Geometric analysis of pathways dynamics: application to versatility of TGF-β receptors

Satya Swarup Samal, Aurélien Naldi, Dima Grigoriev, Andreas Weber, Nathalie Théret, Ovidiu Radulescu

Geometric analysis of pathways dynamics: application to versatility of TGF-β receptors

Satya Swarup Samal[a], Aurélien Naldi[f], Dima Grigoriev[b], Andreas Weber[c], Nathalie Théret[d,e], Ovidiu Radulescu[f,*]

[a] *Algorithmic Bioinformatics, Bonn-Aachen International Center for IT, Dahlmannstraße 2, D-53113, Bonn, Germany*
[b] *CNRS, Mathématiques, Université de Lille, Villeneuve d'Ascq, 59655, France*
[c] *Institut für Informatik II, University of Bonn, Friedrich-Ebert-Allee 144, 53113 Bonn, Germany*
[d] *Inserm UMR1085 IRSET, Université de Rennes 1, Rennes, France*
[e] *CNRS-Université de Rennes1-INRIA, UMR6074 IRISA, Rennes, France*
[f] *DIMNP UMR CNRS 5235, University of Montpellier, Montpellier, France*

Abstract

We propose a new geometric approach to describe the qualitative dynamics of chemical reactions networks. By this method we identify metastable regimes, defined as low dimensional regions of the phase space close to which the dynamics is much slower compared to the rest of the phase space. These metastable regimes depend on the network topology and on the orders of magnitude of the kinetic parameters. Benchmarking of the method on a computational biology model repository suggests that the number of metastable regimes is sub-exponential in the number of variables and equations. The dynamics of the network can be described as a sequence of jumps from one metastable regime to another. We show that a geometrically computed connectivity graph restricts the set of possible jumps. We also provide finite state machine (Markov chain) models for such dynamic changes. Applied to signal transduction models, our approach unravels dynamical and functional

*corresponding author
Email addresses: samal@cs.uni-bonn.de (Satya Swarup Samal), aurelien.naldi@gmail.com (Aurélien Naldi), dmitry.grigoryev@math.univ-lille1.fr (Dima Grigoriev), weber@cs.uni-bonn.de (Andreas Weber), nathalie.theret@univ-rennes1.fr (Nathalie Théret), ovidiu.radulescu@univ-montp2.fr (Ovidiu Radulescu)

capacities of signalling pathways, as well as parameters responsible for specificity of the pathway response. In particular, for a model of TGFβ signalling, we find that the ratio of TGFBR2 to TGFBR1 receptors concentrations can be used to discriminate between metastable regimes. Using expression data from the NCI60 panel of human tumor cell lines, we show that aggressive and non-aggressive tumour cell lines function in different metastable regimes and can be distinguished by measuring the relative concentrations of receptors of the two types.

Keywords: Tropical Geometry, Cancer Systems Biology, Finite State Automaton, Metastability

---

1. Introduction

Networks of biochemical reactions are used in computational biology as models of signalling, metabolism, and gene regulation. For various applications it is important to understand how the dynamics of these models depend on internal parameters, initial data and environment variables. Traditionally, the dynamics of biochemical networks is studied in the framework of chemical kinetics that can be either deterministic (ordinary differential equations) or stochastic (continuous time Markov processes). In order to cope with qualitative data, boolean (Kauffman, 1969; Thomas, 1973) or multi-valued networks (Thomas, 1991; Naldi et al., 2007) are used instead of continuous models. For this reason, many efforts were focused on coarse graining dynamical networks described by ordinary differential equations (ODE) to boolean networks (Davidich and Bornholdt, 2008). However, in spite of their advantages, dynamical properties of large Boolean or multi-valued networks are still difficult to study. The difficulty originates from the number of possible states, which for multi-valued networks with m levels (Boolean networks correspond to $m = 2$) is m^n. Although there are efficient methods for computing attractors of Boolean networks (algorithms based on binary decision diagrams or on satisfiability solvers can handle synchronous networks with hundreds of variables (Dubrova and Teslenko, 2011)), more intricate dynamical properties like the metastable regimes discussed in this paper, ask for comprehensive analysis of the state transition graph which is hard to perform and moreover for analysis of the hierarchy of time scales which is even harder for Boolean and multi-valued networks. The coarse graining method proposed in this paper leads to a drastic reduction of the number of states.

This offers unprecedented possibilities for qualitative analysis of the dynamics.

In this paper we propose a new method for model analysis that uses coarse grained descriptions of continuous dynamics as discrete automata defined on finite states. These states will not be obtained by discretization of network variables, but by identification of a finite number of collective modes describing possible coordinated activity of several variables.

For large networks with ordinary differential equations dynamics and multiple timescales it is reasonable to consider the following property: a typical trajectory consists in a succession of qualitatively different slow segments separated by faster transitions. The slow segments, generally called metastable states or regimes, can be of several types such as attractive invariant manifolds (Gorban and Karlin, 2005), Milnor attractors (Rabinovich et al., 2006) or saddle connections (Rabinovich et al., 2012). The notion of metastability generalizes the notion of attractor. Like in the case of attractors, distant parts of the system can have coordinated activity for metastability. The dynamical states of large networks can be represented as points in a high dimensional space, called phase space. In this representation each coordinate represents the concentration of a molecular species. Coordinated activity means that many of the species concentrations are correlated, which can be geometrically represented by the proximity to a lower dimension hypersurface in the phase space. A system remains in the proximity of an attractor after entering its basin of attraction, but can leave a metastable regime after a relatively long time (much longer than the time needed for transitions between two different regimes). Fig. 1 summarizes this geometrical picture. The term *crazy-quilt* was coined to describe such a patchy landscape of multiscale networks dynamics (Gorban and Radulescu, 2008).

Fig. 1 here

This phenomenon, called itinerancy, received particular attention in neuroscience (Tsuda, 1991). Itinerant behaviour is shown by mathematical models with stable heteroclinic sequences (defined as open chains of saddle fixed points connected by one-dimensional separatrices) and was also observed in transient activity of antennal lobe neurons involved in insect olfaction or in the activity of high vocal centers controlling songbird patterns (Rabinovich et al., 2006). We believe that similar phenomena occur in molecular biology for chemical reaction networks. A well studied example sustaining this picture is the biochemical and gene expression dynamics guiding the orderly progression of the cell cycle. The main feature of this dynamics is the sequential

activation of cyclin dependent kinase/cyclin complexes. More than 30 years since cyclins were discovered the main cell cycle control events are now well understood and it is agreed that each of them involve the collective action of several regulator molecules. In addition, studies of periodic gene expression in synchronized cell division of yeast show the existence of waves of coordinate expression corresponding to different cell cycle phases or transitions (Rustici et al., 2004). Furthermore, mathematical models of the cell cycle machinery (currently more than 150 published models (Weis et al., 2014)) illustrate the stage dependent coordination of biochemical variables. As an example, the structure of steady state branches of the Wee1-Cdc25 module in yeast lead Tyson et al. (2002) to consider that the exit from mitosis is a collective phenomenon that can be described as a saddle-node bifurcation. Our analysis of such models also showed that branches and bifurcations of states occur naturally in the context of cell cycle modelling (Noel et al., 2012). In a more general context, geometric analysis of single-cell expression data from human and mouse tissues showed that gene expression is structured in clusters but also in continua of states within polyhedra whose vertices can be understood as specialized key tasks (Korem et al., 2015). These findings were interpreted in terms of multi-objective optimization solutions (Korem et al., 2015), but could also suggest transient behaviour between specialized states. The idea of associating cell lineage commitment to collective behaviour of gene networks variables was used in many other contexts including cancer genomics where it was proposed that cancer cells are trapped in some abnormal attractors (Huang et al., 2009).

In this paper we propose a method to compute metastable dynamical regimes and the transitions between such regimes for chemical reaction networks. This will provide a precise meaning to the "crazy-quilt" metaphor illustrated in Fig. 1. To this aim we will use tropical geometry methods. Tropical methods (Litvinov, 2007; Maclagan and Sturmfels, 2015), also known as idempotent or max-plus algebras owe their name to the fact that one of the pioneer of the field, Imre Simon, was Brazilian. These methods found numerous applications to computer science (Simon, 1988), physics (Litvinov, 2007), railway traffic (Chang, 1998), and statistics (Pachter and Sturmfels, 2004). Recently we have applied these methods to model order reduction (Noel et al., 2012, 2014; Radulescu et al., 2015b; Samal et al., 2015b). In these works we have used tropical methods to rank monomial terms into rate vectors according to their orders of magnitude and to identify lowest order, dominant terms. When there is only one dominant term or when the domi-

nant terms have all the same sign, the dynamics is fast and the system tends rapidly towards a region in phase space where at least two dominant terms of opposite signs are equilibrated. We have called the latter situation tropical equilibration (Noel et al., 2014; Radulescu et al., 2015b; Samal et al., 2015b). In this paper, we use tropical equilibrations to identify metastable dynamic regimes of chemical reaction networks. We show that tropical equilibrations can be grouped into branches and describe the qualitative network dynamics as a sequence of transitions from one branch to another. The complexity of the qualitative dynamics depends on the number of branches. Therefore, we would like to know how this number depends on the number of chemical species. Although there are theoretical results suggesting that the number of branches should be small, these results are valid in the average in the probabilistic space of all the models. In order to test this property numerically we will compute the branches for a large collection of models of the Biomodels database (Le Novere et al., 2006). The idea to gather tropical equilibrations into branches was introduced in our CMSB 2015 conference paper (Radulescu et al., 2015a). As new work with respect to this paper, we propose the benchmarking of the method on the Biomodels database. We also investigate the biological significance of metastable states in the context of TGFβ signalling in cancer cell lines.

The structure of the paper is as follows. In the *Theory and Methods* section we introduce the branches of tropical equilibrations and discuss briefly how they can be calculated. We also propose an algorithm to learn Markov state models defined on branches of tropical equilibrations. The *Results* section has two parts. First, we apply the computation of branches to models in the Biomodels database. Then, we apply the method to a model of TGFβ signalling and show how the analysis can be used to interpret biological data.

2. Theory and Methods

2.1. Tropical equilibrations of chemical reactions networks with polynomial rate functions

In this section we introduce the main concepts relating geometry and dynamics.

We consider chemical reaction networks described by mass action kinetics. To describe such a reaction network we need the list of components, $\mathcal{A} =$

$\{A_1, \ldots A_n\}$ and the list of reactions (the reaction mechanism):

$$\sum_i \alpha_i^j A_i \rightleftharpoons \sum_k \beta_k^j A_k, \tag{1}$$

where $j \in [1, r]$ is the reaction number, α_i^j and β_k^j are positive integers defining the reaction stoichiometry.

Then, the mass action kinetics is described by a system of differential equations:

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = \sum_{j=1}^{r} k_j S_{ij} x^{\alpha_j},\ 1 \leq i \leq n, \tag{2}$$

where $k_j > 0$ are kinetic constants, x_i are variable concentrations, $S_{ij} = \beta_i^j - \alpha_i^j$ are the entries of the stoichiometric matrix, $\alpha_j = (\alpha_1^j, \ldots, \alpha_n^j)$ are multi-indices, and $x^{\alpha_j} = x_1^{\alpha_1^j} \ldots x_n^{\alpha_n^j}$.

