Tropicalization of systems biology models

Ovidiu Radulescu, Dima Grigoriev, Vincent Noel, and Sergei Vakulenko

Abstract. Systems biology use networks of biochemical reactions as models for cellular process. The dynamics of reaction networks with many well separated time scales, is well captured by asymptotic models obtained by tropicalization of the smooth dynamics, via the Litvinov-Maslov correspondence principle. The tropicalized models can be used to check the global stability and to identify sensitive parameters and rapid variables of the original models.

1. Introduction

In the last decade, systems biology became the playground of several mathematical fields of study, among which algebraic geometry is one of the most important. Cellular biochemistry can be suitably modelled by networks of reactions with rational and polynomial rate functions. The dynamics of these networks can be described by rational or polynomial ODEs, for which some results exist concerning the type and the complexity of the solutions.

Most of the previous algebraic work on reaction networks was dedicated to the study of steady states [**CTF06, Sou03, Son05, RLS$^+$06, RSP$^+$11**]. This issue is important, because networks with multiple stable steady states control biological cell fate decisions in development and differentiation [**T$^+$98, Del49**].

However, biological cell physiology relies preeminently upon network dynamics. To interpret stimuli, adapt to environmental changes, make decisions, the cells lean on the rich dynamical possibilities of regulatory networks. We say that regulatory networks are flexible, because they can support, in principle, any type of attractors and spatial-temporal patterns [**VRv2**]. Networks are also robust, because these patterns resist to perturbations and are maintained for wide ranges of parameter values [**GR07, RGZL08**]. To achieve robust and flexible functioning, biological networks employ hierarchies of biochemical processes with well separated time scales.

Tropical geometry is well adapted for studying robust and flexible, multiscale, biochemical networks. In [**NGVRv2**] we used the Litvinov-Maslov correspondence principle to tropicalize rational or polynomial ODE models of biochemical networks. The tropicalization make it possible to develop geometrical methods for critical parameter identification and for studying the qualitative dependence of the model dynamics on these parameters. These methods are based on arrangements of

2010 *Mathematics Subject Classification.* 14T05, 92C42.

Key words and phrases. systems biology, tropical geometry.

tropical manifolds, reminding combinatorial methods such as polyhedral complexes used in tropical convexity [**AD09**], or geometric analysis of S-systems proposed by [**SCF$^+$09**] in relation to biochemical network steady states design. Another application of the tropicalization is the detection of quasi-steady and quasi-equilibrium conditions, that are very useful for model reduction.

2. Settings

In chemical kinetics, the reagent concentrations satisfy ordinary differential equations:

$$\frac{dx_i}{dt} = F_i(\boldsymbol{x}),\ 1 \leq i \leq n. \tag{2.1}$$

Rather generally, the rates are rational functions of the concentrations and read

$$F_i(\boldsymbol{x}) = P_i(\boldsymbol{x})/Q_i(\boldsymbol{x}), \tag{2.2}$$

where $P_i(\boldsymbol{x}) = \sum_{\alpha \in A_i} a_{i,\alpha}\boldsymbol{x}^\alpha$, $Q_i(\boldsymbol{x}) = \sum_{\beta \in B_i} b_{i,\beta}\boldsymbol{x}^\beta$, are multivariate polynomials. Here $\boldsymbol{x}^\alpha = x_1^{\alpha_1}x_2^{\alpha_2}\dots x_n^{\alpha_n}$, $\boldsymbol{x}^\beta = x_1^{\beta_1}x_2^{\beta_2}\dots x_n^{\beta_n}$, $a_{i,\alpha}, b_{i,\beta}$, are nonzero real numbers, and A_i, B_i are finite subsets of $\mathbb{N}^n$.

Special case are represented by

$$F_i(\boldsymbol{x}) = P_i^+(\boldsymbol{x}) - P_i^-(\boldsymbol{x}), \tag{2.3}$$

where $P_i^+(\boldsymbol{x})$, $P_i^-(\boldsymbol{x})$ are Laurent polynomials with positive coefficients, $P_i^\pm(\boldsymbol{x}) = \sum_{\alpha \in A_i^\pm} a_{i,\alpha}^\pm \boldsymbol{x}^\alpha$, $a_{i,\alpha}^\pm > 0$, $A_i^\pm$ are finite subsets of $\mathbb{Z}^n$. Real powers $A_i^\pm \subset \mathbb{R}^n$ are sometimes used for the so-called S-systems [**SCF$^+$09**].

