Controllability of reaction systems

Sergiu Ivanov1 · Ion Petre2,3

Received: 17 July 2020 / Accepted: 6 October 2020 / Published online: 19 November 2020
© Springer Nature Singapore Pte Ltd. 2020

Abstract
Controlling a dynamical system is the ability of changing its configuration arbitrarily through a suitable choice of inputs. It is a very well-studied concept in control theory, with wide-ranging applications in medicine, biology, social sciences and engineering. We introduce in this article the concept of controllability of reaction systems as the ability of transitioning between any two states through a suitable choice of context sequences. We show that the problem is \textsc{PSPACE}-hard. We also introduce a model of oncogenic signalling based on reaction systems and use it to illustrate the intricacies of the controllability of reaction systems. This study opens up a new line of research on the dynamic properties of reaction systems and it introduces a new, intricate biomedical model based on reaction systems.

Keywords Reaction systems · Controllability · Oncogenic signalling · Computational complexity

1 Introduction

Reaction systems are a biologically inspired model of computing originally introduced in [15]. They capture two fundamental interactions typically present between biochemical entities—activation and inhibition. Reaction systems are dynamical systems: reactions transform a set of reactants into a set of products provided that none of its inhibitors are present, which are then transformed further into other products, etc.

The reactions in reaction systems are governed by two fundamental principles: the \textit{threshold principle} and the \textit{non-permanency principle}. The threshold principle stipulates that when a resource is available it is available in unlimited amounts. This defines reaction systems as a qualitative modelling framework: their states are plain sets of species, bearing no quantitative information about amounts or concentrations. This also means that concurrent competition for resources is not implicitly included in the underlying semantics, but rather must be elucidated via some dedicated reactions and species. The non-permanency principle states that if a resource is not explicitly sustained/produced by reactions, it will disappear. The next state of a reaction system only consists of the species explicitly produced by the reactions enabled in the previous state.

Reaction systems are open systems: there is a notion of context that adds to the current state in each step of its dynamic process. The next state is produced by the reactions applied to the previous state plus the species added by the context.

Since their introduction in 2007, two major research directions on reaction systems have been established. The first direction focuses on their formal properties as a dynamical system: sequences of states [31, 32], power of small systems for various size measures [17, 34–36, 38], cycles and attractors [5, 14, 16], connections to propositional logic [33], etc. The second major direction of research on reaction systems consists in exploring their potential as a modelling framework, in particular for biological applications [3, 4, 10]. This direction sparked interest in model checking for reaction systems, i.e. formally defining relevant properties, evaluating their complexity, and designing algorithms for checking them [1, 2, 26, 28]. These works revealed a whole wealth of properties whose complexity ranges from polynomial to \textsc{PSPACE}-hard.
In this paper, we continue the study of potential applications of reaction systems to biological and medical research, and we focus on controllability. Intuitively, controllability of a dynamical system is the ability of driving this system to any one of its states, starting from any other state. Controllability is a strong property, which has attracted a lot of attention, especially in the case of biological networks (e.g. [24, 25]). Extensive research has been conducted into the practical feasibility of different variants of controllability for biological networks, and exact and approximate algorithms for finding ways to drive them, e.g. [12, 18, 22, 30]. These results have considerable potential for applications. For example, [6] identifies “driver neurons which can provide full control over the network” governing the actions of the C. elegans worm. Also, [18, 22, 25, 37] discuss applications in drug repurposing and personalised medicine.

The main goal of the present paper is defining controllability for reaction systems, and establishing some computational complexity evaluations. Since reaction systems are intrinsically open systems due to the context sequences governing their evolution, introducing controllability is quite natural: indeed, controls typically represent the actions of the environment. In addition to conventional, unrestricted controllability, we further define a restricted variant, similar to the notion typically used in network control theory, i.e. target controllability [12, 18]. We illustrate our definitions on a novel, oncogenic signalling model that we constructed based on the Boolean model of [41]. We show that imposing restrictions on the context sequences or on the allowed observables pushes the complexity of the controllability of reaction systems to PSPACE-hard.

This paper is structured as follows. Section 2 recalls the basic notions, in particular reaction systems, reachability, Boolean networks, and general controllability. Section 3 introduces the running example we will use to illustrate the notions of controllability, and which is translated from the Boolean model in [41]. Section 4 defines controllability for reaction systems, and Sect. 5 defines target controllability. Finally, Sect. 6 evaluates the complexity of the decision problems associated with our definitions.

2 Preliminaries

2.1 Reaction systems

We recall in this section some of the basic concepts of reaction systems. The presentation only aims to fix the notation and is kept brief. For details, we refer to [15].

Basic definitions. We introduce here the basic concepts around reaction systems.

Definition 1 Let $S$ be a finite (so-called background) set.

A reaction in $S$ is a triplet $r = (R, I, P)$ such that $R, I, P \subseteq S$, $R \cap I = \emptyset$. We call $R, I, P$ the reactant, inhibitor, and the product set of reaction $r$, resp. We also say that $R \cup I$ is the resource set of reaction $r$ and denote it $R \cup I = rsc(r)$. For a set $A$ of reactions, its resource set is $rsc(A) = \bigcup_{r \in A} rsc(r)$.

Let $T \subseteq S$. We say that reaction $r = (R, I, P)$ is enabled in $T$ if $R \subseteq T$ and $I \cap T = \emptyset$. The result of reaction $r$ on $T$ is $res_r(T) = P$ if $r$ is enabled on $T$ and it is $res_r(T) = \emptyset$ otherwise.

Let $R$ be a set of reactions in $S$ and $T \subseteq S$. The result of $R$ on $T$ is $res_R(T) = \bigcup_{r \in R} res_r(T)$.

A reaction system $A = (S, A)$ (over $S$) consists of a (finite) set of reactions $A$ in $S$. The result of $A$ on $T$ is $res_A(T) = \bigcup_{r \in A} res_r(T)$.

The states of reaction system $A = (S, A)$ are the subsets of $S$. We say that state $V \subseteq S$ is reachable in $A$ if it is in the codomain of $res_A$, i.e. $V = res_A(U)$, for some $U \subseteq S$.

