Interaction Models for Biochemical Reactions

Mila Majster-Cederbaum, Nils Semmelrock, Verena Wolf

***Abstract*—This paper presents a stochastic modelling framework for complex biochemical reaction networks from a component-based perspective. Our approach takes into account the discrete character of quantities of participating entities (i.e. the individual populations of the involved chemical species) and the inherent probabilistic nature of microscopic molecular collisions. We collect all instances of one species into one component, hence the size of the state space is reduced considerably. Additionally, the formalism allows for a modular description process and a graphical representation. It may serve as a link between very intuitive but informal models from biology and abstract mathematical models serving as input for analysis tools.**

We discuss example networks and illustrate the benefits of our approach compared with other compositional modelling formalisms (e.g. the stochastic π-calculus).

I. Introduction

Phenomena in biological systems observed at high levels of abstraction are determined by complex biochemical reaction networks happening within the living cell and involving various molecules such as DNA, mRNAs and proteins.

Historically, processes occurring within cells were investigated in terms of deterministic rate equations leading to a system of nonlinear ordinary differential equations (ODEs). They are based on information on concentrations of different molecular species. It came into evidence that many crucial events in living cells depend on the interaction of small numbers of molecules and hence are sensitive to the underlying stochasticity of the reaction processes [10]. Since modelling probabilistic outcomes is not possible with the classical ODE approach various other approaches to the modelling of intracellular randomness have been proposed [4], [12], [1], [20], [19], [3], [7], [22], [16].

In this paper we present a probabilistic component-based formalism for the precise and convenient description and analysis of intracellular processes. It is based on quantitative information about the populations of the participating species and rate constants of their potential chemical reactions.

The starting point of our model is some abstract representation of certain processes happening inside some biological compartment. For instance, a set of biochemical reactions is given. From this we construct a (stochastic) interaction model and based on this we may either run a simulation of the model or perform a numerical analysis of the underlying Markov process which can be deduced from the interaction model automatically.

As the formalism is compositional we may exploit the structure of the interaction model induced by the reaction network to use advanced analysis techniques for the underlying Markov process.

The formalism can be applied to different abstraction levels and gives a hierarchical view of the system. It is possible to have a fine-grained view on the system under study where each molecule is modelled separately as well as an abstract view in which all occurrences of a species are modelled by one component. Additionally, what is considered to be a species can be chosen according to the intended abstraction level. Consider a molecule that may occur in different states, as e.g. free or bound. Classical approaches (e.g. ODE and stochastic) model this situation with two species whereas in our approach we have the choice to consider it as one or two components depending on the present view of the system.

The formalism allows for a intuitive graphical representation.

We apply our component-based modelling to various examples and compare it with other techniques. Results that speak in favour of our approach are

- reduction of the state space by providing a population-based concept of a system component,
- compositionality of the formalism, in particular, automatic calculation of the global system's stochastic transition rates based on the weights given by the individual components,
- adding (and removing) of further reactions and substances can be easily achieved,
- free choice of abstraction level,
- any degree of cooperation, i.e. how many different substances are involved in one chemical process, can be modelled straightforwardly using the notion of input and output ports,
- intuitive graphical representation.

The remainder of this paper is organised as follows. Section II provides the theoretical background of the paper and reviews former modelling approaches. In section III we describe our formalism and discuss analysis methods. A comparison to approaches based on stochastic process algebra is presented in Section IV. Finally, Section V concludes the paper and gives directions of further research. Due to length restrictions, an extensive interaction model for parts of the genetic regulatory network of bacteriophage lambda can be found in the appendix.

II. Non-hierarchical Quantitative Models

In this paper, the systems under study are interacting biological objects, more precisely, biochemical reactions between molecules in the living cell.

We consider a fixed volume with constant environment conditions that is a well-stirred mixture and we assume that molecules collide randomly. Thus, reactions between large populations of input substances (*reactants*) occur more frequently than those between smaller populations.

There are various ways to denote complex networks of cellular processes. A well-adopted biological model is that of *stoichiometries* relating reactants and *products* of a reaction quantitatively. A stoichiometric equation describes how many molecules of the input substances are used up if a reaction of this type happens and how many molecules are produced. We will use this model as the starting point throughout the paper.

Example 1: Let E represent one enzyme molecule that may bind to one protein molecule of type S to form a complex molecule of type C. The complex C may either dissociate to yield E and S again or E and a transformed version of S, denoted by P. The system can be described by the equation

$$E + S \rightleftharpoons C \rightarrow E + P.$$

Since the volume contains whole populations of molecules of type E, S, C and P, respectively, these reactions may happen in many instances. We call this system the *enzyme-catalysed substrate conversion* and we will use it henceforth as a running example.