Litvinov and Maslov [**LMS01, LM96**] proposed a heuristic (correspondence principle) allowing to transform mathematical objects (integrals, polynomials) into their quantified (tropical) versions. According to this heuristic, to a Laurent polynomial with positive real coefficients $\sum_{\alpha \in A} a_\alpha \boldsymbol{x}^\alpha$, where $A \subset \mathbb{Z}^n$ is the support of the polynomial, one associates the max-plus polynomial $max_{\alpha \in A}\{log(a_\alpha) + < log(\boldsymbol{x}), \alpha >\}$. This heuristic can be used to associate a piecewise-smooth hybrid model to the system of rational ODEs (2.1), in two different ways.

The first method was proposed in [**NGVRv2**] and can be applied to any rational ODE system defined by (2.1),(2.2):

Definition 2.1. We call complete tropicalization of the smooth ODE system (2.1) the following piecewise-smooth system:

$$\frac{dx_i}{dt} = DomP_i(\boldsymbol{x})/DomQ_i(\boldsymbol{x}), \tag{2.4}$$

where $Dom\{a_{i,\alpha}\boldsymbol{x}^\alpha\}_{\alpha \in A_i} = sign(a_{i,\alpha_{max}})exp[max_{\alpha \in A_i}\{log(|a_{i,\alpha}|) + < \boldsymbol{u}, \alpha >\}]$. $\boldsymbol{u} = (logx_1, \dots, logx_n)$, and $a_{i,\alpha_{max}}$, $\alpha_{max} \in A_i$ denote the coefficient of the monomial for which the maximum is attained.

The second method,proposed in [**SCF$^+$09**], applies to the systems (2.1),(2.3).

Definition 2.2. We call two terms tropicalization of the smooth ODE system (2.1) the following piecewise-smooth system:

$$\frac{dx_i}{dt} = DomP_i^+(\boldsymbol{x}) - DomP_i^-(\boldsymbol{x}), \tag{2.5}$$

The two-terms tropicalization was used in [**SCF$^+$09**] to analyse the dependence of steady states on the model parameters. The complete tropicalization was used for the study of the model dynamics and for the model reduction [**NGVRv2**].

For both tropicalization methods, for each occurrence of the Dom operator, one can introduce a tropical manifold, defined as the subset of $\mathbb{R}^n$ where the maximum in Dom is attained at least twice. For instance, for $n = 2$, such tropical manifold is made of points, segments connecting these points, and half-lines. The tropical manifolds in such an arrangement decompose the space into sectors, inside which one monomial dominates all the others in the definition of the reagent rates. The combinatorial study of the arrangement give hints on the possible steady states and attractors, as well as on their bifurcations.

3. Justification of the tropicalization and some estimates

In the general case, the tropicalization heuristic is difficult to justify by rigorous estimates, however, this is possible in some cases. We state here some results in this direction. Let us consider the class of polynomial systems, corresponding to mass action law chemical kinetics:

$$\frac{dx_i}{dt} = F_i(\boldsymbol{x}, \epsilon) = \sum_{j=1}^{M} F_{ij}(\boldsymbol{x}, \epsilon), \quad F_{ij} = P_{ij}(\epsilon)\boldsymbol{x}^{\alpha_{ij}} \tag{3.1}$$

where α_{ij} are multi-indices, and ϵ is a small parameter. So, the right hand side of (3.1) is a sum of monomials. We suppose that coefficients P_{ij} have different orders in ϵ:

$$P_{ij}(\epsilon) = \epsilon^{b_{ij}}\hat{P}_{ij}, \tag{3.2}$$

where $b_{ij} \neq b_{i'j'}$ for $(i, j) \neq (i', j')$.

We also suppose that the cone $\mathbf{R}_> = \{x : x_i \geq 0\}$ is invariant under dynamics (3.1) and initial data are positive:

$$x_i(0) > \delta > 0.$$

The terms (3.2) can have different signs, the ones with $\hat{P}_{ij} > 0$ are production terms, and those with $\hat{P}_{ij} < 0$ are degradation terms.

From the biochemical point of view, the choice (3.2) is justified by the multi-scaleness of the biochemical processes. Furthermore, we are interested in biochemical circuits that can function "stably" even in extremal conditions. More precisely, we use the permanence concept, borrowed from the theory of species coexistence (the Lotka -Volterra model, see for instance [**Tak96**]).

Definition 3.1. The system (3.1) is permanent, if there are two constants $C_- > 0$ and $C_+ > 0$ such that

$$C_- < x_i(t) < C_+, \quad \textit{for all } t > T_0(x(0)) \textit{ and for every } i. \tag{3.3}$$

We assume that $C_\pm$ and T_0 are uniform in ϵ as $\epsilon \to 0$.

This means that concentrations of all the reagents cannot vanish or become too big, even in extremal conditions. Biological oscillators, such as the circadian clock and the cell cycle, satisfy this condition. We can also consider systems (3.1) that become permanent after rescaling of the concentrations, $x_i = \hat{x}_i\epsilon^{a_i}$, such as systems with quasi-stationary, low concentration, reagents.