Let $A = (S, A)$ be a reaction system and let $n \geq 1$. Let $\gamma = C_0, C_1, …, C_n \subseteq S$ be a sequence of so-called context sets. The interactive process $\pi_\gamma(A)$ in $A$ defined by $\gamma$ consists of two state sequences of length $n$ $\pi_\gamma(A) = (\delta_\gamma(A), \tau_\gamma(A))$ defined in the following way:

- $\delta_\gamma(A) = D_0, D_1, …, D_n$ is the result sequence of $\pi_\gamma(A)$ and $\tau_\gamma(A) = W_0, W_1, …, W_n$ is its state sequence;
- $D_0 = \emptyset$, and $D_i = res_A(C_{i-1} \cup D_{i-1})$ for all $i \in \{1, …, n\}$;
- $W_i = C_i \cup D_i$ for all $i \in \{0, 1, …, n\}$. $W_0$ is called the initial state of $\pi_\gamma(A)$. We say that $\pi_\gamma(A)$ starts in $W_0$ and ends in $W_n$ and denote it $W_n = res_A(W_0)$ (the result along context sequence $\gamma$).
- We say that state $V$ is reachable from state $U$ if there is a context sequence $\gamma$ such that $V = res_A(W_0)$.
- If $\gamma$ consists of empty (context) sets only, then $\pi_\gamma(A)$ is called context-independent.

Example 1 Consider the reaction system $A = (S, A)$, where $S = \{x, y, z\}$ and $A = \{r_1 : ((x), \{y\}, \{z\}), r_2 : ((y), \{x\}, \{z\})\}$. For convenience, we label the reactions in $A$. Reaction $r_1$ is enabled in $\{x\}$, but not in $\{x, y\}$ or $\emptyset$. The result of $r_1$ on $\{x\}$ is $res_A(\{x\}) = \{z\}$. The result of $A$ on $\{x\}$ is $res_A(\{x\}) = \{z\}$ again, because $r_2$ is not enabled in $\{x\}$.

Consider the sequence of contexts $\gamma : C_0 = \{x\}, C_1 = \{y\}, C_2 = \emptyset$. The interactive process $\pi_\gamma(A)$ consists of the of the following state sequences:

- result sequence: $\delta_\gamma(A) : D_0 = \emptyset, D_1 = \{z\}, D_2 = \{z\}$;
- state sequence: $\tau_\gamma(A) : W_0 = \{x\}, W_1 = \{y, z\}, W_2 = \{z\}$.

Therefore, $\{z\}$ is reachable from $\{x\}$ as there exists (at least) the context sequence $\gamma$ driving $A$ from $\{x\}$ to $\{z\}$.
The dynamic processes defined by interactive processes can be seen as state transition systems. We define the state transition graph of a reaction system as the graph having as nodes the states of the reaction system and the edges defined by the result function of the reaction system as follows.

**Definition 2** Let \( A = (S, \mathcal{A}) \) be a reaction system and \( I \subseteq S \). The \( I \)-context graph of reaction system \( A \) is the graph \( G_A^I = (\mathcal{P}(S), E) \), where the set of edges \( E \) is \( E = \{ (X, Y) \mid \exists C \subseteq I : \text{res}_A(X \cup C) = Y \} \).

In the definition of the context graph, we restrict the context sets to be subsets of a given “input” set \( I \) – this will be useful when we introduce the notions of controllability for reaction systems. For \( I = S \), there is no constraint on the context sets and we obtain the usual transition system associated to a reaction system, see [15].

Model checking for reaction systems has been considered in a number of different setups, based on, e.g. temporal logic ([26–28]), and computational complexity ([1, 2, 5]). We recall here the result on the reachability problem [13].

**Theorem 1** ([13]) Deciding if a state \( V \) of a reaction system \( A \) is reachable from a state \( U \) is \( \text{PSPACE-complete} \).

**Example 2** The following is a Boolean network over two variables \( x \) and \( y \), in which \( x \) alternates between 0 and 1, while \( y \) quickly converges to 0 and keeps this value indefinitely:

\[
\begin{align*}
x &= \neg x \\
y &= y \land x
\end{align*}
\]

We recall a standard way of simulating a single Boolean function with reaction systems, proposed in [9, 15]. Let \( f : \mathbb{B}^X \to \mathbb{B} \) be a Boolean function, \( f \not\in X \), and \( \varphi \) a propositional formula in minimal disjunctive normal form implementing \( f: \varphi = \bigvee C_i \), where \( C_i \) are the conjuncts appearing in \( \varphi \). We will use the notation \( \text{pos}(C_i) \) to refer to the set of variables appearing without a negation in \( C_i \) and \( \text{neg}(C_i) \) to refer to the set of variables appearing with a negation in \( C_i \). Motivated by our biological running example, we also use a new variable \( t_f \) to \( f \) that will be used to block the production of \( f \) by inhibiting all reactions producing \( f \). From the mathematical point of view, these extra inhibitors are not needed for the equivalence between Boolean networks and reaction systems.

The following set of reactions corresponds to \( \varphi \):

\[
A_\varphi = \{ (\text{pos}(C_i), \text{neg}(C_i) \cup \{ t_f \}, \{ f \}) \mid C_i \in \varphi \},
\]

where \( t_f \) is an output-specific inhibitor, which we will later use inhibit all reactions producing a specific species \( f \). The reaction system \( A_\varphi = (X \cup \{ t_f \}, A_\varphi) \) simulates \( f \) in the following sense. Consider any truth assignment \( s : X \to \mathbb{B} \) and construct the corresponding subset of variables \( W_s = \{ x \in X \mid s(x) = 1 \} \), using \( s \) as a indicator function. Then \( \text{res}_A(W_s) = \{ f \} \) if and only if \( f(s) = 1 \). Indeed, \( A_\varphi \) only produces \( f \) on \( W_s \), if only if there exists a conjunction \( C_i \in \varphi \) for which \( \text{pos}(C_i) \subseteq W_s \) and \( \text{neg}(C_i) \cap W_s = \emptyset \).

Consider now the Boolean network \( BN = (f_1, \ldots, f_n) \) over the set of variables \( X = \{ x_1, \ldots, x_n \} \), in which \( f_j \) is used to update the variable \( x_j \). The following set of reactions is obtained by iterating the idea above for each Boolean function of the network:

\[
A_{BN} = \{ (\text{pos}(C), \text{neg}(C) \cup \{ x_i \}, \{ x_i \}) \mid C \in \varphi, 1 \leq i \leq n \},
\]

where \( \varphi \) denotes the Boolean formula in minimal disjunctive normal form implementing the update function \( f_j \). The reaction system \( A_{BN} = (X, A_{BN}) \) simulates the *synchronous* model of dynamics of the Boolean network \( BN \).