The species' populations of the example given above are treated as separate entities and species C is treated in the same way as E and S although it represents a complex molecule. Similarly, the direct connection between S and P is not given (molecule P is a transformed version of S). This is why we call this view *non-hierarchical*.

Given a a set $\{R_i \mid 1 \leq i \leq I\}$ of biochemical reactions involving J different species $S_1, S_2, \ldots, S_J$ the general form of the stoichiometry for reaction R_i is given by the equation

$$\begin{aligned} s_{i_1}S_{i_1} + s_{i_2}S_{i_2} + \ldots + s_{i_m}S_{i_m} &\xrightarrow{c_i} \\ s_{i_{m+1}}S_{i_{m+1}} + s_{i_{m+2}}S_{i_{m+2}} &+ \ldots + s_{i_l}S_{i_l} \end{aligned} \tag{1}$$

where $0 \leq m \leq l \leq 2J$, $m \leq J$ and $s_{i_1}, \ldots, s_{i_l} \in \mathbb{N}$ are stoichiometric coefficients. The left-hand substrates are the reactants and the substrates on the right-hand are the products of R. Equation (1) describes how the reaction affects the population vector, i.e. for each $h \in \{1, \ldots, m\}$ the number of molecules of chemical species S_{i_h} that are used up is s_{i_h} and for $h \in \{m+1, \ldots, l\}$ the number of molecules of chemical species S_{i_h} that are produced by the reaction is s_{i_h}. Following [10] the *stochastic reaction rate constant* (or *rate constant* for short) $c_i \in \mathbb{R}_{>0}$ determines the speed of the reaction, as explained in Section II-A.

In most cases, the number of reactants and products is small, i.e. $m \leq 2$ and $l - m \leq 2$ and also the stoichiometric coefficients are mostly equal to one. All other reaction types are rare because the probability that three or more independent molecules collide at the same time or within a small time interval is very small.

Example 2: Recall the enzyme-catalysed substrate conversion of Example 1. Here, the number of participating species is $J = 4$ and the number of reactions is $I = 3$. All stoichiometric coefficients are equal to one. We let $c_1, c_2, c_3 \in \mathbb{R}_{>0}$ and define

$$R_1 : \ E+S \xrightarrow{c_1} C, \ R_2 : \ C \xrightarrow{c_2} E+S, \ R_3 : \ C \xrightarrow{c_3} E+P.$$

The information represented in a set of stoichiometries can be used to carry out a quantitative analysis in two ways: either by establishing a set of differential equations that allow for a continuous approximation of expected concentrations of chemical substances or by taking the stochastic approach that yields a more detailed picture of the system behaviour including information on variances and higher moments of measures of interest.

Both approaches operate on a low level of abstraction and follow the non-hierarchical view as they consider each chemical species as a separate entity. On this level the fact that some species are the result of the reaction of some other species is encoded only implicitly.

A. Stochastic Model

Focusing on the stochastic approach [1], [21], [11] each population is represented as a random variable $X_j(t)$ describing the number of molecules of type j at time instant t. Random vector

$$X(t) = \big(X_{S_1}(t), X_{S_2}(t), \ldots, X_{S_J}(t)\big)$$

represents the state of the whole system at time instant $t \geq 0$. Since the future evolution of the stochastic process $\big(X(t)\big)_{t \geq 0}$ depends only on the current state (and not on the process history or the current time instant) it is a continuous-time Markov chain (or CTMC for short). A CTMC can be viewed as a graph in which each node represents a state $x = (x_{S_1}, x_{S_2}, \ldots, x_{S_J}) \in \mathbb{N}^J$ that is a possible realisation of random vector $X(t)$. An edge (x, y) constitutes a (discrete) state change from x to y triggered by some chemical reaction. Edges are labelled by (stochastic transition) rates meaning that for a sufficient small time interval $[t, t+dt]$ the probability of being in state y at time $t + dt$ while starting in x at time t is given by

$$P\big(X(t+dt) = y \mid X(t) = x\big) = r \cdot dt.$$

The relationship between rate constants of reactions and transition rates of the underlying CTMC is described below. The behaviour of CTMC $\big(X(t)\big)_{t \geq 0}$ is completely described by its transition rates and its initial state.

