in some sense the optimal set for which the desired level of closeness between the transient behaviors of the original and the reduced model is obtained. This method is inspired by Kron reduction method [9] which is a popular method of reduction of resistive electrical network models.

Listed below are the advantages of the method proposed in [16] that cannot all be found in one particular model reduction method that was listed earlier. The model reduction technique of [16] is easy to implement and can be automated. It does not rely on a priori knowledge about the experimental conditions or biological function of the network. Furthermore, the reduced model largely retains the kinetics and structure of the original model. This enables a direct biochemical interpretation and yields insight into which parts of the network have the highest influence on its behaviour. It also accelerates computations by reducing computational effort, especially when we deal with models of huge biochemical reaction networks. For networks governed by mass action kinetics, an improvement of the model reduction approach [16] in terms of the computational effort involved has been proposed recently (see [7]).

The model reduction approach of [16] has one serious limitation which we explain below. Assume that every connected component (*linkage class* in chemical reaction network (CRN) terminology) of the graph of complexes corresponding to a network has only two complexes. In this case, deletion of any complex by the approach of [16] simply results in the deletion of the corresponding linkage class. For example assuming that $X_1, \ldots, X_{10}$ are distinct chemical species, consider the following reaction network with 3 linkage classes.

$$\left.\begin{array}{l} X_1 + X_2 \rightleftharpoons X_3 + X_4 \\ X_4 + X_5 \rightleftharpoons X_6 + X_7 \\ X_7 + X_8 \rightleftharpoons X_9 + X_{10} \end{array}\right\} \tag{1}$$

Deletion of the complex $X_4 + X_5$ results in elimination of the second linkage class. Thus the reduced model will have the first reaction occuring independently of the third reaction and consequently the reduced model will exhibit a dynamic behaviour which is far from that of the original model. Most biochemical networks in real life fall under the category of reaction networks with each linkage class having only two complexes and therefore there is a need to develop a more effective graph-theory based approach for model reduction of biochemical reaction networks. In this paper, we propose a method to overcome this particular limitation. Each step of this method involves clustering two linkage classes and deletion of complexes in the clustered linkage class. The clustering of linkage classes at every step ensures that the resulting linkage class has at least three complexes, so that deletion of a single complex from it does not result in the elimination of the linkage class.

We further consider the application of our model reduction procedure for detailed balanced networks governed by general enzyme-kinetic rate laws. The reader is referred to [8] for a description of properties of such networks. Thermodynamically the assumption of detailed-balancedness for any network without interactions with the external environment is well justified as it corresponds to microscopic reversibility. We show that for such networks, the clustering procedure described in this paper ensures that the space of equilibria of the original network is equivalent to that of the clustered network. As a consequence to this, the sequential application of clustering of complexes and Kron reduction will result into a reduced model where the space of equilibria of the original network is a subspace of that corresponding to the reduced network. This set inclusion property is due to the property of Kron reduction method applied to detailed-balanced networks [21].

Notation: The space of n dimensional real vectors is denoted by $\mathbb{R}^n$ and the space of n dimensional real vectors consisting of all strictly positive entries is denoted by $\mathbb{R}^n_+$. Define the mapping $\mathrm{Ln} : \mathbb{R}^m_+ \to \mathbb{R}^m, \quad x \mapsto \mathrm{Ln}(x)$, as the mapping whose i-th component is given as $(\mathrm{Ln}(x))_i := \ln(x_i)$.

2. PRELIMINARIES

Let us recall the general description of (bio)chemical reaction networks as recently reviewed in [22]. In this paper, we consider reversible (bio)chemical reaction networks that do not interact with the external environment, meaning that there are no inflows from and no outflows to the external environment. Assume that there are m species, c complexes and r reversible reactions in such a network. Define the *complex composition matrix* Z as the $m \times c$ matrix whose α-th column captures the expression of the α-th complex in the m chemical species. For example consider the reversible reaction network given by