### 2.2 Controllability of linear dynamical systems

We introduce here briefly the controllability of linear dynamical systems. For a more detailed presentation, we refer to [12, 22, 30]. We only discuss here a few basic concepts to guide our definitions of the controllability of reaction systems. Intuitively, a linear dynamical system consists of a set of nodes (variables) influencing each other’s dynamics through linear, one-source/one-target interactions. They can also be influenced through external, arbitrary interventions. The goal is to be able to change the configuration of the system from any initial state to any final state through a suitable choice of external interventions (that depend on the initial and desired final state). A system having this property is called controllable. A linear dynamical system is always trivially controllable by adding external interventions on all its nodes, which conveniently change its state as desired. The typical question to ask is one of optimisation: given a linear dynamical system, find the minimal set of external interventions making it controllable. We introduce these concepts formally in the following.
A linear dynamical system is a vector $\mathbf{x}$ of functions $\mathbf{x} : \mathbb{R} \to \mathbb{R}^n$, $n \geq 1$, defined as the solution of the system of ordinary differential equations

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{A}\mathbf{x}(t),$$  \hspace{1cm} (1)$$

for some fixed initial value $\mathbf{x}(0) \in \mathbb{R}^n$. The matrix $\mathbf{A}$ defining the dynamical system is an $n \times n$ real-valued matrix. The $(i, j)$ entry of matrix $\mathbf{A}$ describes the influence of the $j$-th node of the system over its $i$-th node.

A linear dynamical system can also be influenced through an external contribution, thought of as a parameter $m$-dimensional vector $\mathbf{u}$ of real functions, influencing the $n$ nodes of the dynamical system through a matrix $\mathbf{B} \in \mathbb{R}^{m \times n}$. In this case, the linear dynamical system is defined as the solution of the following system of ordinary differential equations:

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{A}\mathbf{x}(t) + \mathbf{Bu}(t),$$  \hspace{1cm} (2)$$

and it is called the $(\mathbf{A}, \mathbf{B})$ linear dynamical system.

One can also define a subset of so-called target nodes of the dynamical system on which the behaviour of the system is observed: $T = \{t_1, t_2, \ldots, t_l\} \subseteq \{1, 2, \ldots, n\}$, $1 \leq l \leq n$, $t_i < t_j$ for $1 \leq i < j \leq n$. The set of target nodes can be defined through its $0/1$-valued characteristic matrix $\mathbf{C}_T \in \mathbb{R}^{l \times n}$ defined as follows: $\mathbf{C}_T(i,j) = 1$ if and only if $t_i = j$. Otherwise, if $T = \{1, 2, \ldots, n\}$, then $\mathbf{C}_T$ is the identity matrix. A targeted linear dynamical system is defined by a triplet $(\mathbf{A}, \mathbf{B}, \mathbf{T})$.

We say that a dynamical system $(\mathbf{A}, \mathbf{B})$ is controllable if for any $\mathbf{x}(0) \in \mathbb{R}^n$ and any $\alpha \in \mathbb{R}^n$, there is an input function vector $\mathbf{u}_{\mathbf{x}(0), \alpha} : \mathbb{R} \to \mathbb{R}^n$ such that the solution $\mathbf{x}$ of (2) satisfies the property $\mathbf{x}(\tau) = \alpha$, for some $\tau \geq 0$.

We say that a targeted dynamical system $(\mathbf{A}, \mathbf{B}, \mathbf{T})$ is target controllable if for any $\mathbf{x}(0) \in \mathbb{R}^n$ and any $\gamma \in \mathbb{R}^l$, there is an input function vector $\mathbf{u}_{\mathbf{x}(0), \gamma} : \mathbb{R} \to \mathbb{R}^n$ such that the solution $\mathbf{x}$ of (2) satisfies the property $\mathbf{C}_T\mathbf{x}(\tau) = \gamma$, for some $\tau \geq 0$. In other words, the solution eventually matches $\gamma$ on its $T$-components. Obviously, for $T = \{1, 2, \ldots, n\}$, target controllability is identical to controllability.

The (target) controllability problem has an elegant algebraic characterisation known as Kalman’s condition.

**Theorem 2** (Kalman’s condition [21]) A targeted linear dynamical system $(\mathbf{A}, \mathbf{B}, \mathbf{T})$ is target controllable if and only if its controllability matrix $[\mathbf{C}_T\mathbf{B}, \mathbf{C}_T\mathbf{AB}, \mathbf{C}_T\mathbf{A}^2\mathbf{B}, \ldots, \mathbf{C}_T\mathbf{A}^{n-1}\mathbf{B}]$ is of full rank.

**Example 3** Consider the matrix $\mathbf{A} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. One solution of (1) is the vector $\mathbf{x}(t) = (\cos(t), \sin(t))^T$, describing the motion of a point along the unit circle. Take now $\mathbf{B} = (1, 1)^T$, i.e. we introduce one input which influences both nodes of the system equally. By Kalman’s condition, the linear system $(\mathbf{A}, \mathbf{B})$ is controllable: the matrix $[\mathbf{B}, \mathbf{AB}] = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ is of full rank (rank 2). Furthermore, if we take $T = \{1\}$, i.e. the only target is the first node of the system, we have $\mathbf{C}_T = (1, 0)$, and the matrix from Kalman’s condition is $[\mathbf{C}_T\mathbf{B}, \mathbf{C}_T\mathbf{AB}] = (1, -1)$, which is trivially full rank.

The controllability of linear dynamical system has found in the last few years many applications in biology and medicine [11, 18, 19, 22, 25]. In this context the linear dynamical system is an interaction (e.g. signalling) network describing the biological process of interest, and the input is in terms of available drugs or small inhibitors. The difficulty with this application of the concept is that the system is only partially defined, with the majority of the interactions impossible to measure, and thus with the matrices $\mathbf{A}$ and $\mathbf{B}$ only partially defined. The solution is a structural formulation of the problem, where controllability is defined in terms of the interaction network and not in terms of their precise strength, see, e.g. [11]. Also, the problem in this context is often given only through the interaction network (the equivalent of matrix $\mathbf{A}$ above). The goal is to identify a suitable set of input nodes making the system controllable, i.e. given matrix $\mathbf{A}$, the problem is to identify a suitable matrix $\mathbf{B}$ such that the linear dynamical system $(\mathbf{A}, \mathbf{B})$ is controllable. Furthermore, for applications in medicine, with the input being thought of as drugs delivered to a patient, there are various conditions imposed on matrix $\mathbf{B}$, such as having a minimal number of columns (corresponding to minimising the number of drugs), or having the non-zero entries of $\mathbf{B}$ only on certain rows (corresponding to selecting the drugs only from a certain set, e.g. FDA-approved drugs, or disease-specific standard drugs). Some examples on applying the controllability problem in medicine are in [11, 22, 25, 37]. Software for solving the controllability problem is available in [23, 30].