Rate Calculation and Target States. As already stated, the probability of a reaction depends on the reactants' populations. Since each transition of state $x = (x_{S_1}, x_{S_2}, \ldots, x_{S_J})$ corresponds to the event of a reaction, the associated transition rate is a function of the population vector x. As well-justified in [11], these rates grows linear in the number of reactant molecules $x_{S_{i_1}}, \ldots, x_{S_{i_m}}$, i.e. for reaction R_i (compare Equation (1)) we have transition rate

$$r_i(x) = \begin{cases} c_i & \text{if } m = 0, \\ c_i \cdot \displaystyle\prod_{j=1}^{m} \binom{x_{S_{i_j}}}{s_{i_j}} & \text{if } \ \forall j \in \{1, \ldots, m\} : \\ & \quad\ x_{S_{i_j}} \geq s_{i_j}, \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

State x has an outgoing R_i transition with label $r_i(x)$ if each reactant population $x_{S_{i_j}}, 1 \leq j \leq m$, is not lower than the number s_{i_j} of molecules of type S_{i_j}. Otherwise $r_i(x)$ is zero which means that there is no transition that corresponds to reaction type R_i. Equation (2) takes into account the number of distinct combinations of R_i reactant molecules.

A state x has at most l outgoing transitions leading to different target states where l equals the number of participating substrates (compare Equation (1)). Consider reaction R_i and assume that $r_i(x) > 0$. State y represents the system after a R_i reaction has happened, i.e. $y = x + v_i$ where v_i is the vector of change with entries $v_i(j) = -s_{i_j}$ if species S_{i_j} is a reactant and $v_i(j) = s_{i_j}$ if S_{i_j} is a product[1]. We indicate such a transition by $x \xrightarrow{r_i(x)} y$.

Example 3: The Markov chain associated with the reaction network of Example 2 is defined by

$$X(t) = \big(X_E(t), X_S(t), X_C(t), X_P(t)\big)$$

and every state $(x_E, x_S, x_C, x_P) \in \mathbb{N}^4$ with $x_C + x_E = x_E^{(0)}$ and $x_S + x_C + x_P = x_S^{(0)}$ is reachable if we assume that the initial state is $x = (x_E^{(0)}, x_S^{(0)}, 0, 0)$, i.e. initially there are no complex molecules and molecules of type P in the system and the number of type E and type S molecules is $x_E^{(0)}$ and $x_S^{(0)}$, respectively. If $X(t) = (x_E, x_S, x_C, x_P)$ the following transitions can be taken:

If $x_E, x_S > 0$,
$(x_E, x_S, x_C, x_P) \xrightarrow{c_1 x_E x_S} (x_E - 1, x_S - 1, x_C + 1, x_P)$.
If $x_C > 0$,
$(x_E, x_S, x_C, x_P) \xrightarrow{c_2 x_C} (x_E + 1, x_S + 1, x_C - 1, x_P)$.
If $x_C > 0$,
$(x_E, x_S, x_C, x_P) \xrightarrow{c_3 x_C} (x_E + 1, x_S, x_C - 1, x_P + 1)$.

In Section III we will use Markov chains as underlying structure for our proposed modelling approach and generate the CTMC that describes the behaviour of a given interaction system.

B. ODE Model

The most widespread formalism used to model intracellular dynamics is that of *rate equations*, also known as ODE approach. Rate equations are based on the generalised mass action law and consist of J ordinary differential equations where J is the number of substances participating in the reaction network. Concentrations of molecules are related according to their possible chemical reactions. The rate of production of molecular species S_j is expressed as a function of the concentrations of other species.

Example 4: Recall Example 1 and suppose that the (non-stochastic) reaction rate constants $k_1, k_2, k_3 \in \mathbb{R}_{>0}$ of the three reactions are given[2]. The ODE approach presupposes that the species' concentrations vary continuously and deterministically in time. If $[S]$ denotes the concentration of species S (in units mol per litre) the corresponding set of ODEs is as follows:

$$\begin{aligned}
\tfrac{d}{dt}[E] &= -k_1[E][S] + (k_2 + k_3)[C]\\
\tfrac{d}{dt}[S] &= -k_1[E][S] + k_2[C]\\
\tfrac{d}{dt}[C] &= k_1[E][S] - (k_2 + k_3)[C]\\
\tfrac{d}{dt}[P] &= k_3[C]
\end{aligned}$$

We do not discuss the ODE model more detailed because the rest of the paper focuses on the stochastic approach.

The structural properties of the biological system are hidden within such low level models. Moreover, once such a model is constructed it is not easy to add further components in a compositional manner.

We argue that there is a need for more powerful modelling techniques operating on a higher abstraction level to facilitate hierarchical composition of sub-models and that are closer to customary graphical notations of processes used by biologists.