For our reasonings, we can replace the exact values of parameters by their orders of magnitude that are supposed to be known. Usually, orders of magnitude are approximations of the parameters by integer powers of ten and serve for rough comparisons. Our definition of orders of magnitude is based on the equation $k_j = \bar{k}_j \varepsilon^{\gamma_j}$, where ε is a small positive number. The exponents γ_j are considered to be integer or rational. For instance, the approximation

$$\gamma_j = \mathrm{round}(\log(k_j)/\log(\varepsilon)), \tag{3}$$

produces integer exponents, whereas $\gamma_j = \mathrm{round}(d \log(k_j)/\log(\varepsilon))/d$ produces rational exponents, where round stands for the closest integer (with half-integers rounded to even numbers) and d is a strictly positive integer. When $\varepsilon = 1/10$, our definition provides the usual decimal orders.

In this study, orders of magnitude of the kinetic parameters are supposed to be known. In contrast, species orders vary in time and have to be computed. To this aim, the species concentrations are first represented by orders of magnitude defined as

$$a_j = \log(x_j)/\log(\varepsilon). \tag{4}$$

More precisely, one has $x_j = \bar{x}_j \varepsilon^{a_j}$, where $\bar{x}_j$ has zero order (unity). Because $\log(\varepsilon) < 0$, Eq.(4) means that species orders and concentrations are anti-correlated (large orders mean small concentrations and vice versa).

Then, network dynamics is described by a rescaled ODE system

$$\frac{\mathrm{d}\bar{x}_i}{\mathrm{d}t} = \sum_j \varepsilon^{\mu_j(a) - a_i} \bar{k}_j S_{ij} \bar{x}^{\alpha_j}, \tag{5}$$

where $\mu_j(a) = \gamma_j + \langle a, \alpha_j \rangle$ (6), and $\langle , \rangle$ stands for the dot product.

The r.h.s. of each equation in (5) is a sum of multivariate monomials in the concentrations. The orders μ_j indicate how large these monomials are in absolute value. A monomial of order μ_j dominates another monomial of order $\mu_{j'}$ if $\mu_j < \mu_{j'}$.

To set these ideas down let us use a simple chemical network example, the Michaelis-Menten kinetics:

$$S + E \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \overset{k_2}{\to} P + E,$$

where S, ES, E, P represent the substrate, the enzyme-substrate complex, the enzyme and the product, respectively.

After using the two conservation laws $E + ES = e_0$ and $S + ES + P = s_0$, we find

$$\begin{aligned} \dot{x}_1 &= -k_1 x_1 (e_0 - x_2) + k_{-1} x_2, \\ \dot{x}_2 &= k_1 x_1 (e_0 - x_2) - (k_{-1} + k_2) x_2. \end{aligned} \tag{7}$$

where x_1, x_2 are the concentrations of S and ES respectively.

Orders of variables and parameters are as follows $x_i = \bar{x}_i \varepsilon^{a_i}$, $1 \leq i \leq 2$, $k_1 = \bar{k}_1 \varepsilon^{\gamma_1}$, $k_{-1} = \bar{k}_{-1} \varepsilon^{\gamma_{-1}}$, $e_0 = \bar{e}_0 \varepsilon^{\gamma_e}$.

The tropical equilibration problem consists in the equality of the orders of at least two monomials one positive and another negative in the differential equations of each species. This condition allows us to compute the concentration orders defined by (4). More precisely, we want to find a vector a such that

$$\min_{j, S_{ij} > 0} (\gamma_j + \langle a, \alpha_j \rangle) = \min_{j, S_{ij} < 0} (\gamma_j + \langle a, \alpha_j \rangle) \tag{8}$$

The equation(8) is related to the notion of *tropical hypersurface*. A *tropical hypersurface* is the set of vectors $a \in \mathbb{R}^n$ such that the minimun $\min_{j, S_{ij} \neq 0} (\gamma_j + \langle a, \alpha_j \rangle)$ is attained for at least two different indices j (with no sign conditions). *Tropical prevarieties* are finite intersections of tropical hypersurfaces. Therefore, our tropical equilibrations are subsets of tropical prevarieties (Maclagan and Sturmfels, 2015). The sign condition in (8) was imposed because species concentrations are real positive numbers. A sum of positive monomials can not be zero for positive real values of the variables.

The system (8) can be seen as a system of equations in min-plus algebra (also known as tropical semiring), where multiplication $\otimes$ is the real numbers addition $x \otimes y = x + y$ and the addition $\oplus$ is the minimum operation $x \oplus y = \min(x, y)$. In order to find the solutions of such system we can explore combinatorially trees of solutions resulting from various choices of

minimal terms and write down inequations for each situation. Because a set of inequations define a polyhedron, the set of tropical equilibration solutions forms a polyhedron in $\mathbb{R}^n$. As a matter of fact, computing tropical equilibrations from the orders of magnitude of the model parameters is a NP-complete problem (Theobald, 2006) and brute force calculation by exploration of combinatorics has exponential complexity. However, methods based on the Newton polytope (Samal et al., 2014) or constraint logic programming (Soliman et al., 2014) exploit the sparseness and redundancy of the system and reduce the combinatorics and the time to compute tropical solutions.

The tropical equilibration equations for the Michaelis-Menten example are obtained by equating minimal orders of positive monomials with minimal orders of negative monomials in (7):

$$\gamma_1 + \gamma_e + a_1 = \min(\gamma_1 + a_1, \gamma_{-1}) + a_2, \tag{9}$$

$$\gamma_1 + \gamma_e + a_1 = \min(\gamma_1 + a_1, \min(\gamma_{-1}, \gamma_2)) + a_2. \tag{10}$$

Species timescales. The timescale of a variable x_i is given by $\frac{1}{x_i}\frac{\mathrm{d}x_i}{\mathrm{d}t} = \frac{1}{\bar{x}_i}\frac{\mathrm{d}\bar{x}_i}{\mathrm{d}t}$ whose order is

$$\nu_i = \min\{\mu_j | S_{ij} \neq 0\} - a_i. \tag{11}$$

The order ν_i indicates how fast is the variable x_i (if $\nu_{i'} < \nu_i$ then $x_{i'}$ is faster than x_i) .

Partial tropical equilibrations. It is useful to extend the tropical equilibration problem to partial equilibrations, that means solving (8) only for a subset of species. This is justified by the fact that slow species do not need to be equilibrated. In order to have a self-consistent calculation we compute the species timescales by (11). A partial equilibration is *consistent* if $\nu_i < \nu$ for all non-equilibrated species i. $\nu > 0$ is an arbitrarily chosen threshold indicating the timescale of interest.

Tropical equilibrations, slow invariant manifolds and metastable dynamic regimes. In dissipative systems, fast variables relax rapidly to some low dimensional attractive manifold called invariant manifold (Gorban and Karlin, 2005) that carries the slow mode dynamics. A projection of dynamic equations onto this manifold provides the reduced dynamics (Maas and Pope, 1992). This simple picture can be complexified to cope with hierarchies of invariant manifolds and with phenomena such as transverse instability, excitability and itineracy. Firstly, the relaxation towards an attractor can have several stages, each with its own invariant manifold. During relaxation towards the attractor, invariant manifolds are usually embedded one

into another (there is a decrease of dimensionality) (Chiavazzo and Karlin, 2011). Secondly, invariant manifolds can lose local stability, which allow the trajectories to perform large phase space excursions before returning in a different place on the same invariant manifold or on a different one (Haller and Sapsis, 2010). We showed elsewhere that tropical equilibrations can be used to approximate invariant manifolds for systems of polynomial differential equations (Noel et al., 2012, 2014; Radulescu et al., 2015b). Indeed, tropical equilibration are defined by the cancelling out of dominant forces acting on the system. The remaining weak non-compensated forces ensure the slow dynamics on the invariant manifold. Tropical equilibrations are thus different from steady states, in that there is a slow dynamics. In this paper we will use them as proxies for metastable dynamic regimes.

More precisely, let us assume that species timescales (defined by Eq.(11)) satisfy the relation $\nu_1 \leq \nu_2 \leq \ldots \leq \nu_n$ and that not all the timescales are the same, i.e. there is $m < n$ such that $\nu_{m+1} - \nu_m > 0$. Then, two groups of variables have separated timescales. The variables $X_r = (x_1, x_2, \ldots, x_m)$ are fast (change significantly on timescales of order of magnitude $\varepsilon^{-\nu_m}$ or shorter. The remaining variables $X_s = (x_{m+1}, x_{m+2}, \ldots, x_n)$ are much slower (have little variation on timescales of order of magnitude $\varepsilon^{-\nu_m}$). The metastable regime means that fast variables have reached quasi-steady state values on a low dimensional hypersurface of the phase space.

Branches of tropical equilibrations and connectivity graph. For each equation i, let us define

$$M_i(a) = \underset{j}{\operatorname{argmin}}(\mu_j(a), S_{ij} > 0) = \underset{j}{\operatorname{argmin}}(\mu_j(a), S_{ij} < 0), \tag{12}$$

in other words M_i denotes the set of monomials having the same minimal order μ_i. We call *tropically truncated system* the system obtained by pruning the system (5), i.e. by keeping only the dominating monomials.

$$\frac{\mathrm{d}\bar{x}_i}{\mathrm{d}t} = \varepsilon^{\mu_i - a_i}\Big(\sum_{j \in M_i(a)} \bar{k}_j \nu_{ji} \bar{x}^{\alpha_j}\Big), \tag{13}$$

The tropical truncated system is uniquely determined by the index sets $M_i(a)$, therefore by the tropical equilibration a. Reciprocally, two tropical equilibrations can have the same index sets $M_i(a)$ and truncated systems. We say that two tropical equilibrations a_1, a_2 are equivalent iff $M_i(a_1) = M_i(a_2)$, for all i. Equivalence classes of tropical equilibrations are called *branches*. A branch B with an index set M_i is *minimal* if $M_i' \subset M_i$ for all i where M_i' is the index set of the branch B' implies $B' = B$ or $B' = \emptyset$.

Closures of equilibration branches are defined by a finite set of linear inequalities, which means that they are polyhedral complexes. Minimal branches correspond to maximal dimension faces of the polyhedral complex. The incidence relations between the maximal dimension faces ($n-1$ dimensional faces, where n is the number of variables) of the polyhedral complex define the *connectivity graph*. More precisely, minimal branches are the vertices of this graph. Two minimal branches are connected if the corresponding faces of the polyhedral complex share a $n-2$ dimensional face. In terms of index sets, two minimal branches with index sets M and M' are connected if there is an index set M'' of an existing non-minimal branch such that $M'_i \subset M''_i$ and $M_i \subset M''_i$ for all i.