$$X_1 + 3X_2 \rightleftharpoons X_3 \rightleftharpoons 2X_1 + 2X_2 \tag{2}$$

For the above reaction network, the complex composition matrix is given by

$$Z = \begin{bmatrix} 1 & 0 & 2 \\ 3 & 0 & 2 \\ 0 & 1 & 0 \end{bmatrix}$$

Note that all the entries of the complex composition matrix are nonnegative integers. We define as in [21] the *graph of complexes* as a directed graph whose nodes are the complexes and whose edges correspond to the reversible reactions of the network. The direction of each edge is arbitrary, usually given by the direction of the forward reaction of the network. Note that the direction of each edge of the graph of complexes does not affect the modelling or the model reduction procedure described in this paper. One can associate an *incidence matrix* B with the graph of complexes. This is a matrix of dimension $c \times r$. The j^{th} column of B corresponds to the j^{th} edge of the graph of complexes and has entries -1 corresponding to the tail vertex, $+1$ corresponding the head vertex and the remaining entries are all 0. Observe that an incidence matrix B corresponding to the reaction network (2) is

$$B = \begin{bmatrix} -1 & 0 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix}$$

For simplifying the description of our approach later, let us introduce further the following notations. For a given reaction edge i, we denote a substrate complex (which is assumed to be the tail vertex of i-th edge) by $\mathcal{S}_i$ and a product complex (which corresponds to the head vertex) by $\mathcal{P}_i$. Correspondingly, $Z_{\mathcal{S}_i}$ and $Z_{\mathcal{P}_i}$ denote the column of Z that is related to the complex $\mathcal{S}_i$ and $\mathcal{P}_i$, respectively.

The connected components of a chemical reaction network are called its *linkage classes*. Assume that a given reversible chemical reaction network has ℓ linkage classes. Then its complex composition matrix Z and incidence

matrix B can be partitioned according to its linkage classes as follows.

$$Z = [Z_1 \ Z_2 \ \dots \ Z_\ell] \tag{3}$$

$$B = \begin{bmatrix} B_1 & 0 & 0 & \dots & 0 \\ 0 & B_2 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & 0 & B_{\ell-1} & 0 \\ 0 & \dots & \dots & 0 & B_\ell \end{bmatrix} \tag{4}$$

We denote by $x \in \mathbb{R}^m_+$, the vector of concentrations of chemical species participating in the given network of reactions and by $v \in \mathbb{R}^r$ the vector of reaction rates of the network. The dynamics of a chemical reaction network is described by the equation $\dot{x} = ZBv$, where v is a function of x that depends on the governing laws of the reactions in the network.

A vector of concentrations $x^* \in \mathbb{R}^m_+$ is called a *thermodynamic equilibrium* for the dynamics $\dot{x} = ZBv(x)$ if $v(x^*) = 0$. At a thermodynamic equilibrium, every reversible reaction of the network is at equilibrium. A chemical reaction network with dynamics given by $\dot{x} = ZBv(x)$ is called *detailed-balanced* if there exists a thermodynamic equilibrium $x^* \in \mathbb{R}^m_+$. It can be shown that every reaction of a detailed-balanced reaction network is necessarily reversible (see [21, Proposition 3.2] for a proof).

In this paper, we mainly consider detailed-balanced biochemical reaction networks that are governed by general enzyme-kinetic rate laws. The reaction rate v_j of the j^{th} reaction of such a network is given by

$$v_j(x) = d_j(x)\left(k_j^{\text{forw}} \exp\left(Z_{\mathcal{S}_j}^T \text{Ln}(x)\right) - k_j^{\text{rev}} \exp\left(Z_{\mathcal{P}_j}^T \text{Ln}(x)\right)\right), \tag{5}$$

where for $j = 1, \dots, r$, $d_j : \mathbb{R}^m_+ \to \mathbb{R}_+$ is a rational function of its argument in case that the governing law for the j^{th} reaction is Michelis-Menten-type kinetics and it is equal to 1 in case the governing law is mass action kinetics. We remark that in practice the function d_j usually depends on the composition of the substrate and the product complex of the j^{th} reaction. For other properties of such reaction networks, we refer interested readers to [8].

3. CLUSTERING LINKAGE CLASSES

Recall as mentioned in the introduction that every step of our model reduction procedure consists of clustering two linkage classes and then deleting complexes from the resulting linkage class. In this section, we explain in detail how the clustering is carried out.