III. Hierarchical Modelling

We have seen that traditional mathematical models for reaction networks such as Markov processes or ordinary differential equations operate on the number of molecules or the concentration of the involved chemical species. For instance, in case of the enzyme-catalysed substrate conversion the corresponding Markov process consists of four random variables, one for each type of molecule (compare Example 3). The fact that the complex molecule C is the formation of a protein and an enzyme molecule is represented only in an implicit way. The type C molecules form a separate population. From the biological viewpoint a molecule of type C consists of two parts, whereas an (unbound) enzyme molecule is an entity in its own right. Usually, in graphical representations complex molecules are not drawn as an extra object. Instead, the two parts forming the complex are connected in some way (e.g. by an arrow). Furthermore, in stoichiometries involving complex formation multi-character symbols are chosen for complex molecules where each letter describes a certain part of the complex molecule[3] A similar problem arises with the product molecules (type P) that are "transformed" S molecules. From the biological point of view, a protein molecule can be in different chemical states: either it behaves like a molecule of type S or certain thermodynamic and/or structural properties are changed by the enzyme-catalysed transformation such that the molecule is then of type P.

In case of a hierarchical view we model the enzyme-catalysed substrate conversion by using only two components. One component corresponds to the substrate/product population, denoted by S, and one the enzyme population, denoted by E. Each individual can be in different states, i.e. enzyme molecules (component E) can be either free (type E) or bound (type $E{\cdot}S$). Elements of S can be free (type S), bound as a part of a complex molecule (type $E{\cdot}S$) or transformed (type P).

[1] It might happen the rare case that there is a species S_h occurring on both sides of reaction R_i, i.e. there exists i_j, i_k such that $1 \le i_j \le m < i_k \le l$ and $i_j = i_k = h$. Then $v_i(h) = -s_{i_j} + s_{i_k}$.

[2] In general, the constants k_1, k_2, k_3 differ from the stochastic rate constants c_1, c_2, c_3 but can be calculated from those.

[3] E.g. in case of Example 1 we write $E + S \rightleftharpoons E{\cdot}S \rightarrow E + P$.

In the following, we define a hierarchical modelling framework based on *interaction systems* and give semantics in terms of a stochastic process. The starting point is a component system defining the components and their properties.

Definition 1: A *component system* CS is a pair $(K, \{A_n\}_{n\in K})$ with

- a finite component set K,
- pairs $A_n = (I_n, O_n)$ of disjoint *port sets* for each component $n \in K$ such that for all ports $a_n \in (I_n \cup O_n)$ it holds that either a_n is an *input port* ($a_n \in I_n$) or an *output port* ($a_n \in O_n$) and a_n can not be a port of some other component $m \in K, m \neq n$.

Informally, at a low abstraction level set K contains the "basic molecule types" of the reaction network and the associated ports of component n describe information about the molecules that are needed to model reactions. Set I_n represents possible actions component n is able to perform (e.g. "bind", "dissociate") and O_n is the set of possible results (e.g. "complex", "free", "transformed").

Example 5: We may choose $K = \{\mathsf{E}, \mathsf{S}\}$ for the components modelling the enzyme-catalysed substrate conversion (see Example 2) and associate suitable pairs of port sets A_E and A_S as follows:

$$A_\mathsf{E} = (I_\mathsf{E}, O_\mathsf{E}) = (\{\mathsf{bind}_\mathsf{E}, \mathsf{diss}_\mathsf{E}\}, \{\mathsf{complex}_\mathsf{E}, \mathsf{free}_\mathsf{E}\})$$

$$A_\mathsf{S} = (I_\mathsf{S}, O_\mathsf{S}) = (\{\mathsf{bind}_\mathsf{S}, \mathsf{diss}_\mathsf{S}\}, \{\mathsf{complex}_\mathsf{S}, \mathsf{free}_\mathsf{S}, \mathsf{transf}_\mathsf{S}\})$$

Let $\mathcal{P}$ denote the power set operator. Interactions between the different components are defined by relating ports of components.

Definition 2: Let $CS = (K, \{A_n\}_{n\in K})$ be a component system. An *interaction* of CS is a triple $R = (c_R, I_R, O_R)$ such that $c_R \in \mathbb{R}_{>0}$ is a constant, $I_R \cup O_R \neq \emptyset$, $I_R \subseteq \mathcal{P}\big(\bigcup_{n\in K} I_n\big)$ and $O_R \subseteq \mathcal{P}\big(\bigcup_{n\in K} O_n\big)$. We may use subscript R to identify the port sets and the constant of R.