Returning to the Michaelis-Menten example let us analyse the quasi-equilibrium situation (Meiske, 1978; Segel, 1988; Segel and Slemrod, 1989; Gorban et al., 2010; Gorban and Shahzad, 2011) when the reaction constant k_{-1} is much faster than the reaction constant k_2. In terms of orders, this condition reads $\gamma_{-1} < \gamma_2$. In this case, the two tropical equilibration equations (9), (10) are identical, because $\min(\gamma_{-1}, \gamma_2) = \gamma_{-1}$. Let $\gamma_m = \gamma_{-1} - \gamma_1$ denote the order of the parameter $K_m = k_{-1}/k_1$. There are two branches of solutions of (9), namely $a_2 = \gamma_e, a_1 \leq \gamma_m$ and $a_2 = a_1 + \gamma_e - \gamma_m, a_1 \geq \gamma_m$ corresponding to $\min(\gamma_1 + a_1, \gamma_{-1}) = \gamma_1 + a_1$ and to $\min(\gamma_1 + a_1, \gamma_{-1}) = \gamma_{-1}$, respectively. Using the relation between orders and concentrations we identify the first branch of solutions with the saturation regime $x_2 \approx e_0$ (the free enzyme is negligible) and $x_1 >> K_m$ (the substrate has large concentration) and the second branch with the linear regime $x_2 << e_0$ (the concentration of the attached enzyme is negligible) and $x_1 << K_m$ (the substrate has low concentration).

The fast truncated system (obtained after removing all dominated monomials from (7)) reads

$$\begin{aligned} \dot{x}_1 &= -k_1 x_1 e_0 + k_{-1} x_2, \\ \dot{x}_2 &= k_1 x_1 e_0 - k_{-1} x_2, \end{aligned} \tag{14}$$

for the linear regime branch and

$$\begin{aligned} \dot{x}_1 &= -k_1 x_1 (e_0 - x_2), \\ \dot{x}_2 &= k_1 x_1 (e_0 - x_2), \end{aligned} \tag{15}$$

for the saturated regime branch.

2.2. Learning a finite state machine from a nonlinear biochemical network

Once the metastable regimes are computed, one can try to reconstruct

the transitions happening between these regimes as illustrated in Fig. 1. By abstraction, we can consider that the metastable regimes (patches in the "crazy-quilt" picture in Fig.1) are represented as nodes of a graph. Two nodes in the graph are connected if and only if there is at least one transition from one regime to the other. We call this abstraction a finite state machine, because the number of regimes is finite. However, given one regime, there may be several different possibilities to leave this regime, each leading to a different metastable regime. Which transition is chosen depends on the initial data. Thus, although the initial model is deterministic the finite state machine abstraction is generally stochastic. From one node can leave several transitions each one having a different probability per unit time. Because the initial system spends a long time on the metastable regime and little time on the transitions, it is natural to expect that the memory of previous transitions is lost and that the stochastic process is Markovian. In the following we propose a method to learn this stochastic, Markov process, from many simulations of the full model, each one starting from a different, randomly chosen, initial state. The method is similar to the Markov state models (Bowman et al., 2009) used to coarse grain phase space for protein folding molecular dynamics. Contrary to energy landscape in protein folding, our "crazy-quilt" is not that rugged as it consists in a small number of patches. Therefore, standard Monte Carlo procedures function well in our case.

In order to construct the finite state machine, we first need a way to map the phase space of the continuous model to a finite set of branches. First, we compute the branches of tropical solutions as subsets of the euclidian space $\mathbb{R}^n$ where n is the number of variables. We are using the algorithm based on constraint solving introduced by (Soliman et al., 2014) to obtain all rational tropical equilibration solutions $a = (a_1, a_2, \ldots, a_n)$ within a box $|a_i| < b$, $b > 0$ and with denominators smaller than a fixed value d, $a_i = p_i/q$, p_i, q are positive integers, $q < d$. The output of the algorithm is a matrix containing all the tropical equilibrations within the defined bounds. A post-processing treatment is applied to this output consisting in computing truncated systems, index sets, and minimal branches. Tropical equilibrations minimal branches are stored as matrices $A_1, A_2, \ldots, A_b$, whose lines are tropical solutions within the same branch. Here b is the number of minimal branches.

Our method computes numerical approximations of the tropical prevariety. Given a value of ε, this approximation is better when the denominator bound d is high. At fixed d, the dependence of the precision on ε follows

more intricate rules dictated by Diophantine approximations. For this reason, we systematically test that the number b and the truncated systems corresponding to minimal branches are robust when changing the value of ε.

Trajectories $x(t) = (x_1(t), \ldots, x_n(t))$ of the smooth dynamical system are generated with different initial conditions, chosen uniformly. For each time t, we compute the Euclidian distance $d_i(t) = \min_{y \in A_i} \|y - log_\varepsilon(x(t))\|$, where $\|*\|$ denotes the Euclidean norm and $\log_\varepsilon(x) = (\log x_1/\log(\varepsilon), \ldots, \log x_n/\log(\varepsilon))$. This distance classifies all points of the trajectory as belonging to a tropical minimal branch. The result is a symbolic trajectory $s_1, s_2, \ldots$ where the symbols s_i belong to the set of minimal branches. In order to include the possibility of transition regions we include an unique symbol t to represent the situations when the minimal distance is larger than a fixed threshold. The choice of this threshold is robust (see discussion in Section 3.2 and Fig.4). We also store the residence times $\tau_1, \tau_2, \ldots$ that represent the time spent in each of the state.

Special care should be taken when the model has a number of conservations laws. A conservation law is a linear combination of species concentrations that is kept constant during the dynamics, in other words an equation of the type $c_1^i x_1 + c_2^i x_2 + \ldots + c_n^i x_n = K_i$. We suppose that the conservation laws are semi-positive (all c_i are positive or nought). Then, several such equations together with the positivity conditions $x_i \geq 0$ define a polyhedron. We want to pick initial conditions for the trajectories $x(t)$ uniformly in such a polyhedron. To start, each component j can take its initial concentration in the range $[0, \min(K_i/c_j^i, \text{ such that } c_j^i > 0)]$, but as we set the initial concentration of one component, the range available for components which are involved in the same constraint is reduced. The last component picked for each constraint must take the maximal value. Thus the sampling depends on the order in which initial concentrations are selected. To avoid introducing a bias related to this order, a random ordering of components is selected for each random initial state. Furthermore, if a component is picked as last in a constraint, its value is enforced (it must take all what remains). Conflicts may arise if this component is part of other constraints as well. To avoid this, we ensure that the last assigned item is specific to the constraint (not involved in any other constraint). This step may create problems with highly interdependent sets of constraints or constraints with less than two specific components. For the models considered here, this sampling works well as constraints have more specific components than overlapping ones.

By this method we generate N symbolic trajectories of length M defin-

ing the vectors of successive states $(s_1^j, s_2^j, \ldots, s_M^j)$ and residence times $(\tau_1^j, \tau_2^j, \ldots, \tau_M^j)$, where $j \in [1, N]$.

The stochastic automaton is learned as a homogenous, finite states, continuous time Markov process, defined by the lifetime (mean sojourn time) of each state T_i, $1 \leq i \leq b$ and by the transition probabilities $p_{i,j}$ from a state i to another state j. We use the following estimators for the lifetimes and for the transition probabilities:

$$T_i = (\sum_{n=1}^{N}\sum_{m=1}^{M} \tau_n \mathbb{1}_{s_m^n=i})/(\sum_{n=1}^{N}\sum_{m=1}^{M} \mathbb{1}_{s_m^n=i}) \quad (16)$$

$$p_{i,j} = (\sum_{n=1}^{N}\sum_{m=1}^{M} \mathbb{1}_{s_m^n=i, s_{m+1}^n=j})/(\sum_{n=1}^{N}\sum_{m=1}^{M} \mathbb{1}_{s_m^n=i}),\ i \neq j, \quad (17)$$

where $\mathbb{1}_C$ stands for the indicator function equal to one if condition C is fulfilled or else equal to nought.

3. Results and discussion

3.1. Benchmarking metastable branch computation on Biomodels database

Data source. For benchmarking, we selected 36 models from the r25 version of Biomodels database (Le Novere et al., 2006). We have considered all the models that have polynomial vector field (mass action kinetics) and satisfy some technical constraints imposed by our Systems Biology Markup Language (SBML) parser (no function definitions, for instance). Models with zero valued parameters or which did not have at least one positive and one negative monomial per ODE were also filtered out (there were two of those).

Computation of minimal branches. The model SBML files are parsed and the polynomial vector fields are extracted. Thereafter, the conservation laws (that are the sum of the variables whose total concentration is invariant) are computed. The vector field along with the conservation laws are the input to an unpublished version of the tropical geometry based algorithm by Samal et al. (2015a) to compute the minimal branches. It should be noted here that due to the conservation laws the number of equations may exceed the number of chemical species.

According to the Eq.(8) and to the geometric interpretation of tropical equilibrations from Sect. 2.1 the tropical solutions are either isolated points or bounded or unbounded polyhedra. Changing the parameter ε is just a way to

approximate the position of these points and polyhedra by lattices or in other words by integer coefficients vectors. Finding the value of ε that provides the best approximation is a complicated problem in Diophantine approximation. For that reason, we preferred an experimental approach consisting in choosing several values of ε and checking the robustness of the results.

A summary of the analysis is presented in Table 1 with seven different choices for ε values (we wanted to have orders of magnitude close to decimal ones and to avoid commensurability between different values of ε; the choice 1/5, 1/7, 1/9, 1/11, 1/17, 1/19, 1/23 seemed good enough for this purpose). In addition, in Fig. 2 we plot the number of minimal branches versus the number of equations in the model. As can be noticed, this number is much lower than the number of states of a Boolean network with the same number of variables, which illustrates the advantage of our coarse graining with respect to other methods that discretize the values of the variables in order to obtain Boolean or multi-value networks (Davidich and Bornholdt, 2008).

The complete list of models and the corresponding statistics is provided as Supplementary File 1.

There are a few theoretical bounds for the number of tropical branches and metastable regimes. For monomolecular networks (linear differential equations) there can be at most one branch of full tropical equilibrations. However, if we allow that slow variables are not equilibrated in metastable regimes (cf. the notion of partial equilibration in Section 2.1) we showed elsewhere (Radulescu et al., 2015a) that there can be more than one metastable regime, but less than n, where n is the number of species. For nonlinear networks, one upper bound can be obtained by the tropical Bézout theorem (Richter-Gebert et al., 2005), which says that the number of tropical equilibrations is exponential in n. However, our numerical experiments show that this bound is largely overestimated.