Without loss of generality, let us discuss our general method for clustering two linkage classes by combining some of the complexes from the two linkage classes. These linkage classes may share common chemical species in two different complexes occuring in these linkage classes. Consider two different complexes C_1 and C_2 from each linkage class that will be combined. Associated to these complexes, we denote Z_{C_1} and Z_{C_2} as the complex composition vectors corresponding to C_1 and C_2, respectively. Let us introduce a new complex C_{12} whose elements contain all species in C_1 and C_2, by allocating additional complex composition vectors z_1 and z_2 with non-negative integer entries such that the following relation holds

$$Z_{C_1} + z_1 = Z_{C_2} + z_2 =: Z_{C_{12}}$$

As an example, let us recall the chemical reaction network in the Introduction in (1). We can combine the complexes $X_3 + X_4$ and $X_4 + X_5$ to form a new complex C_{12} through infinite number of possibilities. For instance, we can add the species X_5 to the first complex and, *vice versa*, the species X_3 to the second complex, where in this case,

$$z_1 = [0\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 0]^T$$
$$z_2 = [0\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0]^T,$$

and thus,

$$Z_{C_{12}} = [0\ 0\ 1\ 1\ 1\ 0\ 0\ 0\ 0\ 0]^T.$$

In the above example, we exploit the fact that there is a common species X_4 in both complexes and what we need to add to each complex is the remaining species from the other complex. Another possibility is by incorporating all species from the other complex. If we consider again the previous example, the vectors z_1 and z_2 are given by

$$z_1 = Z_{C_2} = [0\ 0\ 0\ 1\ 1\ 0\ 0\ 0\ 0\ 0]^T$$
$$z_2 = Z_{C_1} = [0\ 0\ 1\ 1\ 0\ 0\ 0\ 0\ 0\ 0]^T,$$

and the combined complex composition vector $Z_{C_{12}}$ becomes

$$Z_{C_{12}} = [0\ 0\ 1\ 2\ 1\ 0\ 0\ 0\ 0\ 0]^T.$$

Since the choice of z_1 and z_2 is arbitrary, one should determine these vectors in such a way that some of the dynamical properties of the original model can be retained.

Assume that the original network has ℓ linkage classes and its complex composition matrix Z and its incidence matrix B has a partition corresponding to its linkage classes given by equations (3) and (4) respectively. Assume without loss of generality that the linkage classes that are being clustered are the first two linkage classes whose complex composition matrices are given by Z_1 and Z_2. As before, we identify the complexes C_1 and C_2 as the complexes-to-be-combined from the first and second linkage classes, respectively. Furthermore, for simplicity of presentation, we can perform a permutation in the ordering of complexes from each linkage class such that C_1 is the last complex in the first linkage class and C_2 is the first complex in the second linkage class.

Now consider a partition of Z_1 as follows

$$Z_1 = \big[\ \underbrace{Z_{r_1}}_{\mathbb{R}^{m\times(c_1-1)}} \quad \underbrace{Z_{C_1}}_{\mathbb{R}^{m\times 1}}\ \big]$$

where c_1 is the number of complexes in the first linkage class. Similarly consider the partition of Z_2 given by

$$Z_2 = \big[\ \underbrace{Z_{C_2}}_{\mathbb{R}^{m\times 1}} \quad \underbrace{Z_{r_2}}_{\mathbb{R}^{m\times(c_2-1)}}\ \big]$$

where c_2 is the number of complexes in the second linkage class.

In the following, we discuss the construction of a new complex composition matrix $\hat{Z}$ from the original Z such that we are able to preserve some of the equilibrium properties.

Let V_1 denote a matrix with m rows and the same number of columns as Z_{r_1}, each of whose columns is equal to z_1. Let

V_2 denote a matrix with m rows and the same number of columns as Z_{r_2}, each of whose columns is equal to z_2. Then the new complex composition matrix $\hat{Z}$ corresponding to the clustered network is

$$\hat{Z} = [Z_{r_1} + V_1 \;\; Z_{C_{12}} \;\; Z_{r_2} + V_2 \;\; Z_3 \;\; Z_4 \;\; \dots \;\; Z_\ell], \qquad (6)$$

where $Z_3, Z_4, \dots Z_\ell$ are the complex composition sub matrices corresponding to the linkage class $3, 4, \dots \ell$, respectively. One can check easily that the clustered network will have $\ell - 1$ linkage classes since the first three entries of the right hand side of the above equation correspond to one linkage class.