An *interaction set* is a set $\mathcal{I}$ of interactions of CS that covers all ports, i.e.

$$\bigcup_{R\in\mathcal{I}, A\in I_R} A = \bigcup_{n\in K} I_n \quad \text{and} \quad \bigcup_{R\in\mathcal{I}, B\in O_R} B = \bigcup_{n\in K} O_n.$$

Here, interactions have a one-to-one correspondence to chemical reactions. Constant c_R is the stochastic rate constant (compare Section II) of reaction R, the sets of ports that are elements of I_R correspond to R's reactants and those of O_R to the products of reaction R.

Example 6: Recall the running example of the paper (Example 2). We put

$$R_1 = (c_1, \{\{\mathsf{bind}_\mathsf{E}\}, \{\mathsf{bind}_\mathsf{S}\}\}, \{\{\mathsf{complex}_\mathsf{E}, \mathsf{complex}_\mathsf{S}\}\}).$$

We shortly write

$$R_1 : \{\mathsf{bind}_\mathsf{E}\} + \{\mathsf{bind}_\mathsf{S}\} \xrightarrow{c_1} \{\mathsf{complex}_\mathsf{E}, \mathsf{complex}_\mathsf{S}\}.$$

Similarly,

$$R_2 : \quad \{\mathsf{diss}_\mathsf{E}, \mathsf{diss}_\mathsf{S}\} \xrightarrow{c_2} \{\mathsf{free}_\mathsf{E}\} + \{\mathsf{free}_\mathsf{S}\},$$

$$R_2 : \quad \{\mathsf{diss}_\mathsf{E}, \mathsf{diss}_\mathsf{S}\} \xrightarrow{c_3} \{\mathsf{free}_\mathsf{E}\} + \{\mathsf{transf}_\mathsf{S}\}.$$

For each $n \in K$ let $\pi^I(n, R)$ and $\pi^O(n, R)$ be the set of input ports and output ports, respectively, involved in reaction R and offered by component n, i.e. $a_n \in \pi^I(n, R)$ iff $a_n \in I_n$ and $a_n \in (\bigcup_{A\in I_R} A)$ and $a_n \in \pi^O(n, R)$ iff $a_n \in O_n$ and $a_n \in (\bigcup_{B\in O_R} B)$.

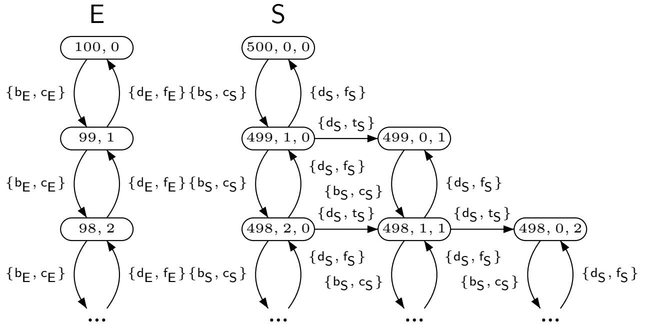

Fig. 1. The local transition systems of the two components that represent the enzyme-catalysed substrate conversion.

Definition 3: Let $CS = (K, \{A_n\}_{n\in K})$ be a component system. A *local transition system* of component $n \in K$ is a tuple $(Q_n, A_n, \rightarrow_n, w_n, s_n)$ with

- a set Q_n of local states, $Q_n \cap Q_m = \emptyset$ for $m \neq n, m \in K$,
- a weight function $w_n : Q_n \times I_n \rightarrow \mathbb{R}_{\geq 0}$,
- an initial state s_n and
- a transition relation $\rightarrow_n \subseteq Q_n \times \mathcal{P}(I_n \cup O_n) \times Q_n$.

Let A be a set of ports. We write $q_n \xrightarrow{A}_n q'_n$ iff $(q_n, A, q'_n) \in \rightarrow_n$ and for $a_n \in I_n$ we call $w_n(q_n, a_n)$ the weight at port a_n of component n in state q_n.

Local transition systems evolve via a set of ports through their state space, meaning that a chemical reaction may operate on several ports of component $n \in K$. Weights provide information for interactions in the global transition system. In our context, the weight equals the reactants' population port a_n is associated with i.e. if q_n is the current state then $w_n(q_n, a_n) \in \mathbb{N}$ equals the current number of type a_n molecules.

