Structural analysis of biochemically motivated nonlinear systems

Theses of the Ph.D. Dissertation

Bernadett Ács

Thesis Advisor:
Dr. Gábor Szederkényi, D.Sc.

Pázmány Péter Catholic University
Faculty of Information Technology and Bionics
Roska Tamás Doctoral School of Sciences and Technology

2018

1 Introduction

The development and/or the maintenance of any kind of device or system requires some kind of knowledge about its possible states. To reveal the connections among properties of different events and to try to predict certain characteristics of the future quantitative mathematical models are successfully applied. These models usually describe only selected properties of the real process, but from an application point of view in most cases it is enough.

The operation of more complicated systems such as living organisms can often be described by complex phenomena, and for the modelling of quantities changing in space and/or in time dynamical systems are the most commonly applied tools. Therefore, this type of modelling has become an intensively studied and frequently applied tool in systems biology.

In many real life problems for example in economic systems, population dynamics or biochemical systems the variables are physically constrained to have only nonnegative values, and therefore the theory of nonnegative systems [1] needs to be applied for their characterization. A dynamical system is called nonnegative if its trajectories remain in the nonnegative orthant whenever the initial value is nonnegative. (If strict positivity is required then it is called a positive system.) A wide subclass of dynamical systems can be transformed into nonnegative form by shifting the coordinates into the nonnegative orthant and then in a further transformed version of the model the trajectories can be kept in a given region, see [2].

A more special class of nonnegative dynamical systems is formed by the quasi polynomial (QP) systems, which was first introduced in [3]. The author has also shown that most smooth dynamical models can algorithmically be transformed into QP form, which property makes such systems suitable for the modelling of dynamical systems belonging to a much wider class.

If the right hand side of the ordinary differential equations of the system can be given in the form of a multivariate polynomial as well, then it is called a polynomial system. The aim of this thesis is the structural and computational analysis of a certain type of nonnegative polynomial systems called kinetic systems, that can describe the dynamics of chemical reaction network (CRN) models obeying the mass action law. Despite the fact that kinetic systems are rather special polynomial systems, these models are versatile tools in modelling. Furthermore, by using suitable

model transformations the majority of nonnegative dynamical systems can be transformed into kinetic form [3, 4].

The different types of dynamical systems and possible transformations between them are shown on the diagram in Figure 1.

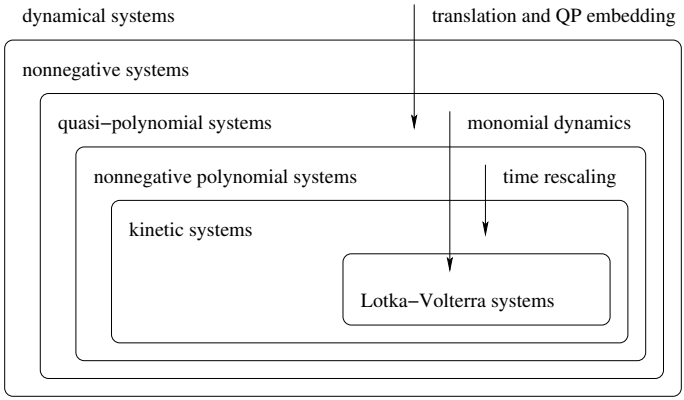

Figure 1: Classes and transformations of dynamic systems.

Chemical reaction networks obeying the mass action law can be originated from the dynamical modelling of chemical and biochemical processes, but they can be applied to describe various kinds of dynamical phenomena. Their applications appear in several different fields of science and engineering, such as the modelling of electrical networks, transportation problems or the spreading of epidemics, therefore these models are so-called universal descriptors [5, 2].

The class of kinetic systems is defined by chemical reaction network models, but for the verification of the kinetic property it is not necessary to compute a suitable CRN, it is enough to examine just the sign pattern of the monomial coefficients, see [6].