We thus add the complex with the composition vector z_1 to every complex in the first linkage class and add the complex with the composition vector z_2 to every complex in the second linkage class. Thus every complex in the first two linkage classes are modified. We show later that in case of a special category of reaction networks called detailed balanced networks which are necessarily reversible, such a modification does not alter the equilibrium properties of the network.

We show now how the incidence matrix corresponding to the graph of complexes is modified by clustering. Consider partitions of B_1 and B_2 given by

$$B_1 = \begin{bmatrix} B_{r_1} \\ B_{C_1} \end{bmatrix} \qquad B_2 = \begin{bmatrix} B_{C_2} \\ B_{r_2} \end{bmatrix}$$

where B_{r_1} denotes the rows of B_1 corresponding to all the complexes of the first linkage class apart from C_1 and similarly B_{r_2} denotes all the rows of B_2 corresponding to all the complexes of the second linkage class apart from C_2. Define

$$\hat{B}_{12} = \begin{bmatrix} B_{r_1} & 0 \\ B_{C_1} & B_{C_2} \\ 0 & B_{r_2} \end{bmatrix}$$

The matrix $\hat{B}_{12}$ defined above represents the new incidence matrix corresponding to the clustered linkage class. The new incidence matrix $\hat{B}$ for the clustered network is given by

$$\hat{B} = \begin{bmatrix} \hat{B}_{12} & 0 & 0 & \dots & 0 \\ 0 & B_3 & 0 & \dots & 0 \\ 0 & 0 & B_4 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & B_\ell \end{bmatrix} \qquad (7)$$

(c.f., the original incidence matrix as in (4)).

Thus the vertices corresponding to the complexes C_1 and C_2 in the original graph of complexes are merged. All the remaining vertices and edges are unchanged. The clustering of the two linkage classes modifies only the complexes belonging to the two linkage classes. It does not change the number, the rate governing laws and the rate constants of the reactions of the network. Since the clustered linkage class will have at least 3 complexes, we can apply Kron reduction [16] to possibly reduce the order of the model. Note that it does not make sense to cluster complexes that do not share common species because the subsequent Kron reduction will then not result in the reduction of the model-order. The following example illustrates the application of our clustering procedure followed by application of Kron reduction in order to reduce the model-order.

Example 1. Consider the following reversible mass action reaction network consisting of two linkage classes

$$X_1 + X_2 \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} X_3 + 2X_4$$
$$X_3 + X_5 \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} X_6 \qquad (8)$$

where $k_1, k_{-1}, k_2, k_{-2} \in \mathbb{R}_+$ are called reaction constants of the network. Notice that the complexes $X_3 + 2X_4$ and $X_3 + X_5$ have one common species X_3 and these can be clustered by the procedure described in the paper. Taking

$$z_1 = [0 \;\; 0 \;\; 0 \;\; 0 \;\; 1 \;\; 0]^T \qquad \text{and}$$
$$z_2 = [0 \;\; 0 \;\; 0 \;\; 2 \;\; 0 \;\; 0]^T,$$

we obtain the following clustered network

$$X_1 + X_2 + X_5 \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} X_3 + 2X_4 + X_5 \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} 2X_4 + X_6 \qquad (9)$$

If we now delete the complex $X_3 + 2X_4 + X_5$ from the above network by Kron reduction as explained in [16], we get

$$X_1 + X_2 + X_5 \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} 2X_4 + X_6 \qquad (10)$$

where $k_3, k_{-3} \in \mathbb{R}_+$ are the reaction constants of the reduced network. It is easy to see that (10) is a reduced order network since the dimension of x is reduced from 6 to 5.