### 3 Running example: a reaction system model for breast cancer dynamics

Our running example is a reaction system modelling oncogenic signalling, with a focus on the receptor tyrosine kinase (RTK) signalling network and on the occurrence of uncontrollable proliferation. We follow the Boolean network model proposed in [41] and give it a correspondent in terms of reaction systems.

The model of [41] is grounded in the context of the breast cancer model of [40]. It consists of a simplified version of the RTK signalling network through the MAPK, PI3K, AKT
and mTORC1 pathways, including the cross-talk between them, and several feedback loops. The network includes both oncogene proteins (RAS, PI3K, mTOR) and tumour suppressors (Rb, FOXO3). The model is illustrated in Fig. 1a (the growth factors GF are not included in the figure). Each of the key proteins in Fig. 1 has a correspondent variable in the Boolean network model, with the update Boolean functions described in Table 1. Variable X gets value 0 if its correspondent protein is inactive and 1 if it is active. The outcome of the model is described by variable Prolif which, as an exception, is ternary-valued, with 0 standing for non-proliferation, 1 for proliferation, and 2 for uncontrolled proliferation. Depending on the signals reaching the output node, the model is in one of these three proliferation modes, as described by the update function for variable Prolif, \( f_{\text{Prolif}} = E2F + (\text{EIF4F} \land \text{S6K}) \). Because of the non-binary nature of this function, we replace the modeling of the proliferation status with two different variables, Prolif that models with true/false the proliferation/non-proliferation status and UProlif that models with true/false the uncontrolled proliferation status. Their Boolean functions are defined as follows:

\[
\begin{align*}
 f_{\text{Prolif}} &= (E2F \land \neg\text{EIF4F}) \lor (E2F \land \neg\text{S6K}) \lor (\neg E2F \land \text{EIF4F} \land \text{S6K}), \\
 f_{\text{UProlif}} &= E2F \land \text{EIF4F} \land \text{S6K}.
\end{align*}
\]

This model can reach a non-proliferation, a proliferation and an uncontrolled proliferation status, depending on the different activation statuses of the MAPK-, PI3K-, AKT- and mTORC1-paths, see [41] and Fig. 1b–d on page 12.

The reaction system associated to the Boolean network in Table 1 has as its background set all the variables of the Boolean network and an inhibitor variable associated to each of them: \( S = \{ \text{GF}, \text{RTK}, \text{R}_{\text{RTK}}, \text{RAS}, \text{R}_{\text{RAS}}, \text{MAPK}, \text{MAPK}, \text{PI3K}, \text{pPI3K}, \text{PI3}, \text{pPI3}, \text{FOXO3}, \text{pFOXO3}, \text{AKT}, \text{iAKT}, \text{cycE/CDK2}, \text{cycE/CDK2}, \text{Rb}, \text{R}_{\text{Rb}}, \text{E2F}, \text{E2F}, \text{TSC}, \text{iTSC}, \text{PRAS40}, \text{iPRAS40}, \text{mTORC1}, \text{i}_{\text{mTORC1}}, \text{EIF4F}, \text{EIF4F}, \text{S6K}, \text{iS6K}, \text{Prolif}, \text{iProlif}, \text{UProlif}, \text{iUProlif} \} \). The set of reactions is listed in Table 2. These reactions were obtained from the Boolean equations given in [41] by applying the conversion technique we describe in Sect. 2.1.

In the reaction system model, the equivalent of a component/signal \( X, X \in \{ \text{GF}, \text{RTK}, \text{RAS}, \text{MAPK}, \text{PI3K}, \text{PI3}, \text{FOXO3}, \text{AKT}, \text{cycE/CDK2}, \text{Rb}, \text{E2F}, \text{TSC}, \text{PRAS40}, \text{mTORC1}, \text{EIF4F}, \text{S6K} \} \) being active/inactive is having (not having, resp.) \( X \) in the current state. Synthetically activating/inactivating \( X \) corresponds to the context adding to the current state \( X \) (\( i_{X} \), resp.). A non-proliferation configuration corresponds to a state that does not include either symbol Prolif or UProlif. A proliferation/uncontrolled proliferation configuration corresponds to a state including Prolif (UProlif, resp.)

Table 3 shows an interactive process of our reaction systems model, running with the constant context \( \{ \text{GF} \} \), and oscillating between proliferation and uncontrolled proliferation. Such interactive processes can be constructed using software tools like [20] or [29].

## 4 Controllability of reaction systems

The basic concept of controllability of a reaction system \( \mathcal{A} = (S,A) \) is that for any \( X, Y \subseteq S \), \( X \neq Y \), there is an interactive process of \( \mathcal{A} \) starting in \( X \) and ultimately reaching \( Y \). More exactly, the controllability problem \( (\mathcal{A}, X, Y) \) consists in finding a context sequence \( C = (C_{i})_{0 \leq i < n} \) for \( \mathcal{A} \) such that the interactive process generated by \( C \) starts in \( X \) and ends in \( Y \). \( \mathcal{A} \) is said to be controllable if the controllability problem \( (\mathcal{A}, X, Y) \) has a solution for any pair \( X, Y \subseteq S \), \( X \neq Y \).

Similarly as in the case of linear dynamical systems, in the absence of constraints on the context sequences, the controllability problem has a trivial solution: for any \( X, Y \subseteq S \), the control sequence \( \gamma_{X \rightarrow Y} = X, S, Y \) leads to an interactive process starting in \( X \), going to the empty state (since all reactions are disabled through getting the full set \( S \) as a context), and then going to \( Y \):

| Context | \( X \) | \( S \) | \( Y \) |
|---------|-------|-------|-------|
| Result  | \( \emptyset \) | \( \text{res}_{X}(X) \) | \( \emptyset = \text{res}_{Y}(S) \) |
| State   | \( X \) | \( S \) | \( Y \) |

Instead, we define controllability of reaction systems as follows.