It is known that in general there are many realizations and different reaction graph structures corresponding to a given kinetic dynamics. This phenomenon is called macro-equivalence or dynamical equivalence [7]. There is also a generalization of dynamical equivalence called linear conjugacy, where a positive definite diagonal linear transformation is applied to the state variables working as if the units of measurement were individually scaled [8]. It is easy to see that linear conjugacy pre-

serves the kinetic property of the system and also the main qualitative dynamical properties like stability, multiplicities or the boundedness of solutions. However, due to the larger degree of freedom introduced by the transformation parameters, in general, it allows a larger set of possible structures compared to dynamical equivalence.

There is a widely applied structure oriented representation of CRNs that is a weighted directed graph called the Feinberg-Horn-Jackson graph. It depicts the reactions which are present in the network, and some other parameters of the network as well that are easier to describe with graph properties. Furthermore, in some cases there is a relation between the dynamics of the network and the reaction graph structure, without considering the actual reaction rates. This has become an important research area in chemical reaction network theory since the 1970s, see [7]. In this topic there are several practice oriented results as well as beautiful mathematical designs.

To determine a possible reaction network structure of a given kinetic system a symbolic method was proposed in [6]. Since this method returns only one particular dynamically equivalent realization called the canonical realization, a different approach must be applied to determine others.

Chemical reaction networks have a simple algebraic characterization, which makes it particularly appealing to develop computational methods for their dynamical and structural analysis [9, 1] or even control [10]. Realizations of a given kinetic dynamics can be defined by linear constraints, that suggests the application of linear optimization methods. Since this is in general a very simple model, several computational methods have already been developed to find linearly conjugate or, as a special case, dynamically equivalent realizations of kinetic systems and also having preferred properties such as density/sparsity , maximal or minimal realizations, complex or detailed balance, weak reversibility or minimum deficiency.

2 Basic notions

For clarity the most important notions and tools considering the topic of the thesis are introduced in this section.

The general form of a **polynomial system** is defined by a function $x : \mathbb{R} \to \mathbb{R}^n$, a coefficient matrix $M \in \mathbb{R}^{n\times p}$ and a monomial-type vector-mapping $\varphi : \mathbb{R}^n \to \mathbb{R}^p$ with coordinate functions of the form $\varphi_j(x) = \sum x_i^{\beta_{ij}}$, where $\beta_{ij} \in \mathbb{N}$ for all $i \in \{1,\dots,n\}$ and $j \in \{1,\dots,p\}$. Using these notations the dynamical equations of the polynomial system can be written as

$$\dot{x} = M \cdot \varphi(x) \tag{1}$$

By definition a nonnegative polynomial system is called kinetic if there is a chemical reaction network (CRN) with the given dynamical behaviour. A **chemical reaction network** can be characterized by three sets.

$$\begin{aligned}
&\textbf{species:} && \mathcal{S} = \{X_i \mid i \in \{1,\dots,n\}\} \\
&\textbf{complexes:} && \mathcal{C} = \{C_j = \sum_{i=1}^{n} \alpha_{ji} \cdot X_i \mid \alpha_{ji} \in \mathbb{N}\} \\
&\textbf{reactions:} && \mathcal{R} \subseteq \{(C_i, C_j) \mid C_i, C_j \in \mathcal{C}\}
\end{aligned}$$

The rate of the reaction (C_i, C_j) for $i, j \in \{1,\dots m\}$, $i \neq j$ is determined by the corresponding **reaction rate coefficient** $k_{ij} \in \mathbb{R}_+$. This reaction is present in the reaction network if and only if $k_{ij} > 0$ holds.

The quantitative properties of chemical reaction networks can be characterized by special matrices. The linear combinations defining the structures of the complexes are decoded by the **complex composition matrix** $Y \in \mathbb{N}^{n\times m}$, where

$$[Y]_{ij} = \alpha_{ji} \qquad i \in \{1,\dots,n\},\ j \in \{1,\dots,m\} \tag{2}$$

The structure of the reaction network is described through the reaction rates by the **Kirchhoff matrix** $A_k \in \mathbb{R}^{m\times m}$ of the CRN as follows:

$$[A_k]_{ij} = \begin{cases} k_{ji} & \text{if } i \neq j \\ -\sum\limits_{l=1, l\neq i}^{m} k_{il} & \text{if } i = j \end{cases} \tag{3}$$