4. DETAILED BALANCED NETWORKS GOVERNED BY GENERAL KINETICS

Any model reduction procedure is supposed to preserve key properties of the network. In the case of Kron reduction procedure [16], it is known that if the governing rate laws of all the reactions of the network are the same, then the reduced network also has the same governing rate law. Furthermore, in the case of detailed balanced biochemical reaction networks governed by general enzyme kinetics, it is known ([8], [21, Proposition 6.3]) that although the space of equilibria expands upon model reduction, it always contains the space of equilibria of the original network. We now study this important class of networks and show that the clustering of linkage classes as described in the previous section completely preserves the space of equilibria for this class of networks, i.e., we show that the space of equilibria is unaltered upon clustering..

Consider now the detailed-balanced network governed by general kinetics as discussed in section 2. Let the first r_1 reactions be the reactions in linkage class 1 and the next $r_2 - r_1$ reactions be those in linkage class 2 and we are clustering linkage classes 1 and 2. If $\hat{v}_j$ denotes the rate of the j^{th} reaction of the clustered network, then it is easy to see that $\hat{v}_j = v_j$ for $j = r_2 + 1, r_2 + 2, \dots, r$, i.e., the reaction rates of reactions in linkage classes 3 onward remain the same after clustering, since the complexes, the rate governing laws, the number and rate constants of the reactions in these linkage classes are not altered. Since the substrates and products of the first r_2 reactions are modified after clustering, for $j = 1, 2, \dots, r_2$, the function d_j can be modified to $\hat{d}_j$ which depends on the species involved in the clustered complexes. Since $\hat{d}_j$ does not influence the equilibrium points, it provides us with an

additional degree of freedom where $\hat{d}_j$ can be determined later to fit with the dynamic behavior of the original network. Using the complex clustering as described in the previous section and using $\hat{d}_j$, the rates for reactions in the first two linkage classes are given by

$$\hat{v}_j = \hat{d}_j(x)\left(k_j^{\text{forw}} \exp\left(\hat{Z}_{\mathcal{S}_j}^T \text{Ln}(x)\right) - k_j^{\text{rev}} \exp\left(\hat{Z}_{\mathcal{P}_j}^T \text{Ln}(x)\right)\right), \tag{11}$$

for all $j = 1 \dots r_2$ where $\hat{Z}$ is the clustered complex composition matrix as in (6).

Proposition 1. Consider a detailed-balanced reaction network governed by general kinetics whose dynamics is described by $\dot{x} = ZBv(x)$, with $v(x)$ denoting the reaction rates vector whose j^{th} component is given by equation (5). Assume that the first two linkage classes of this network are clustered resulting in the new dynamics

$$\dot{x} = \hat{Z}\hat{B}\hat{v}(x), \tag{12}$$

where $\hat{Z}$ and $\hat{B}$ are as in (6) and (7), respectively, and $\hat{v}_j$ is as in (11) for the first r_2 reactions and $\hat{v}_j(x) = v_j(x)$ otherwise. Then the clustered network is detailed balanced and the space of equilibrium points is the same for both the clustered and the original networks.

Proof. Let $\mathcal{E}$ and $\hat{\mathcal{E}}$ denote the space of equilibria of the original and the clustered networks. One of the results that We need to prove is that $\mathcal{E} = \hat{\mathcal{E}}$. Since the original network is detailed-balanced, every equilibrium is a thermodynamic equilibrium as proved in [21, Theorem 4.1]. Thus $\mathcal{E}$ is the set of solutions to $v(x) = 0$. Following equations (5), $\mathcal{E}$ is the set of solutions to

$$k_j^{\text{forw}} \exp\left(Z_{\mathcal{S}_j}^T \text{Ln}(x)\right) - k_j^{\text{rev}} \exp\left(Z_{\mathcal{P}_j}^T \text{Ln}(x)\right) = 0, \tag{13}$$

$j = 1, 2, \dots, r$.