Table. 1 here

Fig. 2 here

3.2. Biological significance of metastable branches for TGF-β signalling

As a case study we consider a nonlinear model of dynamic regulation of Transforming Growth Factor beta TGF-β signalling pathway that we have recently described by Andrieux et al. (2012). TGF-β signalling occurs through association of the ligand with TGF-beta type I (TGFBR1) and type II (TGFBR2) serine/threonine kinase receptors. TGF-β binding to TGFBR2 induces recruitment and phosphorylation of TGFBR1, which in

turn transmits the signal through phosphorylation of SMAD2 transcription factor. Once phosphorylated, the SMAD2 hetero-dimerizes with SMAD4 and the complexes then migrate to the nucleus, where they regulate the transcription of TGF-β-target genes. In that context, the Transcriptional Intermediary Factor 1, TIF1-γ have been shown to function either as a transcriptional repressor or as an alternative transcription factor that promote TGF-β signalling. The apparent controversial effect of TIF1-γ on regulation of the SMAD-dependent TGF-β signalling was solved by a model integrating a ternary complex associating TIF1-γ with SMAD2 and SMAD4 complexes. This model has a dynamics defined by $n = 18$ polynomial differential equations and 25 biochemical reactions (the full set of ordinary differential equations can be found in the Appendix 1, the reaction scheme can be found in by (Andrieux et al., 2012)) or in the SBML model that we provide as Supplementary File 2. The computation of the tropical equilibrations for this model shows that there are 9 minimal branches of full equilibrations (in these tropical solutions all variables are equilibrated). The connectivity graph of these branches (defined in Section 2.1) and the learned finite-state automaton (obtained with the method of Section 2.2) are shown in Fig. 3. The Table2 illustrates the convergence of the estimated transition probabilities when N, the number of Monte-Carlo samples, is increased. In order to illustrate the robustness of the classification of the states in the trajectories for each branch B_i, $i \in [1, 9]$ we computed the distribution of probability of Euclidian distances between randomly chosen states of the model, compatible with the conservation laws, and the branches. In Fig.4 these distributions are compared with the distribution of minimal distances used to classify states on the model's trajectories. The latter distances are smaller and clearly separated from the random states distances.

Fig. 3 here
Table 2 here
Fig. 4 here

The transition probabilities of the automaton are coarse grained properties of the statistical ensemble of trajectories for different initial conditions (cf. Section 2.1). Given a state and a minimal branch close to it, it will depend on the actual trajectory to which other branch the system will be close to next. However, when initial data and the full trajectory are not known, the automaton will provide estimates of where we go next and with which probability. For the example studied and for nominal parameter values, the branch B1 is a globally attractive sink: starting from anywhere, the automa-

ton will reach B1 with probability one. This branch contains the unique stable steady state of the initial model. This calculation illustrates the basic properties of minimal branches of equilibrations. Trajectories of the dynamical system can be decomposed into segments that remain close to minimal branches. Furthermore, all the observed transitions between branches are contained in the connectivity graph resulting from the polyhedral complex of the tropical equilibration branches. This result proves the solidity of our tropical approach, because the geometric connectivity was not enforced to constrain the possible transitions; the fact that it is really the case emerges from the analysis of the trajectories. A change of parameter values can have several consequences: change the connectivity graph, change of the probabilities of transitions and change of the attractor position.

In order to understand the significance of the minimal branches and their relation with dynamic and physiologic properties of the network we have performed an analytic study of the tropical equilibration solutions. We show in the following (see also the Appendix 2) that the most important cause of the multiplicity of branches is the dynamics of the TGFBR1 and TGFBR2 receptors whose internalization and trafficking regulates TGF-β signalling (Le Roy and Wrana, 2005). These two receptors belong to a ligand-receptor module of 6 variables and 12 reactions that is decoupled from the rest of the network. More precisely, the ligand-receptor module activates the SMAD transcription factors but receives no feed-back (see Fig. 5) and can be studied independently from the rest of the variables. This module has been used with little variation in many models of TGF-β signalling (Vilar et al., 2006; Zi and Klipp, 2007; Chung et al., 2009).

Fig. 5 here

We show in the Appendix 2 that the tropical equilibration of the ligand-receptor module form a two dimensional polyhedron conveniently parametrized by the concentration orders a_{12} and a_{13} of the receptors TGFBR1 and TGFBR2 respectively. The branches can be calculated analytically (Eqs.(25) and (29) in Appendix 2). For the nominal values of the model parameters one of these branches is empty and the three remaining branches correspond to B_1, B_2 and B_3. The two other triplets of branches (B_4, B_5, B_6) and (B_7, B_8, B_9) correspond to the same mutual relations of variables in the ligand-receptor module. They are distinguished by the values of the remaining variables (the transcription factors module). Our computation of the automaton showed that the branches B_i, $i \in [4, 9]$ are practically inaccessible from states in branches B_i, $i \in [1, 3]$, therefore we will not discuss

them here.

We have used symbolic computation to determine the steady states of the ligand-receptor module. This module has an unique steady state corresponding to concentrations orders that can be placed inside the polyhedron of tropical solutions using the Eq.(4). The minimal branch containing the steady state is a sink of the coarse grained dynamics. The polyhedron of tropical solutions, its decomposition into minimal branches, and the position of the steady states inside it, depend on model parameters. Among model parameters two are important: k_{18} and k_{19} representing the production rate of the protein receptors TGFBR1 and TGFBR2, respectively. Consequently, these two parameters are correlated to gene expression and account for possible variability in mRNA levels of the two types of receptors. Fig. 6 shows the tropical equilibration branches of the ligand-receptor modules for various parameters k_{19} corresponding to various TGFBR2 expression levels. For the nominal parameters used in the model, the branch $B1$ is a sink i.e. an attractor (the coarse-grained dynamics shows that the probability to leave this state is negligible), and the branches $B2$ and $B3$ are transient i.e. metastable. This means that starting in the branch $B2$ or $B3$ the receptor module will reach the branch $B1$ after a certain time and will remain there. However, over-expression of TGFBR2 modelled by changing the parameter k_{19} and illustrated in the Fig. 6a-c can tilt the balance in favor of large concentrations of receptor of type 2 corresponding to branches $B2$, $B3$ of tropical solutions in the model. Interestingly, this change occurs by a displacement of the position of the steady state from $B1$ to $B2$ and $B3$ and not by a change of the concentration values allowed for these branches.

Fig. 6 here

While Vilar et al. (2006) have speculated that the ligand-receptor module is responsible for the versatility of the response of the TGF-β pathway, no experimental evidence support this hypothesis. Here, we now demonstrate that there are correlations between dynamical specificity characterized by membership to a particular branch of equilibration and cell phenotype. We illustrate such a comparison for the NCI-60 panel of cancer cell lines, a well established tool for tumor comparison and drug screening provided by the National Cancer Institute. Based on microarray analysis, these cell lines were found to cluster into two classes: epithelium-like (non-aggressive) and mesenchymal-like (aggressive) cell lines (Ross et al., 2000).

Using the global proteome analysis of this NCI-60 panel (Gholami et al., 2013) we extracted the protein expression levels of TGFBR1 and TGFBR2

and showed that mesenchymal-like (aggressive) cell lines can be distinguished from epithelial-like (non-aggressive) cell lines by the increased level of TGFBR2 (Fig 6d.

The proteome data was compared to membership to tropical branches. According to Eq.(4) there is a linear relation between opposite concentration orders $-a_i$ and logarithms of concentrations, $-a_i \approx b\log(x_i)$, $i = 1, \ldots, n$ ($b = -1/\log(\varepsilon) > 0$). We used opposite concentration orders $-a_i$ instead of a_i because they change in the same direction as the concentrations (small opposite orders mean small concentrations and large opposite orders mean large concentrations). Therefore, in Fig 6a-c the relation $TGFBR1 = TGFBR2$ is verified on the bissector of the first quadrant, whereas $TGFBR2 > TGFBR1$ and $TGFBR2 < TGFBR1$ are valid above and below the bissector, respectively. When compared the proteome results with the membership to a particular branch of equilibration, we found that the distribution of concentration orders in branches place non-aggressive cancer cell lines in a range covered by branch 1, whereas the aggressive cancer cell lines are placed in a range covered by branches 2 and 3 (Fig 6d). Indeed, the ratio $TGFBR2/TGFBR1$ is small for the branch B_1 and in non-aggressive cancer cell lines, and is much larger for B_2, B_3 and in aggressive cell lines.

Furthermore, we validated the association of up-regulation of TGFBR2 with mesenchymal-like appearance in an independent dataset of 51 breast cancer cell lines (Neve et al., 2006). As we have recently described (Ruff et al., 2015), comparative analyses between Basal B cell lines with mesenchymal-like phenotype and Basal A and Luminal cell lines with epithelial morphology permitted to identify more than 600 differentially expressed genes that include TGFBR2. Gene expression data were now extracted for TGFBR1 and TGFBR2 and we showed that TGFBR2 gene expression is significantly induced in mesenchymal-like cell lines while TGFBR1 did not vary (Supplementary Fig.1). In accordance with our observation, Parker et al. (Parker et al., 2007) have previously reported the association of low TGFBR2 expression with a lower aggressive tumor phenotype.

In summary, the tropical geometry analysis of the TGFβ signalling model first shows that the multiplicity of branches is due to the dynamics of TGFβ receptors. The more important parameters in this ligand-receptor module are the concentration of TGFBR1 and TGFBR2 and three main tropical branches are distinguished by the value of TGFBR2/TGFBR1 ratio (small in B1, intermediate in B2, large in B3). Importantly, we showed that the TGFBR2/TGFBR1 ratio is associated with tumor cell lines phenotype (high

and low TGFBR2/TGFBR1 ratio for aggressive and non aggressive cell lines, respectively). Together these data demonstrated that tropical geometry analysis permits to discriminate between cellular states based on the evaluation of TGFβ receptors concentration. The importance of such up-regulation of TGFBR2 in aggressive cancer cell lines might be related to its implication in SMAD-independent signalling that includes PI3K-Akt, JNK, p38MAPK and Rho-like GTPases and which highly contribute to epithelial-mesenchymal transition (Zhang, 2009; Moustakas and Heldin, 2012).

Together these observations suggest that metastable regimes defined by branches of minimal tropical equilibrations are associated with cell phenotypes. The idea of associating tropical minimal branches with clinical phenotype is similar to the idea of cancer attractors (Huang et al., 2009) where the idea is that cancer cells are trapped in some abnormal attractors.

4. Conclusion

We have presented a method to coarse grain the dynamics of a smooth biochemical reaction network to a discrete symbolic dynamics of a finite state automaton. The coarse graining was obtained using a tropical geometry approach to compute the states. These states correspond to metastable dynamic regimes and to relatively slow segments of the system trajectories. The coarse grained model can be used for studying statistical properties of biochemical networks such as occurrence and stability of temporal patterns, recurrence, periodicity and attainability problems.

Further improvement and evolution is possible for this approach. First, the coarse graining can be performed in a hierarchical way. For the nonlinear example studied in the paper we computed only the full tropical equilibrations that stand for the lowest order in the hierarchy (coarsest model). As discussed in Section 2.1 we can also consider partial equilibrations when slow variables are not equilibrated and thus refine the automaton. Generally, there are more partial equilibrations than total equilibrations and learning an automaton on the augmented state set will produce refinements. Second, and most importantly, the dynamics within a branch could be also described. As shown elsewhere, reductions of the systems of ordinary differential equations are valid locally close to tropical equilibrations (Noel et al., 2012, 2014; Radulescu et al., 2015b; Samal et al., 2015b). Furthermore, the same reduction is valid for all the equilibrations in a branch. This suggests that a hybrid approach, combining reduced ODE dynamics within branch

with discrete transitions between branches is feasible. The transitions can be autonomously and deterministically commanded by crossing the boundaries between branches that are perfectly determined by our approach.