Let the function $x : \mathbb{R} \to \mathbb{R}_+^n$ describe the concentrations of the species depending on time. Assuming mass-action kinetics the dynamics of the concentrations can be characterized by a polynomial system:

$$\dot{x} = Y \cdot A_k \cdot \psi^Y(x) \tag{4}$$

where $\psi^Y : \mathbb{R}^n \to \mathbb{R}^m$ is the **monomial function** of the CRN. The monomials, i.e. the coordinate functions correspond to the complexes

and they are defined as

$$\psi_j^Y(x) = \prod_{i=1}^{n} x_i^{Y_{ij}} \qquad j \in \{1, \dots, m\} \tag{5}$$

A polynomial system (1) is called **kinetic** if there exists a chemical reaction network characterized by the matrix pair (Y, A_k) that is governed by the same dynamics, i.e. the following equation holds.

$$M \cdot \varphi(x) = Y \cdot A_k \cdot \psi^Y(x) \tag{6}$$

In this case the CRN (Y, A_k) is called a dynamically equivalent realizations of the kinetic system (1).

The notion of dynamical equivalence can be extended to the case when the state variable x is transformed by a positive definite diagonal matrix $T \in \mathbb{R}^{n \times n}$ to the form $\bar{x} = T^{-1} \cdot x$ (i.e. $x = T \cdot \bar{x}$).

The reaction network (Y, A_k) is a realization of the transformed model and is called a **linearly conjugate realization** of the kinetic system (1) if there exists a positive definite diagonal matrix $T \in \mathbb{R}^{n \times n}$ so that the following equation holds

$$Y \cdot A_k \cdot \psi^Y(x) = T^{-1} \cdot M \cdot \Phi_T \cdot \varphi(x) \tag{7}$$

where $\Phi_T \in \mathbb{R}^{p \times p}$ is also a transformation matrix depending on T. It has to be mentioned that dynamically equivalent realizations form a special case among linearly conjugate realizations when the matrix T is the unit matrix.

The CRN defined by the sets $\mathcal{S}, \mathcal{C}, \mathcal{R}$ and reaction rate coefficients k_{ij} for all $i, j \in \{1, \dots, m\}, i \neq j$ can be represented by the directed graph $G(V, E)$ with weight function $w : E(G) \to \mathbb{R}_+$ as its reaction graph if

the **vertices** correspond to the complexes	–	$V(G) = \mathcal{C}$
the **directed edges** represent the reactions	–	$E(G) = \mathcal{R}$
the **weights** are the reaction rate coefficients	–	$w((C_i, C_j)) = k_{ij}$

Then there is a directed edge in the reaction graph if and only if the corresponding reaction considering the direction takes place in the reaction network and its weight is the reactions rate constant. If the reaction rates are not considered in the graph then it is called a **reaction graph structure**.

3 Computational methods

By modifying the description of the kinetic system, and fixing the set of complexes one can get an equivalent form of Equation (7) as:

$$Y \cdot A_b = T^{-1} \cdot M \tag{8}$$

where $A_b = A_k \cdot {\Psi_T}^{-1}$ is also a Kirchhoff matrix and it represents the same reaction graph structure as A_k. Since the matrices Y and M are fixed, both sides of the equation are linear functions. From this equation and the properties of the matrices can the linear programming model characterizing linearly conjugate realizations be formed.

The known parameters of the model are the matrices M and Y that define a kinetic system with a fixed set of complexes. While the optimization variables are the entries of the matrices A_b and T^{-1}, since these are more convenient variables and they uniquely define the reaction network. To ensure linear conjugacy Equation (9) must be fulfilled. It is equivalent to Equation (7), where $\mathbf{0}^{n\times m} \in \mathbb{R}^{n\times m}$ is a zero matrix. The Equations (10), (11) and (12) are necessary to ensure that the matrices A_b and T^{-1} meet their definitions.

$$Y \cdot A_b - T^{-1} \cdot M = \mathbf{0}^{n\times m} \tag{9}$$

$$\sum_{\substack{j=1 \\ j\neq i}}^{m} [A_b]_{ji} = -[A_b]_{ii} \qquad i \in \{1,\dots,m\} \tag{10}$$

$$[A_b]_{ij} \geq 0 \qquad i,j \in \{1,\dots,m\},\ i \neq j \tag{11}$$

$$[T^{-1}]_{ll} > 0 \qquad l \in \{1,\dots,n\} \tag{12}$$

If a realization with special properties is required to be determined, further linear constraints can be added to the model, and the objective function of the optimization can also be defined according to the additional requirements.