Example 7: The local transition systems of components E and S (compare Example 4) are party illustrated in Figure 1. Component E has 100 different states if there are initially 100 enzyme molecules and no complex molecules in the system. States are labelled by pairs of natural numbers such that the first entry gives the number of free enzyme molecules and the second gives the number of enzyme molecules that are bound as a part of a complex. Similarly, the local states of component S are labelled by triples such that the first entry equals the current number of free S molecules, the second entry the number of those that are bound and the third entry represents the number of transformed S molecules, i.e. type P molecules. We assume that initially there are 500 free molecules of type S and no transformed or bound molecules. Port names are abbreviated in Figure 1 as follows:

bind_E: b_E	diss_E: d_E	$\mathsf{complex}_\mathsf{E}$: c_E	free_E: f_E
bind_S: b_S	diss_S: d_S	$\mathsf{complex}_\mathsf{S}$: c_S	free_S: f_S
$\mathsf{transf}_\mathsf{S}$: t_S			

Assume that the current state of E is given by $q_\mathsf{E} = (x_E, x_C)$. Weight function w_E is such that $w_\mathsf{E}((x_E, x_C), \mathsf{bind}_\mathsf{E}) = x_E$ meaning that input port bind_E's associated weight equals the current number of free enzyme molecules whereas input port diss_E has weight x_C which equals the current number of complex molecules. Similarly, the input ports of S have weights $w_\mathsf{S}((x_S, x_C, x_P), \mathsf{bind}_\mathsf{S}) = x_S$ and $w_\mathsf{S}((x_S, x_C, x_P), \mathsf{diss}_\mathsf{S}) = x_C$ if component S is in state $q_\mathsf{S} = (x_S, x_C, x_P)$.

Recall that output ports do not provide weights. They are only used to distinguish between several outgoing transitions. Furthermore, if a specie's population occurs in more than one component, e.g. complex molecules, the input ports' weights have to be equal to let the corresponding reaction happen as it is the case for port diss_S and diss_E.

Wlog, we assume that $K = \{1, 2, \ldots, k\}$. The behaviour of the composite system (*global system*) is described by the derived interaction system.

Definition 4: Let $CS = (K, \{A_n\}_{n \in K})$ be a component system and $\mathcal{I}$ an interaction set of CS. An *interaction system* is a triple $(CS, \mathcal{I}, \mathcal{T})$ with a *global transition system* $\mathcal{T} = (\mathcal{Q}, \mathcal{I}, \rightarrow, s)$ such that for each $n \in K$ there exists a local transition system $T_n = (Q_n, A_n, \rightarrow_n, s_n)$ and

- $\mathcal{Q} = Q_1 \times Q_2 \times \ldots \times Q_k$ where elements of $\mathcal{Q}$ are denoted by $(q_1, q_2, \ldots, q_k)$ and called *global states*.
- State $s = (s_1, s_2, \ldots, s_k)$ is the initial state.
- The global transition relation $\rightarrow \subseteq \mathcal{Q} \times \mathcal{I} \times \mathbb{R}_{\geq 0} \times \mathcal{Q}$ is such that there is a global transition
$$q = (q_1, q_2, \ldots, q_k) \xrightarrow{R,w} q' = (q_1', q_2', \ldots, q_k')$$
labelled by reaction $R = (c_R, I_R, O_R)$ and rate $w > 0$ if and only if the following two conditions hold:
    1) For all $n \in K$, if $\pi^I(n, R) \cup \pi^O(n, R) = C \neq \emptyset$ then $q_n \xrightarrow{C}_n q_n'$ and $q_n = q_n'$ otherwise.
    2) For all input port sets $A \in I_R$ there exists a weight $w_A > 0$ such that
        i) for all $\forall n \in K$:
        $$a \in A \cap \pi^I(n, R) \text{ implies } w_n(q_n, a) = w_A,$$
        ii)
        $$w = \begin{cases} c_R \cdot \prod_{A \in I_R} w_A & \text{if } I_R \neq \emptyset, \\ c_R & \text{otherwise.} \end{cases} \qquad (3)$$

The two conditions of the global transition relation can be informally explained as follows:

Condition 1 states that the local state q_n of component n can change triggered by reaction R if and only if there is a non-empty port set of n that participates in R. Moreover, the port set of n involved in R equals the label of the transition leading from q_n to successor q_n'.

The second condition calculates the weight associated with the global transition. Each port set $A \in I_R$ has a unique weight w_A and all ports $a \in A$ describe the same population which might occur in more than one component, e.g. in case of complex formation. Thus, w_A is the current weight of all $a \in A$. The product of these local weights and constant c equals the stochastic transition rate w of reaction R according to the stochastic semantics (compare Section II). If $I_R = \emptyset$ we have $O_R \neq \emptyset$ and the production rate of the substances $B \in O_R$ is constant.