From (6), it follows that for $j = 1, \dots, r_1$

$$\begin{aligned}\hat{v}_j(x) &= \hat{d}_j(x) k_j^{\text{forw}} \exp\left((Z_{\mathcal{S}_j} + z_1)^T \text{Ln}(x)\right) \\ &\quad - \hat{d}_j(x) k_j^{\text{rev}} \exp\left((Z_{\mathcal{P}_j} + z_1)^T \text{Ln}(x)\right) \\ &= \tilde{d}_j(x)\left(k_j^{\text{forw}} \exp\left(Z_{\mathcal{S}_j}^T \text{Ln}(x)\right)\right. \\ &\quad \left. - k_j^{\text{rev}} \exp\left(Z_{\mathcal{P}_j}^T \text{Ln}(x)\right)\right)\end{aligned}$$

where $\tilde{d}_j(x) := \hat{d}_j(x) \exp\left(z_1^T \text{Ln}(x)\right)$. Similarly, for $j = r_1 + 1, \dots, r_2$, it is easy to see that

$$\begin{aligned}\hat{v}_j(x) = \tilde{d}_j(x)&\left(k_j^{\text{forw}} \exp\left(Z_{\mathcal{S}_j}^T \text{Ln}(x)\right)\right. \\ &\left. - k_j^{\text{rev}} \exp\left(Z_{\mathcal{P}_j}^T \text{Ln}(x)\right)\right)\end{aligned}$$

where $\tilde{d}_j(x) := \hat{d}_j(x) \exp\left(z_2^T \text{Ln}(x)\right)$. Since $\hat{d}_j(x) > 0$ for $j = 1, \dots, r_2$, it follows that also $\tilde{d}_j(x) > 0$. Consequently, if $x^* \in \mathcal{E}$, then for $j = 1, \dots, r_1$

$$\begin{aligned}\hat{v}_j(x^*) &= \tilde{d}_j(x^*)\left(k_j^{\text{forw}} \exp\left(Z_{\mathcal{S}_j}^T \text{Ln}(x^*)\right)\right. \\ &\quad \left. - k_j^{\text{rev}} \exp\left(Z_{\mathcal{P}_j}^T \text{Ln}(x^*)\right)\right) \\ &= 0.\end{aligned}$$

Similarly, for $j = r_1 + 1, \dots, r_2$, we have that

$$\hat{v}_j(x^*) = 0$$

Thus $\hat{v}(x^*) = 0$. This implies that x^* is a thermodynamic equilibrium for the clustered network, which in turn implies that the clustered network is detailed-balanced. Hence every equilibrium of the clustered network is a thermodynamic equilibrium, in particular a solution of the set of equations (13). This implies that $\mathcal{E} = \hat{\mathcal{E}}$. □

The result in Proposition 1 shows that clustering preserves both the detailed-balancedness and the space of equilibria of the original network. Recall (see [21, Section 6]) that Kron reduction preserves the detailed-balancedness and retains the space of equilibria of the original network.

It should be noted that the clustering procedure described in this paper does not actually reduce the model order. It is useful in model-order reduction only if combined iteratively with Kron reduction since the latter procedure has the limitation that it can only effectively reduce reactions in the same linkage class. Since both approaches retain the space of equilibria from the original network that they start with, we can apply an iterative reduction procedure involving subsequent applications of clustering and Kron reduction methods and this will ensure that the space of equilibria of the original network will be retained in the final reduced network. The iterative reduction procedure will also provide effective means to reduce a given original detailed-balanced network to any desired extent, which is not the case if only Kron reduction is applied.

5. CONCLUSION AND FUTURE WORK

In this paper, we have proposed a novel model reduction method, which is based on an iterative procedure involving subsequent applications of clustering of chemical complexes and Kron reduction, for simplifying (bio)chemical reaction networks with general kinetics. In case, the original network is detailed-balanced, the proposed method can retain the space of equilibria from the original network and the resulting reduced network is also detailed-balanced and governed by general kinetics.

As discussed in Section 3, there are infinite number of ways of combining two complexes in two linkage classes in order to cluster these linkage classes. Currently, we are investigating how two complexes in separate linkage classes can be combined so as to have minimum effect on the overall dynamics of the network upon clustering. We are also evaluating the efficacy of the proposed method and the applicability of iterative procedure of clustering and Kron-based reduction to a number of curated models of biochemical networks.

REFERENCES

[1] I.P. Androulakis. Kinetic mechanism reduction based on an integer programming approach. *AIChE J*, volume 46, pages 361-371, 2000.

[2] M. Apri, M. de Gee, J. Molenaar. Complexity reduction preserving dynamical behaviour of biochemical networks. *J Theor Biol*, volume 304, pages 16-26, 2012.