For solving an LP problem there are several polynomial time algorithms, the first provably correct solution is the Simplex Algorithm developed by Dantzig in 1947. This algorithm works in most of the practical applications very efficiently. Later several other algorithms have also been designed for the efficient computation of linear optimization problems, such as the criss-cross method or the ellipsoid method [11].

The computation of some special realizations might require the using of integer and continuous variables at the same time, which transforms the model into a mixed integer linear programming (MILP) problem. This problem is known to be NP-complete, which means in practical applications that there is no polynomial-time method for solving it. There are several approximative methods, such as solving the LP-relaxed version, but also exact methods such as the cutting plane method and the Branch and Bound method. Despite the many possible solutions it is still desired to avoid the application of integer variables, since there are much more efficient methods for solving linear optimization problems defined on continuous variables.

4 New scientific results

The problem of computing dense realizations can be formed at first sight as a MILP problem. In the literature there are still only non-polynomial time solutions or ones that work in polynomial-time but only under certain restrictions. However, the application of convex geometry changes the problem to be solvable. The results considering dense realizations are formed in Thesis I.

Weakly reversible realizations form an intensively studied class of CRN realizations where there is a connection between structure and dynamics. One of the most important results in this area is the Deficiency Zero Theorem [12], which states that weakly reversible realizations under some further conditions have locally and in some special cases globally stable equilibrium points. For this reason it is important to have a method for computing even linearly conjugate weakly reversible realizations. The results are summarized in Thesis II.

Is it possible to give a computationally efficient algorithm for determining all possible reaction graph structures representing linearly conjugate CRN realizations of a given kinetic system? The answer is yes, the corresponding results are introduced in Thesis III.

A generalization of kinetic systems has been introduced that is suitable for handling uncertain parameters and also additional linear constraints, whenever the possible values of the unknown parameters can be represented as points of a convex polyhedron. Due to the similar model structure several results developed for the case of non-uncertain kinetic systems can be proven for uncertain models as well. The results are listed in Thesis IV.

Thesis I. ***I have proven new results regarding to dense realizations of kinetic systems, using a geometric approach.***
The realizations represented as points in the Euclidean space form a convex polyhedron, and this property can be utilized efficiently from a computational point of view.
Thesis I.a ***I have proven that a dense linearly conjugate realization of a kinetic system with a fixed set of complexes and an additional finite set of linear constraints determines a superstructure considering all realizations of the constrained model.***
The superstructure property is essential for the correct operation of all the algorithms presented in this dissertation.
The results are described in detail in [14], [18], [19] and in Section 3.1.
Thesis I.b ***I have developed a novel polynomial-time algorithm to compute a dense linearly conjugate realization of a kinetic system with a fixed set of complexes and fulfilling an additional finite set of linear constraints.***
The advantage of the method is that it applies linear optimization methods, it avoids the use of integer variables, and it works for every kinetic system without restrictions on the variables. I have proven that the algorithm returns the dense linearly conjugate realization, or as special case the dense dynamically equivalent realization of any kinetic system. This algorithm is applied as a subroutine in the algorithms presented in Theses II, III.a, III.b and IV.b.
I have shown that even if there are arbitrarily predefined upper bounds considering the variables the set of possible reaction graph structures representing linearly conjugate realizations is the same as in the unbounded case, therefore the computer implementations of the algorithms presented in this dissertation can work accurately.
The results are described in detail in [14], [18] and in Section 3.2.

Thesis II. ***I have proposed a new algorithm for computing a weakly reversible linearly conjugate realization of a kinetic system by extending the method introduced in*** [13]**.**
I have proven that the algorithm runs in polynomial time, and it returns a dense weakly reversible linearly conjugate realization of the kinetic system, if it exists.
I have also shown that the computed dense realization defines a superstructure among all linearly conjugate weakly reversible realizations of the kinetic system.
The results are described in detail in [14], [18], [19] and in Chapter 4.