For permanent systems, we can obtain some results justifying the two procedures of tropicalization. The complete tropicalization reads

$$\frac{d\bar{x}_i}{dt} = Dom(F_i(\bar{x})), \tag{3.4}$$

where $Dom(F_i) = F_{ik(i)}(\boldsymbol{x}, \epsilon), \quad |F_{ik(i)}(\boldsymbol{x}, \epsilon)| > |F_{ij}(\boldsymbol{x}, \epsilon)|, \quad j \neq k_i$ is the dominant term.

Notice that, because of the changing sign, (3.4) has discontinuous right hand side, therefore it is a differential inclusion. We assume here that there is no sliding motion. The situation with sliding motion needs special treatment and is discussed in the last section.

The two terms tropicalization reads

$$\frac{d\bar{x}_i}{dt} = Dom(F_i^+(\bar{x})) - Dom(F_i^-(\bar{x})) = Dom_2(F_i(\bar{x})), \tag{3.5}$$

where F_i^+, F_i^- gather the positive and negative terms of F_i, respectively.

PROPOSITION 3.2. *Assume that system* (3.1) *is permanent. Let us consider the Cauchy problem for (3.1) and (3.4) (or (3.5)), with the same initial data:*

$$x(0) = \bar{x}(0).$$

Then the difference $y(t) = x(t) - \bar{x}(t)$ *satisfies the estimate*

$$|y(t)| < C_1 \epsilon^\gamma \exp(bt), \quad \gamma > 0, \tag{3.6}$$

positive constant C_1, b *is uniform in* ϵ*. If the original system (3.1) is structurally stable in the domain* $\Omega_{C_-,C_+} = \{x : C_- < |x| < C_+\}$*, then the corresponding tropical systems (3.4) and (3.5) are also permanent and there is a orbital topological equivalency* h_ϵ *between the trajectories* $x(t)$ *and* $\bar{x}(t)$ *of the corresponding Cauchy problems close to the identity as* $\epsilon \to 0$*.*

The proof of estimate (3.6) follows immediately by the Gronwall lemma. The second assertion follows from the definition of structural stability.

Permanency property is not easy to check. One of the possible methods is to find an invariant domain I in $\mathbf{R}^n$ such that the vector field $\boldsymbol{F}$ is directed inward I at the boundary ∂I. It is clear that to find such a domain I is simpler for tropicalizations than for the original system (for instance, when dominant monomial ODEs can be integrated). Furthermore, if x is a solution of (3.1), then

$$\frac{dx_i}{dt} \leq K(\epsilon)|Dom(F_i(x, \epsilon))|, \tag{3.7}$$

where $K(\epsilon) = 1 + O(\epsilon^\gamma)$ for small ϵ and $K(1) = M$. Similarly, for the two term tropicalization, we have

$$\frac{dx_i}{dt} \leq K(\epsilon)Dom(F_i^+(x, \epsilon)) - Dom(F_i^-(x, \epsilon)). \tag{3.8}$$

In general, the inequalities (3.8) and (3.7) say nothing about $x(t)$. However, if the family of systems depending on a parameter K

$$\frac{dx_i}{dt} \leq KDom(F_i^+(x, \epsilon)) - Dom(F_i^-(x, \epsilon)). \tag{3.9}$$

defines monotone semiflows then the permanency of $\bar{x}(t)$ can be used to obtain permanency of $x(t)$. In practice, the monotonicity condition is rarely satisfied globally (global validity is incompatible with the possibility of oscillations, and can

not be satisfied by biological clocks), but can be satisfied locally, on some sectors bounded by tropical manifolds. Then, one can combine conditions on $\bar{x}(t)$ and conditions on $x(t)$ piecewisely, in order to prove permanency of $x(t)$.

4. Tropical sliding motions and model reduction

Biologists are attached to details. For the sake of completeness, systems biologists generate large, complex models. However, many details of these models are not important and can be simplified to facilitate model analysis. The model reduction problem is to find a simpler system, whose dynamics approximates the dynamics of the complex system [**RGZL08**]. Current model reduction techniques use quasi-equilibrium and quasi-steady state approximations. In both situations, the trajectories of some fast species satisfy approximate algebraic conditions imposed by the slow species. Given the trajectories $\boldsymbol{x}(t)$ of all species, we call imposed trajectory of the i-th species a real, positive, and stable solution $x_i^*(t)$ of the polynomial equation

$$P_i(x_1(t),\dots,x_{i-1}(t),x_i^*(t),x_{i+1}(t),\dots,x_n(t))=0, \tag{4.1}$$

We say that a species i is slaved if the distance between the trajectory $x_i(t)$ and some imposed trajectory $x_i^*(t)$ is small for some time interval I, $sup_{t\in I}|log(x_i(t))-log(x_i^*(t))|<\delta$, for some $\delta>0$ sufficiently small. The remaining species, that are not slaved, are called slow species.