**Definition 3** Let \( \mathcal{A} = (S,A) \) be a reaction system.

- For some nonnegative integer \( n \) with \( 0 \leq n < |S| \), we say that \( \mathcal{A} \) is \( n \)-controllable if for any \( X, Y \subseteq S \), with \( Y \) reachable in \( \mathcal{A} \), there is a context sequence consisting of sets of cardinality at most \( n \) generating an interactive process starting in \( X \) and ending in \( Y \).

- For some \( I \subseteq S \), we say that \( \mathcal{A} \) is \( I \)-controllable if for any \( X, Y \subseteq S \), with \( Y \) reachable in \( \mathcal{A} \), there is a context sequence consisting of subsets of \( I \) generating an interactive process starting in \( X \) and ending in \( Y \).

We define two versions of the controllability problem for reaction systems as follows. Let \( \mathcal{A} = (S,A) \) be a reaction system.

Problem \( C(A,n) \) For some nonnegative integer \( n \) with \( 0 \leq n < |S| \), decide if \( \mathcal{A} \) is \( n \)-controllable. Also, find the smallest such \( n \).
Controllability of reaction systems
Fig. 1 The signal transduction network of [41] characterising cell proliferation. The MAPK pathway is in yellow, the PIP3 pathway is in green, the AKT-pathway is in dark green, the mTORC1-pathway is in orange, and the cye/CK2-pathway is in blue. Network configurations leading to b uncontrolled proliferation, c proliferation, and d no proliferation. In b–d, a rectangle with blue background denotes an inactive component, and one with yellow denotes an active component.

Problem $C(A, I)$ For some $I \subseteq S$, decide if $A$ is $I$-controllable. Also, find a minimal (with respect to inclusion) such set $I$.

Example 4 The interactive process in Table 4 is an example on how to induce a change of state from the attractor state $S_{19}$ (with uncontrolled proliferation) in Table 3 to state $\{\text{RTK, RAS, MAPK, FOXO3, cye/CDK2, E2F, TSC, PRAS40, Prolif}\}$ (with Proliferation). We use just one additional symbol in the context sequence, $t_{\text{PI3K}}$, inhibiting all reactions producing PI3K.

While this example shows that, in this particular setting, the reaction system can be driven to a desired target state, it does not show whether the reaction system is controllable. Controllability would require the possibility to drive the system from any state to any state.

5 Target controllability of reaction systems

The concept of target controllability of reaction systems is focused on a given set of target nodes $T \subseteq S$ of a reaction system $A = (S, A)$. The objective in this problem is to be able, through a suitable context sequence, to transition between any two subsets of $T$. The caveat here is that, with the focus set on $T$ only, elements from $S \setminus T$ may be present arbitrarily in the interactive process, including in the initial and final states. We formalise this concept as follows.

Definition 4 Let $A = (S, A)$ be a reaction system and $T \subseteq S$ the set of targets.

- For some nonnegative integer $n$ with $0 \leq n < |S|$, we say that $(A, T)$ is $n$-target controllable if for any $X, Y \subseteq T$, with $Y$ or a superset of it reachable in $A$, there is a context sequence consisting of sets of cardinality at most $n$ generating an interactive process starting in a state $W_0$ with $W_0 \cap T = X$ and ending in a state $W_r$ with $W_r \cap T = Y$, for some $r \geq 0$.
- For some $I \subseteq S$, we say that $(A, T)$ is $I$-target controllable if for any $X, Y \subseteq T$, with $Y$ or a superset of it reachable in $A$, there is a context sequence consisting of subsets of $I$ generating an interactive process starting in a state $W_0$ with $W_0 \cap T = X$ and ending in a state $W_r$ with $W_r \cap T = Y$, for some $r \geq 0$.

We define two versions of the target controllability problem for reaction systems as follows. Let $A = (S, A)$ be a reaction system and $T \subseteq S$ the set of targets.

Problem $TC(A, T, n)$ For some nonnegative integer $n$ with $0 \leq n < |S|$, decide if $(A, T)$ is $n$-target controllable. Also, find the smallest such $n$.

Problem $TC(A, T, I)$ For some $I \subseteq S$, decide if $(A, T)$ is $I$-target controllable. Also, find a minimal (with respect to inclusion) such set $I$.

Example 5 A natural target of our running example is the proliferation symbol. Consider for example driving the model away from the state $S_{19}$ in Table 3, characterised by uncontrolled proliferation, and into a steady state or an attractor consisting of states characterised by proliferation or no proliferation. There are several ways of doing that with constant context sequences that still include the GF symbol, including the following options:

- the context set $\{\text{GF, PRAS40}\}$ drives the model into an attractor consisting of 10 states, all of them with Prolif;
- the context set $\{\text{GF, } t_{\text{cytE/CDK2}}, \text{PRAS40}\}$ drives the model into an attractor consisting of 11 states, 6 with Prolif, and 5 without;
- the context set $\{\text{GF, } t_{\text{cytE/CDK2}}, \text{PRAS40}\}$ drives the model into an attractor consisting of 10 states, all of them with no proliferation;
- the context set $\{\text{GF, } t_{\text{PI3K}}, t_{\text{cytE/CDK2}}\}$ drives the model into a steady state with no proliferation, with the interactive process shown in Table 5.

Similarly to Example 4, Example 5 shows that particular states can be reached from the attractor state $S_{18}$, without proving whether the reaction system is controllable or target controllable.

6 Complexity results for the controllability of reaction systems

In this section, we show that while unrestricted controllability is trivial, it may become very complex if additional requirements are imposed on context sequences. This observation is also valid for target controllability.
Table 1 The Boolean network model of [41] for oncogenic signalling in disjunctive normal form

| $f_{RTK}$ | $(G \land \neg FOXO3) \lor (G \land \neg S6K \land \neg MAPK)$ |
| $f_{RAS}$ | $RTK$ |
| $f_{MAPK}$ | $RAS \lor PIP3$ |
| $f_{PI3K}$ | $RTK \lor PIP3$ |
| $f_{PI3K}$ | $\neg AKT \lor \neg MAPK$ |
| $f_{PI3K}$ | $PI3K$ |
| $f_{MAPK}$ | $RAS \lor PIP3$ |
| $f_{cycle/CDK2}$ | $(AKT \land \neg FOXO3) \lor E2F$ |
| $f_{RB}$ | $\neg RB$ |
| $f_{TSC}$ | $\neg MAPK \lor \neg AKT$ |
| $f_{PRAS40}$ | $\neg AKT$ |
| $f_{mTORC1}$ | $\neg TSC \land \neg PRAS40$ |
| $f_{E1F4F}$ | $mTORC1$ |
| $f_{S6K}$ | $mTORC1$ |

6.1 Controllability is PSPACE-hard

We will show that $I$-controllability is at least as hard as reachability, which is PSPACE-complete [13].