The most important result of this paper is the extension of the notion of attractor to metastable regimes of chemical reaction networks and the proposition of a practical recipe to compute metastability. Metastable regimes correspond to low-dimensional hypersurfaces of the phase space, along which the dynamics is relatively slower. Most likely, metastable regimes have biological importance because the network spends most of its time in these states. The itinerancy of the network, described as the possibility of transitions from one metastable regime to another is paramount to the way neural networks compute, retrieve and use information (Tsuda, 1991) and can have similar role in biochemical networks. Our approach based on tropical geometry provides an algorithmic method to test these ideas further. The extension of this approach i.e. making use of statistical methods to compute the association of the tropical minimal branches with clinical phenotypes based on "-omics" data remains a topic of future research.

Acknowledgements. O.R and A.N are supported by INCa/Plan Cancer grant N°ASC14021FSA. D.G. is grateful to the grant RSF 16-11-10075 and to MCCME for wonderful working conditions and inspiring atmosphere.

References

Andrieux, G., Fattet, L., Le Borgne, M., Rimokh, R., Théret, N., 2012. Dynamic regulation of Tgf-B signaling by Tif1γ: a computational approach. PloS One 7 (3), e33761.

Bowman, G. R., Beauchamp, K. A., Boxer, G., Pande, V. S., 2009. Progress and challenges in the automated construction of markov state models for full protein systems. The Journal of Chemical Physics 131 (12), 124101.

Chang, C.-S., 1998. On deterministic traffic regulation and service guarantees: a systematic approach by filtering. IEEE Transactions on Information Theory 44 (3), 1097–1110.

Chiavazzo, E., Karlin, I., 2011. Adaptive simplification of complex multiscale systems. Physical Review E 83 (3), 036706.

Chung, S.-W., Miles, F. L., Sikes, R. A., Cooper, C. R., Farach-Carson, M. C., Ogunnaike, B. A., 2009. Quantitative modeling and analysis of the transforming growth factor β signaling pathway. Biophysical Journal 96 (5), 1733–1750.

Davidich, M., Bornholdt, S., 2008. The transition from differential equations to boolean networks: a case study in simplifying a regulatory network model. Journal of Theoretical Biology 255 (3), 269–277.

Dubrova, E., Teslenko, M., 2011. A sat-based algorithm for finding attractors in synchronous boolean networks. IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB) 8 (5), 1393–1399.

Gholami, A. M., Hahne, H., Wu, Z., Auer, F. J., Meng, C., Wilhelm, M., Kuster, B., 2013. Global proteome analysis of the nci-60 cell line panel. Cell Reports 4 (3), 609–620.

Gorban, A., Karlin, I., 2005. Invariant manifolds for physical and chemical kinetics. Vol. 660 of Lect. Notes Phys. Springer.

Gorban, A. N., Radulescu, O., 2008. Dynamic and static limitation in reaction networks, revisited. In: Guy B. Marin, D. W., Yablonsky, G. S. (Eds.), Advances in Chemical Engineering – Mathematics in Chemical Kinetics and Engineering. Vol. 34 of Advances in Chemical Engineering. Elsevier, pp. 103–173.

Gorban, A. N., Radulescu, O., Zinovyev, A. Y., 2010. Asymptotology of chemical reaction networks. Chemical Engineering Science 65 (7), 2310–2324.

Gorban, A. N., Shahzad, M., 2011. The Michaelis-Menten-Stueckelberg theorem. Entropy 13 (5), 966–1019.

Haller, G., Sapsis, T., 2010. Localized instability and attraction along invariant manifolds. SIAM Journal on Applied Dynamical Systems 9 (2), 611–633.

Huang, S., Ernberg, I., Kauffman, S., 2009. Cancer attractors: A systems view of tumors from a gene network dynamics and developmental perspective. Seminars in Cell & Developmental Biology 20 (7), 869 – 876.

Kauffman, S. A., 1969. Metabolic stability and epigenesis in randomly constructed genetic nets. Journal of Theoretical Biology 22 (3), 437–467.

Korem, Y., Szekely, P., Hart, Y., Sheftel, H., Hausser, J., Mayo, A., Rothenberg, M. E., Kalisky, T., Alon, U., 2015. Geometry of the gene expression space of individual cells. PLoS Comput Biol 11 (7), e1004224.

Le Novere, N., Bornstein, B., Broicher, A., Courtot, M., Donizelli, M., Dharuri, H., Li, L., Sauro, H., Schilstra, M., Shapiro, B., Snoep, J. L., Hucka, M., 2006. BioModels database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems. Nucleic Acids Research 34 (suppl 1), D689–D691.

Le Roy, C., Wrana, J. L., 2005. Clathrin-and non-clathrin-mediated endocytic regulation of cell signalling. Nature Reviews Molecular Cell Biology 6 (2), 112–126.

Litvinov, G., 2007. Maslov dequantization, idempotent and tropical mathematics: a brief introduction. Journal of Mathematical Sciences 140 (3), 426–444.

Maas, U., Pope, S. B., 1992. Simplifying chemical kinetics: intrinsic low-dimensional manifolds in composition space. Combustion and Flame 88 (3), 239–264.

Maclagan, D., Sturmfels, B., 2015. Introduction to tropical geometry. Vol. 161 of Graduate Studies in Mathematics. American Mathematical Society, RI.

Meiske, W., 1978. An approximate solution of the michaelis-menten mechanism for quasi-steady and state quasi-equilibrium. Mathematical Biosciences 42 (1), 63–71.

Moustakas, A., Heldin, C.-H., 2012. Induction of epithelial–mesenchymal transition by transforming growth factor β. In: Seminars in Cancer Biology. Vol. 22. Elsevier, pp. 446–454.

Naldi, A., Thieffry, D., Chaouiya, C., 2007. Decision diagrams for the representation and analysis of logical models of genetic networks. In: Computational Methods in Systems Biology. Springer, pp. 233–247.

Neve, R. M., Chin, K., Fridlyand, J., Yeh, J., Baehner, F. L., Fevr, T., Clark, L., Bayani, N., Coppe, J.-P., Tong, F., et al., 2006. A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. Cancer Cell 10 (6), 515–527.

Noel, V., Grigoriev, D., Vakulenko, S., Radulescu, O., 2012. Tropical geometries and dynamics of biochemical networks application to hybrid cell cycle models. In: Feret, J., Levchenko, A. (Eds.), Proceedings of the 2nd International Workshop on Static Analysis and Systems Biology (SASB 2011). Vol. 284 of Electronic Notes in Theoretical Computer Science. Elsevier, pp. 75–91.

Noel, V., Grigoriev, D., Vakulenko, S., Radulescu, O., 2014. Tropicalization and tropical equilibration of chemical reactions. In: Litvinov, G., Sergeev, S. (Eds.), Tropical and Idempotent Mathematics and Applications. Vol. 616 of Contemporary Mathematics. American Mathematical Society, pp. 261–277.

Pachter, L., Sturmfels, B., 2004. Tropical geometry of statistical models. Proceedings of the National Academy of Sciences of the United States of America 101 (46), 16132.

Parker, A. S., Lohse, C. M., Wu, K., Kreinest, P., Copland, J. A., Hilton, T., Wehle, M., Cheville, J. C., Blute, M., 2007. Lower expression levels of the transforming growth factor beta receptor type ii protein are associated with a less aggressive tumor phenotype and improved survival among patients with clear cell renal cell carcinoma. Human Pathology 38 (3), 453–461.

Rabinovich, M. I., Afraimovich, V. S., Bick, C., Varona, P., 2012. Information flow dynamics in the brain. Physics of Life Reviews 9 (1), 51–73.

Rabinovich, M. I., Varona, P., Selverston, A. I., Abarbanel, H. D., 2006. Dynamical principles in neuroscience. Reviews of Modern Physics 78 (4), 1213.

Radulescu, O., Samal, S. S., Naldi, A., Grigoriev, D., Weber, A., 2015a. Symbolic dynamics of biochemical pathways as finite states machines. In: Roux, O., Bourdon, J. (Eds.), Computational Methods in Systems Biology: 13th International Conference, CMSB 2015, Nantes, France, September 16-18, 2015, Proceedings. Vol. 9308 of Lecture Notes in Computer Science. Springer, pp. 104–120.

Radulescu, O., Vakulenko, S., Grigoriev, D., 2015b. Model reduction of biochemical reactions networks by tropical analysis methods. Mathematical Model of Natural Phenomena 10 (3), 124–138.

Richter-Gebert, J., Sturmfels, B., Theobald, T., 2005. First steps in tropical geometry. Contemporary Mathematics 377, 289–318.

Ross, D. T., Scherf, U., Eisen, M. B., Perou, C. M., Rees, C., Spellman, P., Iyer, V., Jeffrey, S. S., Van de Rijn, M., Waltham, M., et al., 2000. Systematic variation in gene expression patterns in human cancer cell lines. Nature Genetics 24 (3), 227–235.

Ruff, M., Leyme, A., Le Cann, F., Bonnier, D., Le Seyec, J., Chesnel, F., Fattet, L., Rimokh, R., Baffet, G., Théret, N., 2015. The disintegrin and metalloprotease adam12 is associated with tgf-β-induced epithelial to mesenchymal transition. PloS One 10 (9), e0139179.

Rustici, G., Mata, J., Kivinen, K., Lió, P., Penkett, C. J., Burns, G., Hayles, J., Brazma, A., Nurse, P., Bähler, J., 2004. Periodic gene expression program of the fission yeast cell cycle. Nature Genetics 36 (8), 809–817.

Samal, S. S., Grigoriev, D., Fröhlich, H., Radulescu, O., Sep. 2015a. Analysis of reaction network systems using tropical geometry. In: Gerdt, V. P., Koepf, W., Seiler, W. M., Vorozhtsov, E. V. (Eds.), Computer Algebra in Scientific Computing – 17th International Workshop (CASC 2015). Vol. 9301 of Lecture Notes in Computer Science. Springer, Aachen, Germany, pp. 422–437.

Samal, S. S., Grigoriev, D., Fröhlich, H., Weber, A., Radulescu, O., 2015b. A geometric method for model reduction of biochemical networks with polynomial rate functions. Bulletin of Mathematical Biology 77 (12), 2180–2211.

Samal, S. S., Radulescu, O., Grigoriev, D., Fröhlich, H., Weber, A., 2014. A tropical method based on newton polygon approach for algebraic analysis of biochemical reaction networks. In: 9th European Conference on Mathematical and Theoretical Biology.

Segel, L. A., 1988. On the validity of the steady state assumption of enzyme kinetics. Bulletin of Mathematical Biology 50 (6), 579–593.

Segel, L. A., Slemrod, M., 1989. The quasi-steady-state assumption: a case study in perturbation. SIAM Review 31 (3), 446–477.

Simon, I., 1988. Recognizable sets with multiplicities in the tropical semiring. In: Mathematical Foundations of Computer Science 1988. Springer, pp. 107–120.

Soliman, S., Fages, F., Radulescu, O., 2014. A constraint solving approach to model reduction by tropical equilibration. Algorithms for Molecular Biology 9 (1), 24.

Theobald, T., 2006. On the frontiers of polynomial computations in tropical geometry. Journal of Symbolic Computation 41 (12), 1360–1375.

Thomas, R., 1973. Boolean formalization of genetic control circuits. Journal of Theoretical Biology 42 (3), 563–585.

Thomas, R., 1991. Regulatory networks seen as asynchronous automata: a logical description. Journal of Theoretical Biology 153 (1), 1–23.

