Algebraic tools for networks of interacting elements

Elisenda Feliu

My research focuses on the development of mathematical theory to analyse mathematical models that arise from studying systems of (bio)chemical reactions and also to formalise and make sense of biologists' intuition.

My main object of interest is the polynomial system associated with a reaction network. A reaction in the mathematical sense can be seen as an "arrow" between linear combinations of species. For example, the representation

$$A + B \to C$$

indicates that two species A, B react to form a third species C. Then a **reaction network** is simply a finite number of reactions of this type, or, with other words, a directed graph $(\mathcal{C}, \mathcal{R})$ whose nodes are (nonnegative and integer) linear combinations of species in a set $\mathcal{X} = \{X_1, \ldots, X_n\}$. For example, the following are reaction networks

$$X_1 \xrightarrow{\kappa_1} X_2$$
$$X_1 + X_2 \xrightarrow{\kappa_3} 2X_2 \xrightarrow{\kappa_2} 2X_1$$

$$X_1 \xrightarrow{\kappa_1} X_2$$
$$X_2 + X_3 \xrightarrow{\kappa_2} X_1 + X_4$$
$$X_4 \xrightarrow{\kappa_3} X_3.$$

The labels written on the edges of the networks are positive numbers called *reaction rate constants*, and have an important role on modelling how the concentrations of the species in the network change with time.

Given a reaction network as one of the above, one associates an autonomous ODE system that models the concentration x_i of species X_i in time t. The most general modelling framework is that of **mass-action**, which is constructed as follows. We denote a generic reaction as $y \to y'$, and its reaction rate constant as $\kappa_{y \to y'} > 0$. Recall that y is a linear combination of the species, and hence we can identify $y = y_1 X_1 + \cdots + y_n X_n$ with the vector $(y_1, \ldots, y_n)$ in $\mathbb{N}^n$. Then, given $x = (x_1, \ldots, x_n) \in \mathbb{R}_{\geq 0}^n$, we let $x^y = x_1^{y_1} \cdot \ldots \cdot x_n^{y_n}$, with the understanding that $0^0 = 1$. The *mass-action system* associated with the network is

$$\dot{x}_i = \sum_{y \to y' \in \mathcal{R}} \kappa_{y \to y'} (y' - y)_i x^y, \quad i = 1, \ldots, n, \quad x \in \mathbb{R}_{\geq 0}^n,$$

(where $\dot{x}_i = \frac{dx_i}{dt}$). The right-hand side of this expression is a polynomial in x with monomials x^y and coefficients $\kappa_{y \to y'}(y' - y)_i$. Note that $(y' - y)_i$ is precisely the net production of species X_i in the reaction. For example, the two mass-action systems associated with the networks above are respectively:

$$\dot{x}_1 = -\kappa_1 x_1 + \kappa_2 x_2^2 - \kappa_3 x_1 x_2$$
$$\dot{x}_2 = \kappa_1 x_1 - \kappa_2 x_2^2 + \kappa_3 x_1 x_2.$$

$$\dot{x}_1 = -\kappa_1 x_1 + \kappa_2 x_2 x_3$$
$$\dot{x}_2 = \kappa_1 x_1 - \kappa_2 x_2 x_3$$
$$\dot{x}_3 = -\kappa_2 x_2 x_3 + \kappa_3 x_4$$
$$\dot{x}_4 = \kappa_2 x_2 x_3 - \kappa_3 x_4.$$

This modelling framework is applied within chemistry, where species are chemical species, in biochemistry, where species are often proteins, mRNA or similar, in epidemiology, where species

are infected, ill, etc people (the standard SIR model is of this type above), or even in ecology, where species are animals or plants (the famous Lotka-Volterra model for a predator-prey system is also of this type).

My interest in studying these systems arises from biochemistry, where, in practice, the values of the reaction rate constants $\kappa_{y \to y'} > 0$ are **unknown**. One strategy to study these models is to try to estimate the values of these parameters from data arising from experiments. However, these systems have often many variables and parameters, making this task difficult, specially if one takes into account that there are experimental (measurement) errors and errors arising from the choice of model.

To bypass this, one can alternatively take a **qualitative approach**, that is, study the system for *all* parameter values. For instance, it is not hard to show that for the mass-action system on the right-hand side of the display above, all trajectories converge to an equilibrium.

This is what my research focuses on: providing tools to decide what properties a given mass-action system has. I work mainly on properties related to equilibria. Since, as mentioned, the ODE systems under consideration are polynomial, the equilibria are the *nonnegative* zeros of the system of **polynomial equations**:

$$0 = \sum_{y \to y'} \kappa_{y \to y'} (y' - y)_i x^y, \quad i = 1, \dots, n, \quad x \in \mathbb{R}^n_{\geq 0}.$$

My work is thus of **algebraic nature**. Since classic algebraic geometry deals mainly with varieties over the complex numbers (or algebraically closed fields), or, to some extend, over the real numbers without restricting to the nonnegative part, classical tools like Gröbner bases are of limited use. My research aims thus at exploiting the specific structure of the polynomial systems under consideration to be able to study the equilibria.

Since the end-users of the tools we develop are biologists/biochemists, it is not enough to come up with nice theorems that are not constructive or such that it is not easy to verify the assumptions required to apply them. Ideally, my (dream) **goal** would be to have a "black-box" (not requiring user intervention), whose input is a reaction network, and output a list of properties of the system, something we could call a **qualitative profile**:
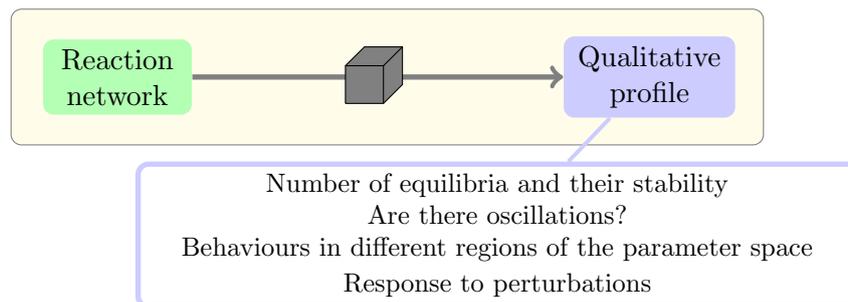
We want this black-box to give exact answers not relying on numerical simulations nor parameter sampling. The relevance of such a tool is two-sided:

- On one hand, it can be used in *synthetic biology*, where small organisms are created *de novo* in the lab. One could use a qualitative profile to screen several networks before deciding which ones to synthesise in the lab.

- On the other, it can be used in *systems biology* to assess the validity of a model or mechanism, or to design experiments to discriminate between two models. That is, given two

candidate networks for some cell process under investigation, one could run the two profiles, find the differences, and design experiments targeting the differential aspects of the two models:
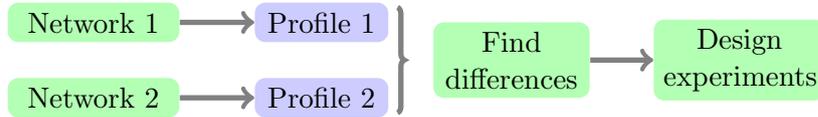
The truth is that we are very far from this dream goal, but progress is being done. There are several black-box methods that address for example whether there are at least two steady states (a question which is relevant to understand with applications in the study of cell decision making). The research questions are easy to explain, but difficult to solve. So, we need bright mathematicians to join in. *Come and knock on my door!*