Identifying slaved species is a first step of model reduction algorithms. Tropical geometry can be used for identification of slaved species without having to simulate the system and compute the trajectories. As first proposed in [**NGVRv2**], the existence of slaved species implies the existence of attractive sliding modes of the complete tropicalization, defined as stable motions on the tropical manifold. Attractive sliding modes are possible in the theory of piecewise-smooth systems [**FA88**] provided that the following condition is satisfied, for $\boldsymbol{x}$ on the tropical manifold:

$$< n_+(x), f_+(x) >< 0, \quad < n_-(x), f_-(x) >< 0 \tag{4.2}$$

where f_+, f_- are the dominant vector fields on the two sides of a tropical hypersurface and $n_+=-n_-$ are the normals to the interior faces.

References

[AD09] F. Ardila and M. Develin. Tropical hyperplane arrangements and oriented matroids. *Mathematische Zeitschrift*, 262(4):795–816, 2009.

[CTF06] G. Craciun, Y. Tang, and M. Feinberg. Understanding bistability in complex enzyme-driven reaction networks. *Proceedings of the National Academy of Sciences*, 103(23):8697–8702, 2006.

[Del49] M. Delbrück. Discussion: Unitées biologiques douées de continuité génétique. In *Actes du colloque international du CNRS*, pages 33–3, Paris, 1949. Editions du CNRS.

[FA88] A.F. Filippov and FM Arscott. *Differential equations with discontinuous righthand sides*, volume 18. Springer, 1988.

[GR07] A. N. Gorban and O. Radulescu. Dynamical robustness of biological networks with hierarchical distribution of time scales. *IET Systems Biology*, 1:238–246, 2007.

[LM96] G.L. Litvinov and V.P. Maslov. Idempotent mathematics: a correspondence principle and its applications to computing. *Russian Mathematical Surveys*, 51(6):1210–1211, 1996.

[LMS01] G.L. Litvinov, V.P. Maslov, and G.B. Shpiz. Idempotent functional analysis: an algebraic approach. *Mathematical Notes*, 69(5):696–729, 2001.

[NGVRv2] V. Noel, D. Grigoriev, S. Vakulenko, and O. Radulescu. Tropical geometries and dynamics of biochemical networks. Application to hybrid cell cycle models. *Electronic Notes in Theoretical Computer Science*, to appear, arXiv:1109.4085v2.

[RGZL08] O. Radulescu, A. N. Gorban, A. Zinovyev, and A. Lilienbaum. Robust simplifications of multiscale biochemical networks. *BMC Systems Biology*, 2(1):86, 2008.

[RLS+06] O. Radulescu, S. Lagarrigue, A. Siegel, P. Veber, and M. Le Borgne. Topology and static response of interaction networks in molecular biology. *Journal of The Royal Society Interface*, 3(6):185–196, 2006.

[RSP+11] O. Radulescu, A. Siegel, E. Pécou, C. Chatelain, and S. Lagarrigue. Genetically regulated metabolic networks: Gale-Nikaido modules and differential inequalities. *Transactions on computational systems biology XIII*, pages 110–130, 2011.

[SCF+09] M.A. Savageau, P.M.B.M. Coelho, R.A. Fasani, D.A. Tolla, and A. Salvador. Phenotypes and tolerances in the design space of biochemical systems. *Proceedings of the National Academy of Sciences*, 106(16):6435, 2009.

[Son05] E.D. Sontag. Molecular systems biology and control. In *Control 11: 396–435. of Boolean networks 15*. Citeseer, 2005.

[Sou03] C. Soulé. Graphic requirements for multistationarity. *Complexus*, 1(123-133), 2003.

[T+98] R. Thomas et al. Laws for the dynamics of regulatory networks. *International Journal of Developmental Biology*, 42:479–485, 1998.

[Tak96] Y. Takeuchi. *Global dynamical properties of Lotka-Volterra systems*. World Scientific Pub Co Inc, 1996.

[VRv2] S. Vakulenko and O. Radulescu. Flexible and robust patterning by centralized gene networks. *Fundamenta Informaticae*, to appear, arXiv:1110.4724v2.

DIMNP UMR CNRS 5235, University of Montpellier 2, Montpellier, France

CNRS, Mathématiques, Université de Lille, 59655, Villeneuve d'Ascq, France

IRMAR UMR 6625, University of Rennes 1, Rennes, France

Saint Petersburg State University of Technology and Design, St.Petersburg, Russia