Modelling biochemical reaction networks in the framework of interaction systems gives rise to a variety of benefits. For a given interaction system IS and local transition systems T_n it is easy to add further components and interactions to the system. Moreover, it is also possible to switch to a higher abstraction level by applying *hide* or *join* operations to abstract from ports or to join several components[4].

Appendix A presents a more complex example of an interaction system modelling a genetic reaction network that is part of the λ phage regulatory system.

A. Analysis of Interaction Systems

We are interested in the temporal behaviour amongst large numbers of molecules to understand the functional activity of the network. In this section we give details on how interaction systems can be used to calculate transient and long-run measures of the biological system under study.

Interaction systems evolve from one global state to another by a set of transitions that are related to interactions i.e. to the chemical reactions. The associated weights of these global transitions equal the stochastic transition rates of the underlying Markov Process as defined in Equation (2), Section II, because they are the product of the local weights and the stochastic reaction rate constant. The weight function w_n of the local transition systems has to be defined appropriately, such that the local weights equal the factors a component contributes to the global rate. The global transition system can be mapped onto a continuous-time Markov chain having the same state space and the same transitions but no reaction labels, i.e. transition $q \xrightarrow{R,w} q'$, $w > 0$ corresponds to transition $q \xrightarrow{w} q'$ in the underlying CTMC and w is the stochastic transition rate[5].

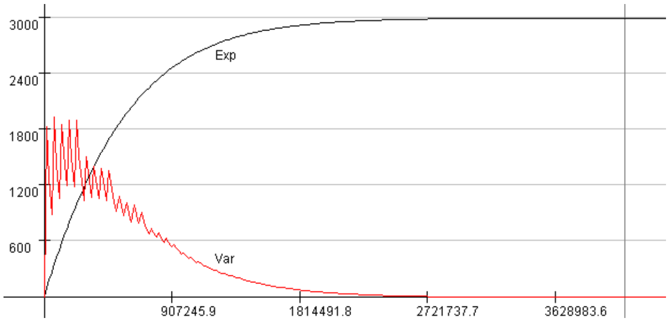

Fig. 2. Simulation results of the enzyme-catalysed substrate conversion.

a) Stochastic Simulation: Discrete-event simulation of interaction systems can be accomplished on the basis of the components' local transition systems T_n and the interaction set $\mathcal{I}$. In each simulation step we check for all reactions $R \in \mathcal{I}$ and all components $n \in K$ if T_n fulfils the conditions

[4] Due to space limitations we omit further details about the *hide* and *join* operation.

[5] In general, multiple transitions might occur. However, in our setting we assume that two reactions of different types always lead to different target states in the global transition system. This ensures that there is at most one transition between two global states.

of Definition (4). We calculate R's transition rate $w = w_R$ according to Equation (3) and summing up all w_R yields the parameter needed to compute the length of the next time step. In analogy to Gillespie's direct method [11] the reactions' transition probabilities are calculated and once the next time step and the next reaction is chosen (by generating two random numbers) each participating local transition system is updated by performing the corresponding local transitions.

We implemented a stochastic simulator for interaction systems. Figure 2 illustrates some results of the simulation of the enzyme-catalysed substrate conversion (compare Example 7). We averaged over 5000 simulation runs and chose parameters

$$c_1 = 10^{-4}, c_2 = 1, c_3 = 10^{-4}, x_E^{(0)} = 220 \text{ and } x_E^{(0)} = 3000.$$

The curve labelled by 'Exp' shows the mean of the number of type P molecules at time t, i.e. we observed the third entry of the local state of component S. The curve with label 'Var' gives the sample variance of the P population at time t. Clearly, at the beginning the variance is high because only a small number of transformed molecules are in the system. The expected time until all S molecules have been transformed is indicated by the vertical line at time $t = 4,123,845$.

b) Numerical Analysis: As already stated, numerical analysis is often not feasible because of the inherent complexity of realistic biological examples. The state space of the global transition system grows exponentially in the number of components which leads to an enormous space complexity. One idea to manage complexity and largeness of state spaces is the exploitation of model and system structure in the analysis of the underlying CTMC.

A lot of structural analysis approaches from the field of performance analysis of computer systems exist, e.g. on the basis of stochastic automata networks and implemented in tools like PEPS [2], APNN [5] and SMART [8]. It is important to notice that our component-based modelling framework can be seen as a generalisation of the stochastic automata network approach in [22]. There, a non-hierarchical, but modular representation of biochemical reaction networks in terms of a tensor algebra is given which also conserves the regular biological structure. As our formalism is also structure preserving, it is predestined to be the starting point of a structural analysis.

