Families of Polynomials in the Study of Biochemical Reaction Networks

Alicia Dickenstein(✉)

Department of Mathematics, FCEN, University of Buenos Aires and IMAS (UBA-CONICET), C. Universitaria, Pab. I, C1428EGA Buenos Aires, Argentina
alidick@dm.uba.ar

Abstract. The standard mass-action kinetics modeling of the dynamics of biochemical reaction networks gives rise to systems of ordinary polynomial differential equations with (in general unknown) parameters. Attempts to explore the parameter space in order to predict properties of the associated systems challenge the standard current computational tools because even for moderately small networks we need to study families of polynomials with many variables and many parameters. These polynomials have a combinatorial structure that comes from the digraph of reactions. We show that different techniques can be strengthened and applied for biochemical networks with special structure.

1 Introduction

The basic definitions and properties of chemical reaction networks, together with the features of some important biochemical networks, can be found in the surveys [7–9] and Chap. 5 of the book [6], as well as in the book [12]. The starting information is a finite directed graph with r labeled edges that correspond to the reactions and nodes that correspond to complexes, given by nonnegative integer linear combinations of a set of s chemical species. The concentrations $x = (x_1, \ldots, x_s)$ of the chemical species are viewed as functions of time. Under mass-action kinetics, the labels of the edges are positive numbers called reaction rate constants and x is assumed to satisfy an autonomous system of ordinary differential equations $\frac{dx}{dt} = f(x)$. Here $f = (f_1, \ldots, f_s)$ is a vector of real polynomials that reflects the combinatorics of the graph.

The reaction rate constants are in general unknown or difficult to measure. Standard methods in other sciences involve exhaustive sampling. Instead, we think the vector κ of reaction rate constants as a vector of parameters. In general, there are further parameters involved in this setting. Linear relations describing the span S of the difference of the complexes on each side of a reaction give rise to linear conservation constants of the dynamics. This means that given a basis $\ell_1, \ldots, \ell_d$ of the orthogonal subspace $S^\perp$, any solution x defined in an interval satisfies linear constraints of the form $\ell_1(x) = T_1, \ldots, \ell_d(x) = T_d$. We say that $T = (T_1, \ldots, T_d)$ is a vector of total amounts and we consider (κ, T) as parameters.

F. Boulier et al. (Eds.): CASC 2021, LNCS 12865, pp. 1–5, 2021.
https://doi.org/10.1007/978-3-030-85165-1_1

The steady states of the the system $\frac{dx}{dt} = f(x)$ are the constant trajectories, that is the values of x^* for which $f(x^*) = 0$. If a trajectory converges, its limit is a steady state. Stable steady states attract nearby trajectories and unstable steady states also drive the dynamics. Multistationarity is a crucial property for chemical reaction networks modeling biological processes, since it allows for different "responses" of the cell. It corresponds to the existence of more than one *positive* steady state with the same total amounts, that is, to the existence of at least two positive zeros of the ideal $\langle f_1 \dots, f_s, \ell_1 - T_1, \dots, \ell_d - T_d \rangle$.

We look at these systems as special families of polynomial ordinary differential equations in s variables with $r + d$ parameters. Our aim is to explore the parameter space in order to predict properties of the systems associated to networks studied in systems biology, which usually have too many variables and too many parameters. There are many useful mathematical and computational tools, but we are forced to extend the mathematical results and to understand the structure of the networks to make the computations feasible.

In the following sections, I will very briefly summarize two of these recent advances. Besides consulting the references, the reader is invited to attend my lecture or to watch later the video for more information.

2 The ERK Pathway

As an example of more general results, we discuss the ERK pathway. It is an enzymatic network that consists of a cascade of phosphorylation of proteins in the cell that communicates a signal from a receptor on the outside membrane to the DNA inside the nucleus. It controlls different responses such as cell division [19]. It is known that the ERK pathway has the capacity for multistationarity and there are oscillatory solutions.