**Theorem 3** Let $A = (S, A)$ be a reaction system and $I \subseteq S$. The problem $C(A, I)$ of deciding whether $A$ is $I$-controllable is PSPACE-hard.

**Proof** We will consider the particular case of $I$-controllability in which $I = \emptyset \subseteq S$. Deciding that $A$ is $\emptyset$-controllable is equivalent to deciding whether, for any pair of subsets $X, Y \subseteq S$, $Y$ is reachable from $X$. Since reachability for reaction systems is PSPACE-complete, we conclude that $\emptyset$-controllability and therefore $I$-controllability of reaction systems is PSPACE-hard.

Similarly, reachability can be reduced to $n$-controllability.

**Theorem 4** Let $A = (S, A)$ be a reaction system and $0 \leq n < S$. The problem $C(A, n)$ of deciding whether $A$ is $n$-controllable is PSPACE-hard.

**Proof** As in the proof of Theorem 3, we consider the particular case $n = 0$, which only allows empty sets as contexts. In this case, $n$-controllability of $A$ is equivalent to deciding whether, for any pair of subsets $X, Y \subseteq S$, $Y$ is reachable from $S$. This implies that $n$-controllability is PSPACE-hard.

6.2 Target controllability is PSPACE-hard

Target controllability can be reduced to “full” controllability by allowing any species to be a control target: $T = S$. This directly implies that $I$-target controllability and $n$-target controllability are both PSPACE-hard.

**Corollary 1** Let $A = (S, A)$ be a reaction system and $I, T \subseteq S$. The problem $TC(A, T, I)$ of deciding whether $A$ is $I$-controllable is PSPACE-hard.

**Corollary 2** Let $A = (S, A)$ be a reaction system, $I \subseteq S$, and $0 \leq n \leq |S|$. The problem $TC(A, T, n)$ of deciding whether $A$ is $n$-controllable is PSPACE-hard.

These results are degenerate in the sense that they focus on the situations in which control inputs are disallowed: $I = \emptyset$ or $n = 0$. We will now show that $I$-controllability for reaction systems is PSPACE-hard even when some control inputs must be provided.
Table 3 An interactive process of the reaction systems model for oncogenic signalling

| Context | GF | GF | GF |
|---------|----|----|----|
| State | \( S_1 = \{ \text{RTK, FOXO3, Rb, E2F, TSC, PRAS40, mTORC1} \} \) | \( S_2 = \{ \text{RTK, RAS, MAPK, PI3K, PIP3, AKT, cycE/CDK2, TSC, PRAS40, EIF4F, S6K, Prolif} \} \) | \( S_3 = \{ \text{RTK, RAS, MAPK, PI3K, PIP3, FOXO3, TSC, PRAS40, Prolif} \} \) |
| Status | No proliferation | Proliferation | Proliferation |
| Context | GF | GF | GF |
| State | \( S_4 = \{ \text{RTK, RAS, MAPK, PI3K, PIP3, AKT, cycE/CDK2, E2F, TSC, PRAS40} \} \) | \( S_5 = \{ \text{RTK, RAS, MAPK, PI3K, PIP3, AKT, cycE/CDK2, E2F, Rb, TSC, PRAS40, Prolif} \} \) | \( S_6 = \{ \text{RTK, RAS, MAPK, PI3K, PIP3, AKT, cycE/CDK2, E2F, mTORC1, Prolif} \} \) |
| Status | No proliferation | Proliferation | No proliferation |
| Context | GF | GF | GF |
| State | \( S_7 = \{ \text{MAPK, PI3K, PIP3, AKT, cycE/CDK2, E2F, mTORC1, EIF4F, S6K} \} \) | \( S_8 = \{ \text{MAPK, AKT, cycE/CDK2, E2F, mTORC1, EIF4F, S6K, Prolif} \} \) | \( S_9 = \{ \text{MAPK, AKT, cycE/CDK2, E2F, mTORC1, EIF4F, S6K, Prolif} \} \) |
| Status | No proliferation | Uncontr. prolif. | Uncontr. prolif. |
| Context | GF | GF | GF |
| State | \( S_{10} = \{ \text{cycE/CDK2, E2F, mTORC1, EIF4F, S6K, Uprolif} \} \) | \( S_{11} = \{ \text{FOXO3, cycE/CDK2, E2F, TSC, PRAS40, EIF4F, S6K, Uprolif} \} \) | \( S_{12} = \{ \text{RTK, FOXO3, cycE/CDK2, E2F, TSC, PRAS40, EIF4F, S6K, Uprolif} \} \) |
| Status | Uncontr. prolif. | Uncontr. prolif. | Uncontr. prolif. |
| Context | GF | GF | GF |
| State | \( S_{13} = \{ \text{RTK, RAS, MAPK, PI3K, PIP3, FOXO3, cycE/CDK2, E2F, TSC, PRAS40, Uprolif} \} \) | \( S_{14} = \{ \text{RTK, RAS, MAPK, PI3K, PIP3, FOXO3, cycE/CDK2, E2F, TSC, PRAS40, Prolif} \} \) | \( S_{15} = \{ \text{RTK, RAS, MAPK, PI3K, PIP3, AKT, cycE/CDK2, E2F, TSC, PRAS40, Prolif} \} \) |
| Status | Uncontr. prolif. | Proliferation | Proliferation |
| Context | GF | GF | GF |
| State | \( S_{16} = \{ \text{RTK, RAS, MAPK, PI3K, PIP3, AKT, cycE/CDK2, E2F, Prolif} \} \) | \( S_{17} = \{ \text{RTK, RAS, MAPK, PI3K, PIP3, AKT, cycE/CDK2, E2F, mTORC1, Prolif} \} \) | \( S_{18} = \{ \text{MAPK, PI3K, PIP3, AKT, cycE/CDK2, E2F, mTORC1, EIF4F, S6K, Prolif} \} \) |
| Status | Proliferation | Proliferation | Proliferation |
| Context | GF | GF | GF |
| State | \( S_{19} = \{ \text{MAPK, PIP3, AKT, cycE/CDK2, E2F, mTORC1, EIF4F, S6K, Uprolif} \} = S_8 \) | \( S_{20} = \{ \text{...} \} \) | \( S_{21} = \{ \text{...} \} \) |
| Status | Uncontr. prolif. | ... | ... |
Table 4 An interactive process of the reaction systems model for oncogenic signalling switching from an uncontrolled proliferation status to a proliferation status