Tsuda, I., 1991. Chaotic itinerancy as a dynamical basis of hermeneutics in brain and mind. World Futures: Journal of General Evolution 32 (2-3), 167–184.

Tyson, J. J., Csikasz-Nagy, A., Novak, B., 2002. The dynamics of cell cycle regulation. Bioessays 24 (12), 1095–1109.

Vilar, J., Jansen, R., Sander, C., 2006. Signal processing in the TGF-beta superfamily ligand-receptor network. PLoS Comput Biol 2 (1), e3.

Weis, M. C., Avva, J., Jacobberger, J. W., Sreenath, S. N., 2014. A data-driven, mathematical model of mammalian cell cycle regulation. PloS One 9 (5), e97130.

Zhang, Y. E., 2009. Non-Smad pathways in TGF-β signaling. Cell Research 19 (1), 128–139.

Zi, Z., Klipp, E., 2007. Constraint-based modeling and kinetic analysis of the Smad dependent TGF-beta signaling pathway. PLoS One 2 (9), e936–e936.

Appendix 1: Description of the TGFb model used in this paper.

The model is described by the following system of differential equations

$$
\begin{aligned}
\frac{\mathrm{d}x_1}{\mathrm{d}t} &= k_2x_2 - k_1x_1 - k_{16}x_1x_{11} \\
\frac{\mathrm{d}x_2}{\mathrm{d}t} &= k_1x_1 - k_2x_2 + k_{17}k_{34}x_6 \\
\frac{\mathrm{d}x_3}{\mathrm{d}t} &= k_3x_4 - k_3x_3 + k_7x_7 + k_{33}k_{37}x_{18} - k_6x_3x_5 \\
\frac{\mathrm{d}x_4}{\mathrm{d}t} &= k_3x_3 - k_3x_4 + k_9x_8 - k_8x_4x_6 \\
\frac{\mathrm{d}x_5}{\mathrm{d}t} &= k_5x_6 - k_4x_5 + k_7x_7 + 2k_{11}x_9 - 2k_{10}x_5^2 - k_6x_3x_5 + k_{16}x_1x_{11} \\
\frac{\mathrm{d}x_6}{\mathrm{d}t} &= k_4x_5 - k_5x_6 + k_9x_8 + 2k_{13}x_{10} - 2k_{12}x_6^2 - k_{17}k_{34}x_6 + k_{31}k_{36}x_8 - k_8x_4x_6 \\
\frac{\mathrm{d}x_7}{\mathrm{d}t} &= k_6x_3x_5 - x_7(k_7 + k_{14}) \\
\frac{\mathrm{d}x_8}{\mathrm{d}t} &= k_{14}x_7 - k_9x_8 - k_{31}k_{36}x_8 + k_8x_4x_6 \\
\frac{\mathrm{d}x_9}{\mathrm{d}t} &= k_{10}x_5^2 - x_9(k_{11} + k_{15}) \\
\frac{\mathrm{d}x_{10}}{\mathrm{d}t} &= k_{15}x_9 - k_{13}x_{10} + k_{12}x_6^2 \\
\frac{\mathrm{d}x_{11}}{\mathrm{d}t} &= k_{23}x_{14} - k_{30}x_{11} \\
\frac{\mathrm{d}x_{12}}{\mathrm{d}t} &= k_{18} - x_{12}(k_{20} + k_{26}) + k_{30}x_{11} + k_{27}x_{15} - k_{22}k_{35}x_{12}x_{13} \\
\frac{\mathrm{d}x_{13}}{\mathrm{d}t} &= k_{19} - x_{13}(k_{21} + k_{28}) + k_{30}x_{11} + k_{29}x_{16} - k_{22}k_{35}x_{12}x_{13} \\
\frac{\mathrm{d}x_{14}}{\mathrm{d}t} &= k_{22}k_{35}x_{12}x_{13} - x_{14}(k_{23} + k_{24} + k_{25}) \\
\frac{\mathrm{d}x_{15}}{\mathrm{d}t} &= k_{26}x_{12} - k_{27}x_{15} \\
\frac{\mathrm{d}x_{16}}{\mathrm{d}t} &= k_{28}x_{13} - k_{29}x_{16} \\
\frac{\mathrm{d}x_{17}}{\mathrm{d}t} &= k_{31}k_{36}x_8 - k_{32}x_{17} \\
\frac{\mathrm{d}x_{18}}{\mathrm{d}t} &= k_{32}x_{17} - k_{33}k_{37}x_{18} \qquad (18)
\end{aligned}
$$

These variables are as follows:

- Receptors on plasma membrane: x_{12} = RI (receptor 1), x_{13} = RII (receptor 2), x_{14} = LR (ligand-receptor complex).
- Receptors in the endosome: x_{11} = LRe, x_{15} = RIe, x_{16} = RIIe.
- Transcription factors and complexes in cytosol: x_1 = S2c, x_3 = S4c, x_5 = pS2c, x_7 = pS24c, x_9 = pS22c, x_{18} = S4ubc.
- Transcription factors and complexes in the nucleus: x_2 = S2n, x_4 = S4n, x_6 = pS2n, x_8 = pS24n, x_{10} = pS22n, x_{17} = S4ubn.

Appendix 2: Calculation of tropical equilibration branches for the ligand-receptor module.

Tropical equilibration solutions for the variables x_{11},x_{12}, x_{13}, x_{14}, x_{15}, x_{16} (the submodel in Fig. 5) can be computed independently from the rest of the variables of the TGFβ model. The ordinary differential equations for these variables form a subsystem that is decoupled (receives no feed-back) from the rest of the equations.

We can reduce the system of 6 tropical equations to a simplified system of 3 tropical equations using the following two general properties.

Property 4.1 (binomial species). *Y is a binomial species if the ordinary differential equation defining its rate of variation contains only one positive monomial term and only one negative monomial term*

$$\frac{\mathrm{d}Y}{\mathrm{d}t} = M_1(X)Y^{n_1} - M_2(X)Y^{n_2},$$

where X denotes the other variables. We further assume that $n_1 < n_2$. Then, the species Y can be eliminated and the resulting simplified tropical system has the same tropical equilibration solutions as the full system. The simplification is performed by eliminating the equation for Y and replacing everywhere Y by $(M_1/M_2)^{1/(n_2-n_1)}$.

Proof of Property 4.1. The proof follows from the fact that the tropical equation for the order a of Y has the unique solution $a = \frac{1}{(n_2-n_1)}(\mu_1 - \mu_2)$.

Property 4.2 (dominated first order reactions). *If a species Y is consumed by several first order reactions of kinetic constants $k_1, k_2, \ldots, k_r$ and if $\gamma_1 \leq \gamma_2 \leq \ldots \leq \gamma_p < \gamma_{p+1} \leq \gamma_{p+2} \leq \ldots \leq \gamma_r$, then the reactions $k_{p+1}, \ldots, k_r$ can be eliminated and the resulting simplified tropical system has the same tropical equilibration solutions as the full system.*

Proof of Property 4.2. The proof follows from the following obvious property of the min operation $\min(\gamma_1, \ldots, \gamma_p, \ldots, \gamma_r) = \min(\gamma_1, \ldots, \gamma_p)$.

Using $\gamma_{26} < \gamma_{20}$, $\gamma_{28} < \gamma_{21}$ (a condition satisfied by the nominal model parameters and meaning that internalization is more rapid than degradation for both receptors 1 and 2) and the Properties 4.1,4.2 we can justify the reduction illustrated in Fig. 7. Because the reduced model has the same tropical solutions as the full, larger model, it is enough to solve the tropical equilibration problem for the reduced model. This reads

$$\min(\gamma_{18}, a_{14} + \gamma_{23}, a_{12} + \gamma_{26}) = \min(a_{12} + a_{13} + \gamma_{22} + \gamma_{35}, a_{12} + \gamma_{26}) \quad (19)$$

$$\min(\gamma_{19}, a_{14} + \gamma_{23}, a_{13} + \gamma_{28}) = \min(a_{12} + a_{13} + \gamma_{22} + \gamma_{35}, \gamma_{28} + a_{13}) \quad (20)$$

$$\min(\gamma_{24}, \gamma_{25}, \gamma_{23}) + a_{14} = a_{12} + a_{13} + \gamma_{22} + \gamma_{35} \quad (21)$$

Suppose now that the following condition is true

$$\min(\gamma_{24}, \gamma_{25}, \gamma_{23}) = \gamma_{23}. \quad (22)$$

This condition is satisfied by the nominal parameters and, like the previous condition, means that receptors have relatively large life-times. Then from (21) we got $a_{14} = a_{12} + a_{13} + \gamma_{22} + \gamma_{35} - \gamma_{23}$ and the equations (19),(20) become

$$\min(\gamma_{18}, a_{14} + \gamma_{23}, a_{12} + \gamma_{26}) = \min(a_{14} + \gamma_{23}, a_{12} + \gamma_{26}) \quad (23)$$

$$\min(\gamma_{19}, a_{14} + \gamma_{23}, a_{13} + \gamma_{28}) = \min(a_{14} + \gamma_{23}, a_{13} + \gamma_{28}) \quad (24)$$

The solutions of (23), (24) can be easily found and form the following polyhedron

$$\begin{aligned} (\{a_{12} + a_{13} + \gamma_{22} + \gamma_{35} \leq \gamma_{18}\} \cup \{\gamma_{26} + a_{12} \leq \gamma_{18}\}) \cap \\ (\{a_{12} + a_{13} + \gamma_{22} + \gamma_{35} \leq \gamma_{19}\} \cup \{\gamma_{28} + a_{13} \leq \gamma_{19}\}), \\ a_{14} = a_{12} + a_{13} + \gamma_{22} + \gamma_{35} - \gamma_{23}. \end{aligned} \quad (25)$$

The orders of the remaining variables can be found as indicated in Prop. 4.1:

$$a_{15} = a_{12} + \gamma_{26} - \gamma_{27}, \quad (26)$$

$$a_{16} = a_{13} + \gamma_{28} - \gamma_{29}, \quad (27)$$

$$a_{11} = a_{12} + a_{13} + \gamma_{22} + \gamma_{35} - \gamma_{30}. \quad (28)$$

The polyhedron of tropical solutions defined by Eq.(25) can be partitioned into minimal branches (also polyhedra). This can be done by checking which term is dominant in the ordinary differential equations for the variables x_{12}, x_{13} and x_{14} (see Eqs. (18)). The result is that there are at most four minimal branches defined by one of the conditions

$$\begin{aligned}
&\{a_{12}+\gamma_{26} < a_{12}+a_{13}+\gamma_{22}+\gamma_{35}\} \cap \{a_{13}+\gamma_{28} < a_{12}+a_{13}+\gamma_{22}+\gamma_{35}\} \\
&\{a_{12}+\gamma_{26} < a_{12}+a_{13}+\gamma_{22}+\gamma_{35}\} \cap \{a_{13}+\gamma_{28} > a_{12}+a_{13}+\gamma_{22}+\gamma_{35}\} \\
&\{a_{12}+\gamma_{26} > a_{12}+a_{13}+\gamma_{22}+\gamma_{35}\} \cap \{a_{13}+\gamma_{28} < a_{12}+a_{13}+\gamma_{22}+\gamma_{35}\} \\
&\{a_{12}+\gamma_{26} > a_{12}+a_{13}+\gamma_{22}+\gamma_{35}\} \cap \{a_{13}+\gamma_{28} > a_{12}+a_{13}+\gamma_{22}+\gamma_{35}\} \quad (29)
\end{aligned}$$

Fig. 7 here

Supplementary material captions.