Thesis III. ***I have achieved new results on computing all possible reaction graph structures representing linearly conjugate realizations of a kinetic system.***

Thesis III.a ***I have proven the correctness of a new algorithm for computing all possible reaction graph structures representing linearly conjugate realizations of a kinetic system on a fixed set of complexes.***

The algorithm is the first method in the literature for computing all the reaction graph structures realizing a given kinetic dynamics.

The computation might require exponential time because of the large number of possible structures, however, between the determination of two different structures polynomial time is elapsed. Furthermore, it is possible to apply parallel implementation of the algorithm using e.g. many core architectures.

The results are described in detail in [15], [20], [21] and in Section 5.1.

Thesis III.b ***I have designed a new efficient algorithm for computing all structurally different linearly conjugate realizations of a kinetic system.***

I have proven that this algorithm also returns all possible reaction graph structures representing linearly conjugate realizations of a kinetic system.

I have also shown that the algorithm returns every realization only once, furthermore, it is also suitable for parallel implementation.

The performance of the new algorithm has been compared to that of the algorithm in Thesis III.a, and considering all the examples the number of required optimization steps decreased by more than 80% in the case of the new algorithm.

The results are described in detail in [16], [22] and in Section 5.2.

Thesis IV. ***I have proven new results regarding to special uncertain kinetic system models, where the parameters are in a convex polyhedron.***

The introduced model is a generalization of the original kinetic model that can include a finite set of additional linear constraints as well.

Thesis IV.a ***I have shown that the superstructure property of dense realizations holds also in the case of uncertain kinetic systems.***

This property depends on the fact that the set of solutions of an uncertain kinetic model is a convex polyhedron.

The results are described in detail in [17], [23] and in Section 6.1.2

Thesis IV.b ***I have proven that the algorithms designed for computing the dense realization, the set of core reactions and all realization structures of a given kinetic system can be extended for the case of uncertain kinetic systems.***
I have also shown that the algorithm developed for computing all structurally different realizations and presented in Thesis III.b is suitable for parallel implementation.
The results are described in detail in [17], [23] and in Section 6.2.

5 Application possibilities

The possibilities of applying existing algorithms is wide, since these can often be used as parts of other computational methods. The algorithm developed for computing dense realizations has already been applied as subroutine in all the other algorithms introduced in this thesis.

The algorithm extended for the computation of weakly reversible linearly conjugate realizations is the first method for solving this problem, which will possibly generate some new ideas and interesting structure based results considering these special realizations. For example one can take advantage of the fact that the algorithm returns the dense weakly reversible realization which defines a superstructure among weakly reversible realizations of the kinetic system. Furthermore, the algorithm can be generalized to the case of constrained kinetic systems, and by using this kind of computation steps it is possible to design an algorithm for determining every weakly reversible realization corresponding to a given kinetic system.

The algorithms designed for determining the set of possible reaction graph structures can be applied for the accurate computation of realizations which are more difficult to characterize, for example sparse realizations. An other application of this computational method might be the CRN design based on dynamics.

It is clear that the kinetic model defined with uncertain parameters has the biggest potential in practical applications. For example in the case of a system model identified by the application of noisy measurements the kinetic model with polyhedric uncertainty introduced in this thesis can be defined using the estimated values of the parameters.

References

[1] W. M. Haddad, VS. Chellaboina, and Q. Hui. *Nonnegative and Compartmental Dynamical Systems.* Princeton University Press, 2010.

[2] N. Samardzija, L. D. Greller, and E. Wassermann. Nonlinear chemical kinetic schemes derived from mechanical and electrical dynamical systems. *Journal of Chemical Physics*, 90 (4):2296–2304, 1989.

[3] L. Brenig. Complete factorisation and analytic solutions of generalized Lotka-Volterra equations. *Physics Letters A*, 133:378–382, 1988.

[4] A. Figueiredo, T. M. Rocha Filho, and L. Brenig. Necessary conditions for the existence of quasi-polynomial invariants: the quasi-polynomial and Lotka-Volterra systems. *Physica A*, 262:158–180, 1999.

[5] 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.