Deciding mulstistationarity is a question in real algebraic geometry that can be effectively decided in practice, but the associated family has 21 variables and $36 = 30 + 6$ parameters. So, how is it that we can study it with an algebro-geometric approach? This important signaling cascade, as most popular models in systems biology, has a *MESSI structure* [21]. There is a partition of the set of species and only certain type of reactions occur. Using this structure, we give combinatorial conditions on the network that ensure the following:

- There are no relevant boundary steady states. That is, there are no steady states (zeros of the polynomials $f_1, \dots, f_{21}$) in the boundary of the nonnegative orthant which lie in the closure of the linear variety $S_T = \{\ell_1(x) = T_1, \dots, \ell_d(x) = T_6\}$, for any choice of $\kappa \in \mathbb{R}^{30}_{>0}$ and T such that S_T intersects the positive orthant.
- The intersections $S_T \cap \mathbb{R}^{21}_{\geq 0}$ are compact and so the system is conservative.
- The system is *linearly binomial*, a concept introduced in [11], which implies that there is a system of binomial generators of the ideal $\langle f_1, \dots, f_{21} \rangle$ obtained by linear algebra operations over $\mathbb{Q}(\kappa)$, involving rational functions whose denominators do not vanish over $\mathbb{R}^{30}_{>0}$.

– The positive points of the steady state variety $\{x \in \mathbb{R}^{21}_{>0} : f(x) = 0\}$ can be cut out by explicit binomials, and thus parametrized by explicit monomials with coefficients in $\mathbb{Q}(\kappa)$ as above.

One way to approximate the dynamics of biological models while dealing with less variables and parameters, is the elimination of the *intermediate complexes* [14]. Following [24], one could ask which are the minimal sets (with respect to inclusion) of intermediates that still give rise to multistationarity. These sets are termed *circuits of multistationarity*. We show in our forthcoming paper [10] that systems like the ERK pathway without intermediates cannot be multistationary and we use a computer algebra system to find all the corresponding circuits of multistationarity. We can also identify the circuits of multistationarity for phosphorylation networks with any number of species. The theoretical results are based on [5,22].

3 Degenerations and Open Regions of Multistationarity

In the beautiful paper [3], regular subdivisions of the (convex hull of the) set of exponents of a polynomial system are used to get a lower bound on the number of positive solutions, with combinatorial arguments to get new lower bounds in terms of the number of variables and the difference between the cardinality of the support and the number of variables. This is based on classical results on degenerations that were used in [25] to study real roots of complete intersections. The idea is to add a parameter u raised to the different heights of a fixed lifting whose projection produces the given regular subdivision, thus giving a deformation of the coefficients of the system along a curve. For small positive values of u, one obtains a degeneration of the original system for which a lower bound on the number of positive roots can be given in terms of *decorated* simplices in the regular subdivision. Again, this is in general unfeasible in practice when there are many variables and many monomials.

On one side, we show how to replace a deformation using a single parameter with an open set defined in terms of the cone of all height functions that produce the regular subdivision. This way, we get an open region in parameter space where multistationarity occurs [2]. Even if deciding if simplices are part of a same regular subdivision is algorithmic, in order to do this when the dimension or the number of monomials is big, we use the simple idea that if two simplices share a facet, then this is always the case. Moreover, we heavily use results about the structure of *s-toric* MESSI systems from [21]. This allows us to find these open regions for cascades with any number of layers in [15], but the lower bound that we get is three. Regions of multistationarity with higher lower bounds are in general unknown, except for the case of sequential distributive phosphorylation networks [16]. There is also a degeneration approach with one parameter using arguments from geometric singular perturbation theory in [13].

4 Other Computational Approaches

There are several other computational approaches to study these systems. Of course, symbolic software using Gröbner bases and in particular real algebraic geometry libraries, as well as Cylindrical Algebraic Decomposition software. Also numerical methods in algebraic geometry can be used [17,18], as well as tropical tools to separate time scales [23]. Machine learning tools started to be used to improve both the symbolic and numeric calculations [1,4,20].

Acknowledgments. We acknowledge the support of ANPCyT PICT 2016-0398, UBACYT 20020170100048BA and CONICET PIP 11220150100473, Argentina.