Supplementary File 1. The complete list of models from the Biomodels database with the results of the tropical branches calculations.

Supplementary File 2. SBML version of the TGFβ model used in this study.

Supplementary Fig. 1. Gene expression levels for TGFBR2 and TGFBR1 in breast cancer cell lines from (Neve et al. 2006).

Table 1: Summary of analysis on Biomodels database. The benchmarked models have a number of dimensions (i.e. number of variables along with number of conservation laws) ranging from 2 to 41. Model BIOMD0000000289 has tropical branches at ε value 1/5, 1/7, 1/9, 1/11 but none at 1/17, 1/19, 1/23.

ε value	Total models considered	Models without tropical branches	Models with tropical branches	Average running time (in secs)	Average number of minimal branches	Min number of minimal branches	Max number of minimal branches
1/5	36	0	36	200.76	15.08	1	423
1/7	36	0	36	177.01	14.41	1	406
1/9	36	0	36	195.01	13.02	1	340
1/11	36	0	36	169.59	12.36	1	322
1/17	36	1	35	175.56	11.02	0	287
1/19	36	1	35	187.42	11.08	0	287
1/23	36	1	35	184.80	11.08	0	287

Table 2: Estimation of the transition probabilities between branches of the TGFβ model, for different values of N, the number of Monte-Carlo samples.

N	1→1	1→4	2→1	2→3	2→5	3→2	3→6	4→1	5→1	5→2	5→3
100	1.0000	0.0000	0.9836	0.0164	0.0000	1.0000	0.0000	1.0000	0.0227	0.2727	0.0227
200	1.0000	0.0000	0.9675	0.0325	0.0000	1.0000	0.0000	1.0000	0.0217	0.2391	0.0109
500	1.0000	0.0000	0.9614	0.0386	0.0000	1.0000	0.0000	1.0000	0.0216	0.2381	0.0173
1000	1.0000	0.0000	0.9703	0.0264	0.0033	1.0000	0.0000	1.0000	0.0280	0.2258	0.0086
2000	1.0000	0.0000	0.9585	0.0365	0.0050	1.0000	0.0000	1.0000	0.0246	0.2173	0.0086
3000	1.0000	0.0000	0.9561	0.0389	0.0050	1.0000	0.0000	1.0000	0.0295	0.2342	0.0084
5000	0.9996	0.0004	0.9542	0.0399	0.0059	0.9996	0.0004	1.0000	0.0340	0.2460	0.0093
10000	0.9994	0.0006	0.9556	0.0363	0.0081	0.9973	0.0027	1.0000	0.0344	0.2370	0.0110

N	5→4	5→6	6→3	7→4	8→4	8→5	8→6	8→7	8→9	9→6
100	0.4091	0.2727	1.0000	1.0000	0.1373	0.4902	0.0784	0.0784	0.2157	1.0000
200	0.4130	0.3152	1.0000	1.0000	0.1300	0.5200	0.0600	0.0600	0.2300	1.0000
500	0.4156	0.3074	1.0000	1.0000	0.1098	0.5294	0.0784	0.0980	0.1843	1.0000
1000	0.4280	0.3097	1.0000	1.0000	0.1051	0.5333	0.0949	0.0909	0.1758	1.0000
2000	0.4497	0.2998	1.0000	1.0000	0.1121	0.5354	0.1020	0.0859	0.1646	1.0000
3000	0.4439	0.2840	1.0000	1.0000	0.1122	0.5315	0.0984	0.0951	0.1627	1.0000
5000	0.4227	0.2880	1.0000	1.0000	0.1160	0.5272	0.0969	0.0965	0.1633	1.0000
10000	0.4304	0.2872	1.0000	1.0000	0.1132	0.5344	0.0997	0.0932	0.1595	1.0000

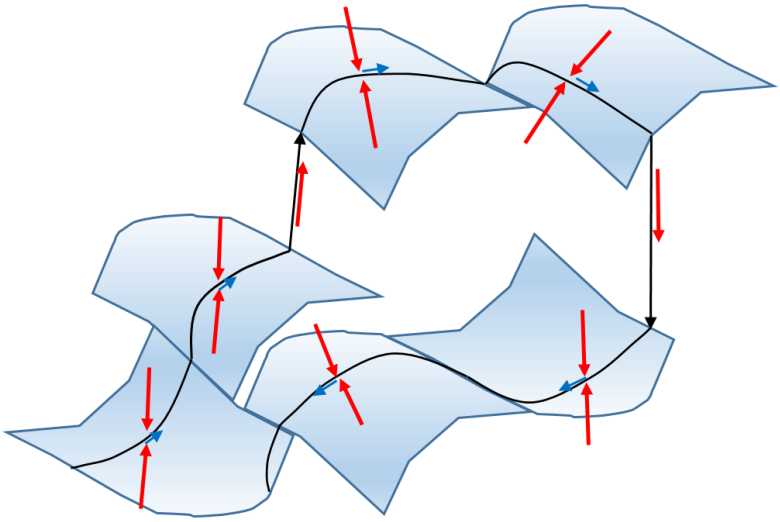

Figure 1: Abstract representation of metastability as itinerant trajectory in a patchy phase space landscape. Dominant vector fields (red arrows) confine the trajectory to low dimensional patches on which act weak uncompensated vector fields (blue arrows). A typical trajectory contains slow segments within patches where dominant vector fields cancel, and transitions between patches in the fast direction of uncancelled dominant vector fields. Continuous (but non smooth) connections are also possible, corresponding to role reversal between dominant and dominated vector fields. The term *crazy-quilt* was coined to describe such a patchy landscape (Gorban and Radulescu, 2008).

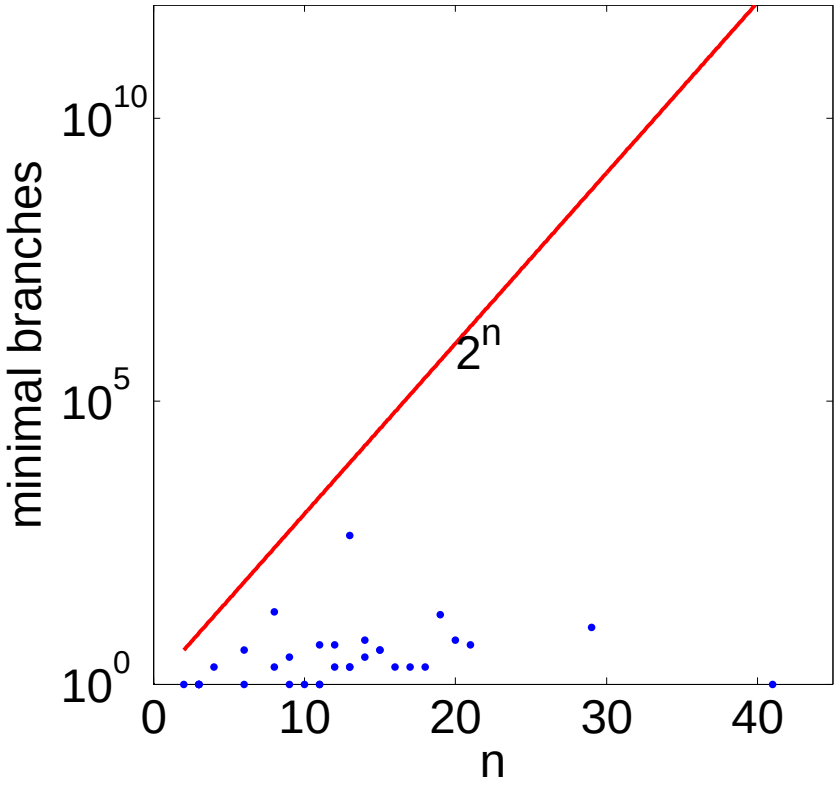

Figure 2: Semi-log plot of the minimal branches versus the number of equations in the models from Biomodels repository for $\varepsilon = 1/5$. Comparison with a binary network number of states 2^n suggests sub-exponential scaling.

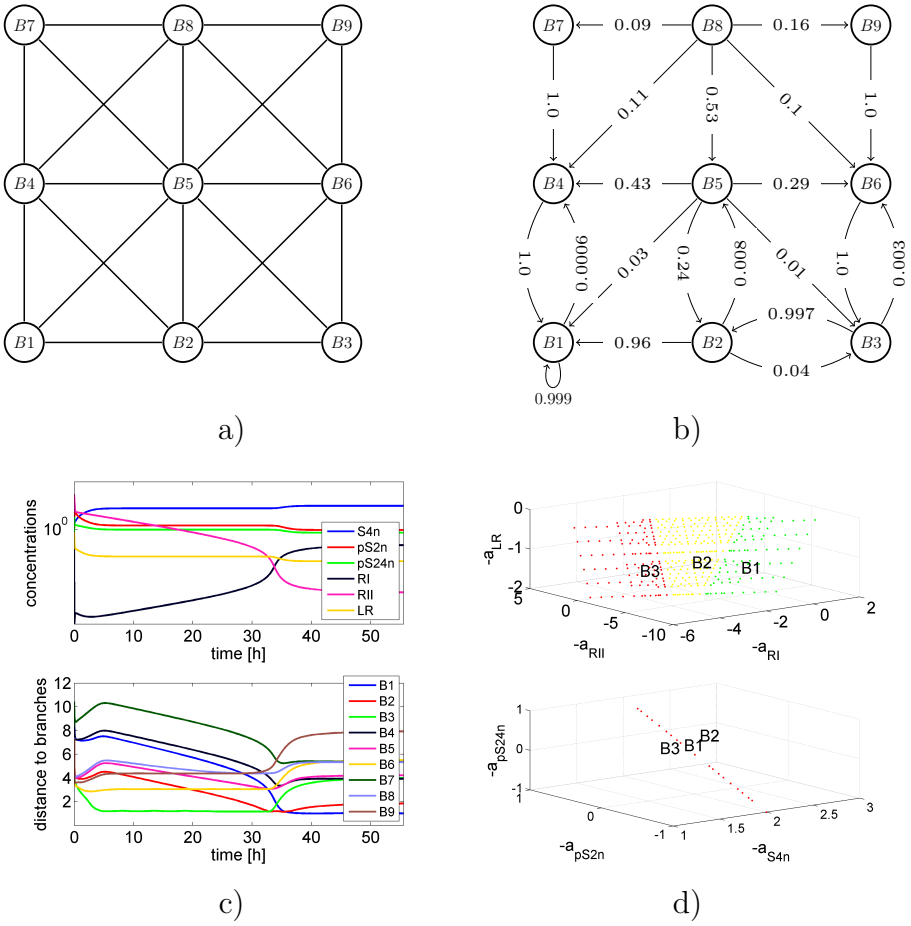

Figure 3: Summary of tropical geometry analysis of the TGFβ model. a) Connectivity graph of tropical minimal branches; b) finite state automaton; c) a single trajectory of the system (starting from initial data chosen randomly close to the branch B_3) is represented by plotting the concentration of different species vs. time (upper sub-figure); the distances to different branches of solutions vs. time (lower sub-figure) shows that the sequence of branches for this trajectory is B_3, B_2, B_1 (all points of the trajectory are close to one of these three branches and significantly more distant to the other branches). d) The different branches of solutions are defined by allowed concentrations of different variables, represented here by orders of magnitudes a_i; the opposite concentration orders $-a_i$ are proportional to the logarithms of concentrations $-a_i \sim \log(x_i)$. All the order calculations were performed using $\varepsilon = 1/11$. The most used branches B_1, B_2, B_3 are shown in projection onto sets of three variables. The variables RI, RII, LR are plasma membrane receptors and ligand-receptor complex (signalling input layers), whereas pS2n, S4n, pS24n are nuclear transcription factors and complexes (effectors). The structure of tropical branches shows that composition of input layers is more flexible (varies on planar domains that are disjoint for different branches) than the concentrations of effectors (vary on linear intervals that overlap for different branches).