[6] V. Hárs and J. Tóth. On the inverse problem of reaction kinetics. In M. Farkas and L. Hatvani, editors, *Qualitative Theory of Differential Equations*, volume 30 of *Coll. Math. Soc. J. Bolyai*, pages 363–379. North-Holland, Amsterdam, 1981.

[7] F. Horn and R. Jackson. General mass action kinetics. *Archive for Rational Mechanics and Analysis*, 47:81–116, 1972.

[8] M. D. Johnston and D. Siegel. Linear conjugacy of chemical reaction networks. *Journal of Mathematical Chemistry*, 49:1263–1282, 2011.

[9] D. Angeli. A tutorial on chemical network dynamics. *European Journal of Control*, 15:398–406, 2009.

[10] E. Sontag. Structure and stability of certain chemical networks and applications to the kinetic proofreading model of T-cell receptor signal transduction. *IEEE Transactions on Automatic Control*, 46:1028–1047, 2001.

[11] A. Schrijver. *Theory of Linear and Integer Programming.* Wiley & Sons, 1998.

[12] M. Feinberg. *Lectures on chemical reaction networks.* Notes of lectures given at the Mathematics Research Center, University of Wisconsin, 1979.

[13] G. Szederkényi, K. M. Hangos, and Zs. Tuza. Finding weakly reversible realizations of chemical reaction networks using optimization. *MATCH Commun. Math. Comput. Chem.*, 67:193–212, 2012.

The author's journal publications

[14] B. Ács, G. Szederkényi, Z. A. Tuza, and Zs. Tuza. Computing linearly conjugate weakly reversible kinetic structures using optimization and graph theory. *MATCH Commun. Math. Comput. Chem.*, 74:481–504, 2015. IF: 3.858, Q1.

[15] B. Ács, G. Szederkényi, Zs. Tuza, and Z. A. Tuza. Computing all possible graph structures describing linearly conjugate realizations of kinetic systems. *Computer Physics Communications*, 204:11–20, 2016. IF: 3.936, D1.

[16] B. Ács, G. Szederkényi, and D. Csercsik. A new efficient algorithm for determining all structurally different realizations of kinetic systems. *MATCH Commun. Math. Comput. Chem.*, 77:299–320, 2017. IF: 3.139, Q1.

[17] B. Ács, G. Szlobodnyik, and G. Szederkényi. A computational approach to the structural analysis of uncertain kinetic systems. *submitted*, 2017. `https://arxiv.org/abs/1704.08633`.

The author's conference publications

[18] B. Ács. A new method for computing linearly conjugate weakly reversible structures of kinetic polynomial systems. In C. Wiuf and E. Feliu, editors, *Workshop on Mathematical Trends in Reaction Network Theory, July 1-3, 2015, Copenhagen, Denmark*, page 21, 2015.

[19] B. Ács. Computing linearly conjugate weakly reversible kinetic structures using graph theory and optimization. In *PhD Proceedings Annual Issues of the Doctoral School, Faculty of Information Technology and Bionics*, pages 69–71, 2014.

[20] Z. A. Tuza, B. Ács, G. Szederkényi, and F. Allgöwer. Efficient computation of all distinct realization structures of kinetic systems. In *6th IFAC Conference on Foundations of Systems Biology in Engineering, October 9-12, 2016 Magdeburg, Germany, IFAC-PAPERSONLINE 49:(26)*, pages 194–200, 2016.

[21] B. Ács. Computational analysis of kinetic systems. In *PhD Proceedings Annual Issues of the Doctoral School, Faculty of Information Technology and Bionics*, pages 107–110, 2015.

[22] B. Ács. Case studies for the structural analysis of biochemically motivated nonlinear systems. In *PhD Proceedings Annual Issues of the Doctoral School, Faculty of Information Technology and Bionics*, pages 117–120, 2016.

[23] G. Szederkényi, B. Ács, and G. Szlobodnyik. Structural analysis of kinetic systems with uncertain parameters. In *2nd IFAC Workshop on Thermodynamic Foundations for a Mathematical Systems Theory - TFMSTII, September 28-30, 2016, Vigo, Spain, IFAC-PAPERSONLINE 49:(24)*, pages 24–27, 2016.