[3] B. Bhattacharjee, D.A. Schwer, P.I. Barton, W.H. Green Optimally reduced kinetic models: reaction elimination in large-scale kinetic mechanisms. *Combustion and Flame*, volume 135, pages 191-208, 2003.

[4] B. Bollobas *Modern Graph Theory. Graduate Texts in Mathematics 184*. Springer, New York, 1998.
[5] B.L. Clarke. General method for simplifying chemical networks while preserving overall stoichiometry in reduced mechanisms. *J Chem Phys*, volume 97, pages 4066-4071, 1992.
[6] A.N. Gorban, I.V. Karlin. Method of invariant manifold for chemical kinetics. *Chem Eng Sci*, volume 58, pages 4751-4768, 2003.
[7] B. Jayawardhana, S. Rao, W. Sikkema, B. Bakker. Handling Biological complexity using Kron reduction. Mathematical Control Theory I, Lecture Notes in Control and Information Sciences, volume 461, pp. 73-93, 2015.
[8] B. Jayawardhana, S. Rao, A.J. van der Schaft. Balanced chemical reaction networks governed by general kinetics. 20th Int. Symp. on *Mathematical Theory of Networks and Systems* (MTNS), July 9-13, 2012, Melbourne, Australia.
[9] G. Kron. *Tensor Analysis of Networks*, Wiley, New York 1939.
[10] G. Liu, M.T. Swihart, S. Neelamegham. Sensitivity, principle component and flux analysis applied to signal transduction: the case of epidermal growth factor mediated signaling. *Bioinformatics*, volume 21, pages 1194-1202, 2005.
[11] M.R. Maurya, S.J. Bornheimer, V. Venkatasubramanian, S. Subramaniam. Reduced-order modelling of biochemical networks: application to the GTPase-cycle signalling module. *IET Syst Biol*, volume 152, pages 229-242, 2005.
[12] I.E. Nikerel, W.A. van Winden, P.J.T. Verheijen, J.J. Heijnen. Model reduction and *a priori* kinetic parameter identifiability analysis using metabolome time series for metabolic reaction networks with linlog kinetics. *Metab Eng*, volume 11, pages 20-30, 2009.
[13] L. Petzold, W. Zhu. Model reduction for chemical kinetics: an optimization approach. *AIChE J*, volume 45, pages 869-886, 1999.
[14] O. Radulescu, A.N. Gorban, A. Zinovyev, V. Noel. Reduction of dynamical biochemical reaction networks in computational biology. *Front Genet*, 3:00131, 2012.
[15] O. Radulescu, A.N. Gorban, A. Zinovyev, A. Lilienbaum. Robust simplifications of multiscale biochemical networks. *BMC Syst Biol*, 2:86, 2008.
[16] S. Rao, A.J. van der Schaft, K. van Eunen, B.M. Bakker, B. Jayawardhana. Model reduction of biochemical reaction networks. *BMC Syst. Biol.*, 8:52, 2014.
[17] M.R. Roussel, S.J. Fraser. Invariant manifold methods for metabolic model reduction. *Chaos*, volume 11, pages 196-206, 2001.
[18] I.H. Segel. Enzymes. *Biochemical Calculations: How to solve mathematical problems in general biochemistry*. 2nd edition. New York: John Wiley and sons, pages 208-323, 1968.
[19] L.A. Segel, M. Slemrod. The quasi-steady-state assumption: A case study in perturbation. *SIAM Rev Soc Ind Appl Math*, volume 31, pages 446-477, 1989.
[20] M. Sunnåker, G. Cedersund, M. Jirstrand. A method for zooming of nonlinear models of biochemical systems. *BMC Syst Biol*, 5:140, 2011.
[21] A.J. van der Schaft, S. Rao, B. Jayawardhana. On the Mathematical Structure of Balanced Chemical Reaction Networks Governed by Mass Action Kinetics. *SIAM J. Appl. Math.*, volume 73, issue 2, pp. 953-973, 2013.
[22] A.J. van der Schaft, S. Rao, B. Jayawardhana. A Network Dynamics Approach to Chemical Reaction Networks. *Int. J. Contr.*, volume 89, issue 4, pp. 731-745, 2016.