| Context | GF | GF, t_{P3K} | GF, t_{P3K} |
|---------|----|-------------|-------------|
| State  | X₀ = S₁₀ = \{MAPK, PIP3, AKT, cycE/CDK2, E2F, mTORC1, EIF4F, S6K, UProlif\} | X₁ = \{MAPK, AKT, cycE/CDK2, E2F, mTORC1, EIF4F, S6K, UProlif\} | X₇ = \{cycE/CDK2, E2F, mTORC1, EIF4F, S6K, UProlif\} |
| Status  | Uncontr. prolif. | Uncontr. prolif. | Uncontr. prolif. |
| Context | GF, t_{P3K} | GF, t_{P3K} | GF, t_{P3K} |
| State  | X₃ = \{FOXO3, cycE/CDK2, E2F, TSC, PRAS40, mTORC1, EIF4F, S6K, UProlif\} | X₄ = \{RTK, FOXO3, cycE/CDK2, E2F, TSC, PRAS40, EIF4F, S6K, UProlif\} | X₅ = \{RTK, RAS, FOXO3, cycE/CDK2, E2F, TSC, PRAS40, UProlif\} |
| Status  | Uncontr. prolif. | Uncontr. prolif. | Uncontr. prolif. |
| State  | X₆ = \{RTK, RAS, MAPK, FOXO3, cycE/CDK2, E2F, TSC, PRAS40, Prolif\} | X₇ = X₆ | X₇ = X₆ |
| Status  | Proliferation | Proliferation | Proliferation |

Table 5 An interactive process of the reaction systems model for oncogenic signalling switching from an uncontrolled proliferation to a steady state with no proliferation

| Context | GF | GF, t_{P3K} \cdot t_{cycE/CDK2} | GF, t_{P3K} \cdot t_{cycE/CDK2} |
|---------|----|-------------------------------|-------------------------------|
| State  | Y₀ = S₁₀ = \{MAPK, PIP3, AKT, cycE/CDK2, E2F, mTORC1, EIF4F, S6K, UProlif\} | Y₁ = \{MAPK, AKT, E2F, mTORC1, EIF4F, S6K, UProlif\} | Y₇ = \{RB, E2F, mTORC1, EIF4F, S6K, UProlif\} |
| Status  | Uncontr. prolif. | Uncontr. prolif. | Uncontr. prolif. |
| Context | GF, t_{P3K} \cdot t_{cycE/CDK2} | GF, t_{P3K} \cdot t_{cycE/CDK2} | GF, t_{P3K} \cdot t_{cycE/CDK2} |
| State  | Y₃ = \{FOXO3, Rb, TSC, PRAS40, mTORC1, EIF4F, S6K, UProlif\} | Y₄ = \{RTK, FOXO3, Rb, TSC, PRAS40, EIF4F, S6K, Prolif\} | Y₆ = \{RTK, RAS, FOXO3, Rb, TSC, PRAS40, Prolif\} |
| Status  | Uncontr. prolif. | Proliferation | Proliferation |
| Context | GF, t_{P3K} \cdot t_{cycE/CDK2} | GF, t_{P3K} \cdot t_{cycE/CDK2} | GF, t_{P3K} \cdot t_{cycE/CDK2} |
| State  | Y₆ = \{RTK, RAS, MAPK, FOXO3, Rb, TSC, PRAS40\} | Y₇ = Y₆ | Y₇ = Y₆ |
| Status  | No proliferation | No proliferation | No proliferation |

Theorem 5 Let \( \mathcal{A} = (S, A) \) be a reaction system and \( I, T \subseteq S \), such that \( I \neq \emptyset \). The problem \( TC(A, T, I) \) of deciding whether \( \mathcal{A} \) is \( I \)-controllable is \( \text{PSPACE} \)-hard.

Proof Consider an arbitrary reaction system \( \mathcal{A} = (S, A) \), an extension of the background set \( S' \supseteq S \), and the reaction system \( \mathcal{A}' = (S', A) \) over the extended alphabet, but with the same reactions as \( \mathcal{A} \). Let \( I = S' \setminus S \) and \( T = S \). Then \( (\mathcal{A}', T) \) is \( I \)-target controllable if and only if \( \mathcal{A} \) is \( \emptyset \)-controllable.

Indeed, by construction the species from \( I = S' \setminus S \) have no influence on the reactions in \( A \). Therefore, given any pair of sets \( X, Y \subseteq S = T \), if \( \mathcal{A}' \) can reach \( Y \) from \( X \) with contexts from \( I \), it can reach \( Y \) from \( X \) with an empty context sequence. Since \( \mathcal{A}' \) and \( \mathcal{A} \) have exactly the same reactions, this also means that if \( \mathcal{A}' \) can reach \( Y \) from \( X \), \( \mathcal{A} \) can reach \( Y \) from \( X \) as well. Conversely, if \( \mathcal{A} \) can reach \( Y \) from \( X \) with a sequence of empty contexts, \( \mathcal{A}' \) can reach \( Y \) from \( X \) with a sequence of empty contexts, or indeed with any sequence of contexts from \( I \) of the same length.

We conclude that, for non-empty \( I \), \( I \)-target controllability is at least as hard as \( I \)-controllability, which proves the statement of the theorem. \( \square \)

The previous proof relies on restricting the set \( I \) to symbols not actually appearing in the reactions. We will now show that \( I \)-target controllability is \( \text{PSPACE} \)-hard even
when contexts are allowed to inject symbols appearing in
the reactants and the inhibitors of reactions.

Before stating and proving this result, we formulate
the following helper notion. Consider the reaction system
\(A = (S, A)\) and take an extension of the background set
\(S' \supseteq S\). We define the \textit{nonce-extension} \(\text{nonce}_{S \to S'}(A)\) of \(A\)
from \(S\) to \(S'\) in the following way:

\[
\text{nonce}_{S \to S'}(A) = \{(R \cup R', I \cup I', P) \mid a : (R, I, P) \in A, R' \subseteq S' \setminus S, I' \subseteq S' \setminus S\}.
\]

Note that \(A \subseteq \text{nonce}_{S \to S'}(A)\), because \(\emptyset \subseteq S' \setminus S\).