References

1. Bernal, E., Hauenstein, J., Mehta, D, Regan, M., Tang, T.: Machine learning the discriminant locus. Preprint available at arXiv:2006.14078 (2020)
2. Bihan, F., Giaroli, M., Dickenstein, A.: Lower bounds for positive roots and regions of multistationarity in chemical reaction networks. J. Algebra **542**, 367–411 (2020)
3. Bihan, F., Santos, F., Spaenlehauer, P.-J.: A polyhedral method for sparse systems with many positive solutions. SIAM J. Appl. Algebra Geom. **2**(4), 620–645 (2018)
4. Böhm, J., Decker, W., Frühbis-Krüger, A., Pfreundt, F.-J., Rahn, M., Ristau, L.: Towards massively parallel computations in algebraic geometry. Found. Comput. Math. **21**(3), 767–806 (2020). https://doi.org/10.1007/s10208-020-09464-x
5. Conradi, C., Feliu, E., Mincheva, M., Wiuf, C.: Identifying parameter regions for multistationarity. PLoS Comput. Biol. **13**(10), e1005751 (2017)
6. Cox, D.A.: Applications of Polynomial Systems, with contributions by C. D'Andrea, A. Dickenstein, J. Hauenstein, H.Schenck, and J. Sidman. Co-publication of the AMS and CBMS (2020)
7. Dickenstein, A.: Biochemical reaction networks: an invitation for algebraic geometers. In: MCA 2013, Contemporary Mathematics, vol. 656, pp. 65–83 (2016)
8. Dickenstein, A.: Algebra and geometry in the study of enzymatic cascades. In: Araujo, C., Benkart, G., Praeger, C.E., Tanbay, B. (eds.) World Women in Mathematics 2018. AWMS, vol. 20, pp. 57–81. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-21170-7_2
9. Dickenstein, A.: Algebraic geometry tools in systems biology. Notices Amer. Math. Soc. **67**(11), 1706–1715 (2020)
10. Dickenstein, A., Giaroli, M., Pérez Millán, M., Rischter, R.: Detecting the multistationarity structure in enzymatic networks. Manuscript (2021)
11. Dickenstein, A., Pérez Millán, M., Shiu, A., Tang, X.: Mutistationarity in structured reaction networks. Bull. Math. Biol. **81**, 1527–1581 (2019)
12. Feinberg, M.: Foundations of Chemical Reaction Network Theory. AMS, vol. 202. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-03858-8
13. Feliu, E., Rendall, A., Wiuf, C.: A proof of unlimited multistability for phosphorylation cycles. Nonlinearity **33**(11), 5629 (2020)
14. Feliu, E., Wiuf, C.: Simplifying biochemical models with intermediate species. J. R. Soc. Interface **10**, 20130484 (2013)
15. Giaroli, M., Bihan, F., Dickenstein, A.: Regions of multistationarity in cascades of Goldbeter-Koshland loops. J. Math. Biol. **78**(4), 1115–1145 (2019)

16. Giaroli, M., Rischter, R., Pérez Millán, M., Dickenstein, A.: Parameter regions that give rise to $2\lfloor \frac{n}{2} \rfloor + 1$ positive steady states in the n-site phosphorylation system. Math. Biosci. Eng. **16**(6), 7589–7615 (2019)
17. Gross, E., Harrington, H.A., Rosen, Z., Sturmfels, B.: Algebraic systems biology: a case study for the Wnt pathway. Bull. Math. Biol. **78**(1), 21–51 (2015). https://doi.org/10.1007/s11538-015-0125-1
18. Nam, K., Gyori, B., Amethyst, S., Bates, D., Gunawardena, J.: Robustness and parameter geography in post-translational modification systems. PLoS Comput. Biol. **16**(5), e1007573 (2020)
19. Patel, A., Shvartsman, S.: Outstanding questions in developmental ERK signaling. Development **145**(14), dev143818 (2018)
20. Peifer, D., Stillman, M., Halpern-Leistner, D.: Learning selection strategies in Buchberger's algorithm. In: Proceedings of the 37th International Conference on Machine Learning, Online, PMLR 119, pp. 7575–7585 (2020)
21. Pérez Millán, M., Dickenstein, A.: The structure of MESSI biological systems. SIAM J. Appl. Dyn. Syst. **17**(2), 1650–1682 (2018)
22. Pérez Millán, M., Dickenstein, A., Shiu, A., Conradi, C.: Chemical reaction systems with toric steady states. Bull. Math. Biol. **74**(5), 1027–1065 (2012)
23. Radulescu, O.: Tropical geometry of biological systems (*Invited Talk*). In: Boulier, F., England, M., Sadykov, T.M., Vorozhtsov, E.V. (eds.) CASC 2020. LNCS, vol. 12291, pp. 1–13. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-60026-6_1
24. Sadeghimanesh, A., Feliu, E.: The multistationarity structure of networks with intermediates and a binomial core network. Bull. Math. Biol. **81**, 2428–2462 (2019)
25. Sturmfels, B.: On the number of real roots of a sparse polynomial system. In: Hamiltonian and Gradient Flows, Algorithms and Control, Fields Inst. Commun., 3, Amer. Math. Soc., Providence, RI, pp. 137–143 (1994)