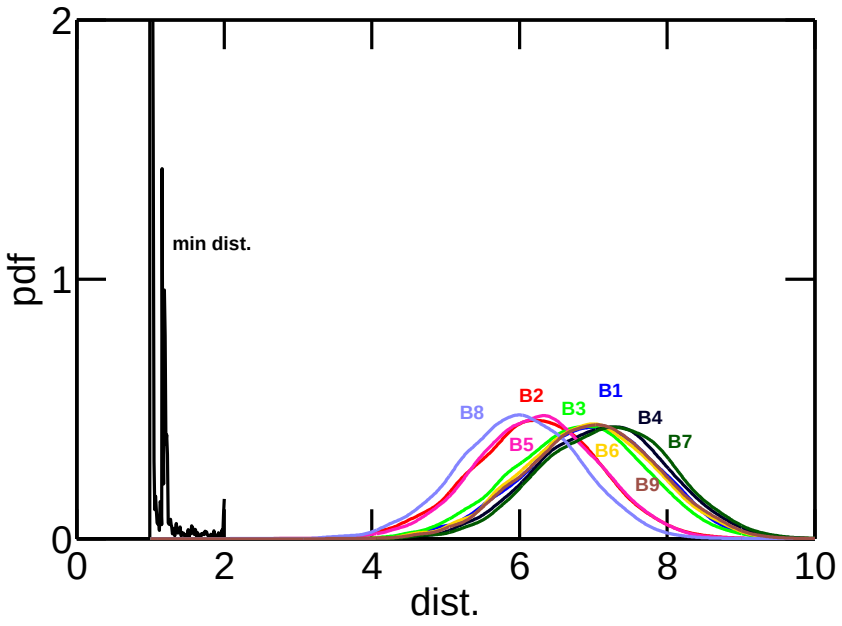

Figure 4: Distribution of distances to tropical branches of the TGFβ model. For each branch B_i, $i \in [1, 9]$ we represent the probability density function of Euclidian distances between randomly chosen states of the model, compatible with the conservation laws, and the branches ($dist = \min_{i\in[1,9]} d_i$). The minimal distances computed for states on model's trajectories ($mindist$) are smaller and clearly separated from the random states distances. In order to compute the transitions represented in Fig.3b) we have used the threshold $dist = 3$ (states with $dist > 3$ were declared transient, whereas states with $dist < 3$ were classified as belonging to some branch.)

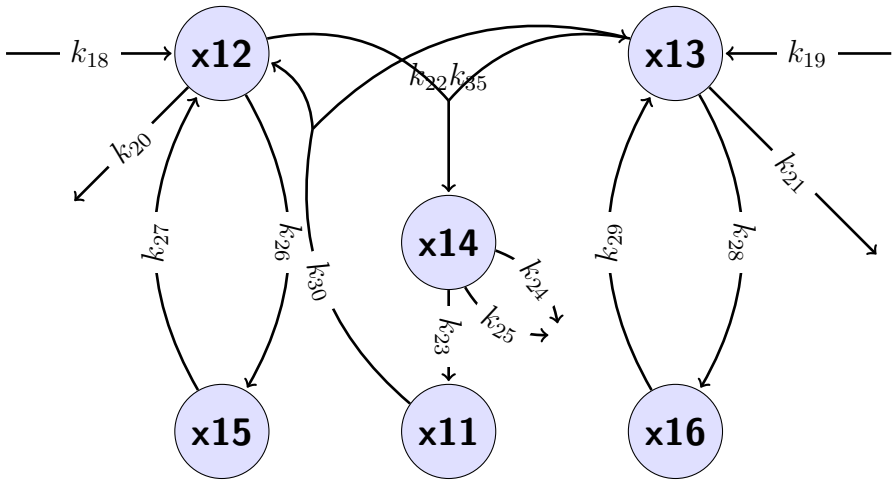

Figure 5: Graphic representation of the ligand-receptor module of the TGF-β full model. The different variables mean: x_{12} : RI (TGBR1), x_{13} : RII (TGFBR2), x_{14} : LR (ligand-receptor complex), x_{15} : RIe (TGFBR1 in endosome), x_{16} : RIIe (TGFBR2 in endosome), x_{11} : LRe (LR in endosome).

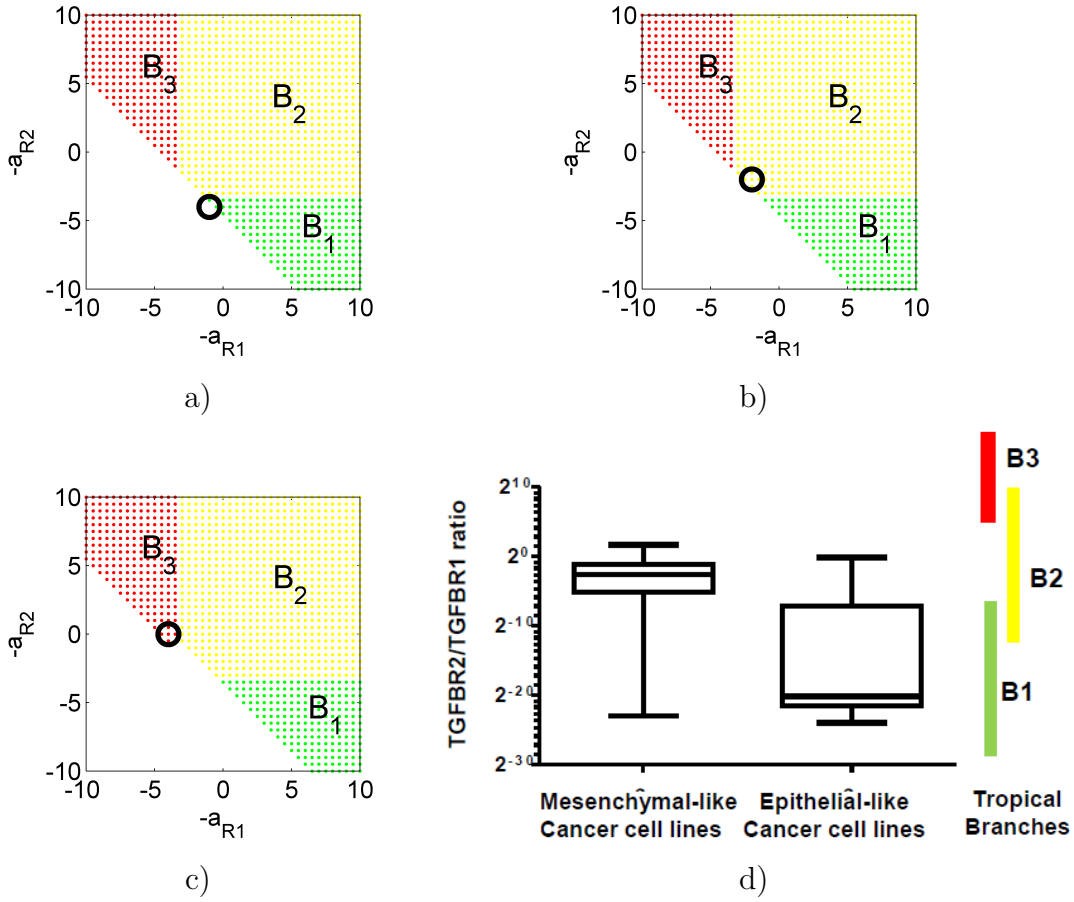

Figure 6: Tropical equilibrations of the ligand-receptor modules for various values of TGFBR2 (R2) gene expression are represented in projection in the plane (a_{R1}, a_{R2}) (a point in this plane provides the orders of concentrations of the protein receptors) (a,b, and c) and comparison with proteomics data from (Gholami et al., 2013) (d). Branches of tropical equilibrations are calculated for a) nominal value k_{19} (TGFBR2 expression) (this is the same as Fig. 3d in projection onto the plane (a_{R1}, a_{R2})), b) $\times 2$ TGFBR2 overexpression , and c) $\times 10$ TGFBR2 overexpression. The circle represents the position of the stable steady state and the branch containing is an attractor of the finite-state automaton. Like in Fig. 3 large opposite concentration orders $-a_i$ indicate large concentrations. All order calculations were performed using $\varepsilon = 1/11$. d) Proteomic data from NCI-60 cancer cell lines. Aggressive lines cover ratios of receptor concentrations intervals (indicated as bars at the right side of the sub-figure) corresponding to branches B3 (red) and B2 (yellow), whereas non-aggressive lines correspond to low expression of TGFBR2 in branch B1 (green). The receptors concentration ratios are well separated in the two classes (Mann-Whitney test, p-value 0.0006).

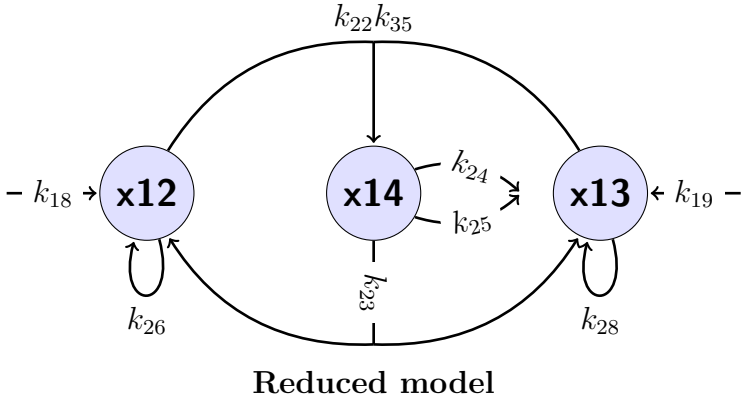

Figure 7: In order to compute the branches of tropical equilibration of the ligand-receptor module we use a reduced model. The reduced model is not necessarily a good approximation of the full dynamics, but has exactly the same tropical solutions as the full model. The different variables mean: x_{12} : RI (TGBR1), x_{13} : RII (TGFBR2), x_{14} : LR (ligand-receptor complex), x_{15} : RIe (TGFBR1 in endosome), x_{16} : RIIe (TGFBR2 in endosome), x_{11} : LRe (LR in endosome).