**Theorem 6** Let \(A = (S, A)\) be a reaction system and \(E, T \subseteq S\)
such that \(E \cap \text{rsc}(A) \neq \emptyset\). The problem \(\text{TC}(A, T, E)\) of deciding
whether \(A\) is \(E\)-controllable is \(\text{PSPACE}\)-hard.

**Proof** Consider a reaction system \(A = (S, A)\) and an
extension of the background set \(S' \supseteq S\). Construct now
a new reaction system \(A' = (S', A')\) with the properties
\(A \subseteq A' \subseteq \text{nonce}_{S \to S'}(A)\). Let \(E = S' \setminus S\) and \(T = S\). Then
\((A', T)\) is \(E\)-target controllable if and only if \(A\) is \(\emptyset\)-control-
lable. Stronger yet, for any \(Z \subseteq S'\), \(\text{res}_{A'}(Z) = \text{res}_A(Z \cap S)\).

Indeed, suppose reaction \(a' = (R, I, P) \in A'\) is enabled by
\(Z\), i.e. \(R \subseteq Z\) and \(I \cap Z = \emptyset\). Then, trivially, the reaction
\(a = (R \cap S, I \cap S, P)\) is enabled by \(Z \cap S\), and \(a \in A\) by
construction of \(A'\). On the other hand, if reaction \(a \in A\)
is enabled by \(Z \cap S\), then it is also enabled by \(Z\), because
\((Z \setminus S) \cap (R \cup I) = \emptyset\). \(\Box\)

7 Conclusions

The controllability problem is of high interest in dynamical
systems, having as its aim the ability to change the
system’s configuration through well chosen sequences of
external interventions. We initiated in this paper the study of
controllability for reaction systems. The reaction systems
framework has all the key ingredients necessary for a nat-
ural definition of the controllability problem: system dyna-
mics through interactive processes, state transitions, external
interventions through context sequences. We defined several
natural variants of the controllability problem for reaction
systems.

Introducing the concept of controllability to reaction
systems has many important applications, e.g. in model-
ling of complex diseases. Indeed, the environmental fac-
tors are naturally captured as adversary control actions on
the complex systems governing the function of the organ-
ism, its organs, or individual cells. On the other hand,
medical therapies can be represented as control inputs
meant to bring these systems back to the normal state.

We introduced the first reaction system-based onco-
genic signalling model in the literature, a model that we
believe will be of independent interest to the reaction
systems community. The model includes several of the
best studied cancer signalling pathways and follows their
interplay leading to tumour proliferation, both in the case
of external growth factor signals and in their absence.

We used this example to show how much diversity of options
there is in the concept of controllability for reaction sys-
tems. The complexity of dynamics shown through this
example anticipated the computational complexity results
we proved in this article, showing that the controllability
problem is \(\text{PSPACE}\)-hard.

Several topics of interest remain to be explored around
controllability of reaction systems, for example, and in no
particular order:

1. Are there formulations of the problem that are computa-
tionally easy (in the sense of computational complexity
theory), perhaps based on minimal reaction systems?
2. Define and study the concept of stable controllability,
where a constant (or an ultimately constant) context
sequence leads to the desired state, that moreover is a
steady state of the reaction system with the given con-
stant context sequence.
3. Find efficient heuristics for the controllability of reaction
systems, identifying (not necessarily optimal) context
sequences solving a given controllability problem.

We believe that these topics should give further insight into
the potential of reaction systems as a qualitative framework
for biomodelling.

**Funding** Sergiu Ivanov was partially supported by Computer Science
Network of Paris Île-de-France Region (project AAP DIM RFSI 2018-
03). Ion Petre was partially supported by a grant of the Romanian
Ministry of Education and Research, CCCDI - UEFISCDI (project number
PN-III-P2-2.1-PED-2019-2391), within PNCDI III.

**Data availability** All data analysed in this study (the reaction systems
model and the interactive processes) is described in full in the article.

**Compliance with ethical standards**

**Conflict of interest** The authors declare that they have no conflict of
interest.
networks. In: C. Cherifi, H. Cherifi, M. K., & Mirco, M. (Eds.), Complex networks & their applications VI—Proceedings of complex networks 2017 (the sixth international conference on complex networks and their applications), COMPLEX NETWORKS 2017, Lyon, France, November 29–December 1, 2017. Studies in computational intelligence (Vol. 689, pp. 567–579). Springer.

40. Zañudo, J. G. T., Scaltriti, M., & Albert, R. (2017). A network modeling approach to elucidate drug resistance mechanisms and predict combinatorial drug treatments in breast cancer. Cancer Convergence, 1(1), 5.

41. Zañudo, J. G. T., Steinway, S. N., & Albert, R. (2018). Discrete dynamic network modeling of oncogenic signaling: Mechanistic insights for personalized treatment of cancer. Current Opinion in Systems Biology, 9, 1–10.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Dr. Sergiu Ivanov is an associate professor (maître de conférences) of computer science at the Department of Computer Science of Université d’Évry, Université Paris-Saclay. He is currently member of the IBISC Laboratory and belongs to team COSMO, who study the fundamental properties of complex systems, and notably biological systems. Dr. Sergiu Ivanov defended his PhD thesis focused on the computational power of biologically inspired models of computing at Université Paris-Est in 2015. His principal research interest lies in applying formal models to gain insight into medical and biological problems. He is currently working on employing formal networks to help infer cancer therapies and better understand the molecular and evolutionary causes of this disease (e.g., resilience, atavistic theory of cancer, etc.). Dr. Sergiu Ivanov’s main research domains are: precision medicine, theoretical biology, complex systems, and bio-inspired models of computing.

Prof. Ion Petre is a professor of complex systems at Department of Mathematics and Statistics of University of Turku. He leads the Computational Biomodeling Laboratory at the Turku Centre for Computer Science in Finland. His research has a focus on computational systems biology and data science in the biomedical field. He has a strong research record with several international research awards. He is an author of over 100 peer-reviewed publications, and a speaker at many international conferences. His main research domains are: computational modeling and analysis of complex systems, network modeling, systems biology, bioinformatics, and data science. His recent projects include network controllability for cancer medicine, refinement-based modeling for large-scale dynamical models, and dynamic modeling for the heat shock response.