Moreover, systems may suffer from different time scales, meaning that some reactions occur extremely rarely compared to others. Thus, advanced analysis techniques have to be applied such as aggregation methods [6]. Since these methods mostly rely on a partitioning of the global state space, the component-wise description gives useful hints for an appropriate segmentation of the state space into aggregates or even induces a partitioning leading to accurate results directly.

IV. Comparison between Interaction Systems and Stochastic Process Algebra

This section provides a short discussion about the similarities and differences between our component-based modelling framework and stochastic process calculi which have been applied to systems in biology recently.

The two most popular process algebras used to construct probabilistic models in the biological context are the stochastic π-calculus [18], [17] and PEPA [14], [7]. Due to lack of space our comparison concentrates on the π-calculus but also gives short remarks about PEPA.

Our modelling approach focuses on the temporal evolution of populations of certain molecule types. Expectations and variances of these at specific time points are calculated to deduce functional relationships. By contrast, stochastic π-calculus models are constructed from the perspective that considers each single molecule individually by representing it as a process term instead of representing a whole population as one term. Since the stochastic semantics of these process terms leads to Markov models with extremely large state spaces, using stochastic process algebra for cellular dynamics seems to be only meaningful if one wants to distinguish molecules of the same type.

Example 8: We present a stochastic π-calculus model of the enzyme-catalysed substrate conversion (see Example 2) in analogy to [3]. One enzyme molecule is represented as

$$E() \quad = \quad \nu d \; \nu t \; !c(d,t).\big(!d.E() + !t.E()\big)$$

and one substrate molecule as

$$S() \quad = \quad ?c(d,t).\big(?d.S() + ?t.P()\big).$$

Here, operator ν creates a new channel, $!a$ and $?a$ mean output and input on channel a, respectively. Channel d models reaction R_2 (decomplexation), channel t reaction R_3 (transformation) and channel c reaction R_1 (complexation). If we, for instance, create three instances of type E and five of type S, we let

$$Q() \quad = \quad E() \mid E() \mid E() \mid S() \mid S() \mid S() \mid S() \mid S().$$

Process Q keeps track of which subprocess $E()$ interacts with which of the five $S()$ subprocesses. Thus, the model's state space is of size $\mathcal{O}(2^{i_E} \cdot 3^{i_S})$ if i_E and i_S are the initial numbers of type E and type S molecules which can be in two and three different states, respectively. In general, the state space grows exponentially in the maximum population size. In the interaction model we have, by contrast, $\mathcal{O}({i_E}^2 \cdot {i_S}^3)$. Here, the size of the underlying Markov chain grows exponentially in the number of different states a molecule can be in, which is in nearly all cases much smaller than the maximum population size.

Clearly, numerical analysis of large populations is only possible if the state space size of the underlying model is in a tractable range. Moreover, also for simulative techniques the population-based view is of benefit if one is interested in the temporal evolution of populations (and not of individual molecules). On the other hand, for systems involving only small populations (e.g. signalling pathways) both, our approach and the stochastic process algebra approach are adequate.

Further differences concern the manner interactions are described. In stochastic π-calculus one distinguishes active and passive participants of interactions (according to $!a$ and

$?a$). In addition, it is a two-way communication[6] which is in contrast to the multiway communication structure of interaction systems.

V. Conclusion

We have proposed a component-based modelling framework which has been shown to be suitable for the representation and analysis of biochemical reaction networks. Our formalism is situated between the rather informal descriptions used by biologists and mathematical tools for simulative and numerical analysis.

The enzyme-catalysed substrate conversion and a part of the λ phage system have served as examples in the paper.

We have highlighted the advantages of our modelling approach and argued that interaction systems allow for abstraction leading to significant state space reductions compared to approaches based on stochastic process algebra.

By maintaining the biological structure of the system under study enhanced analysis techniques can be applied to interaction models.

As future research we plan to present a more detailed component-based model of the λ phage life cycle including further regulatory proteins.

References

[1] A. Arkin, J. Ross, and H. H. McAdams. Stochastic kinetic analysis of developmental pathway bifurcation in phage lambda-infected escherichia coli cells. *Genetics*, 149:1633–1648, 1998.

[2] A. Benoit, P. Fernandes, B. Plateau, and W. Stewart. The PEPS Software Tool. In *13th International Conference on Modelling Techniques and Tools for Computer Performance Evaluation TOOLS 2003*, pages 98–115, Urbana, Illinois, USA, 2003.

[3] R. Blossey, L. Cardelli, and A. Phillips. A compositional approach to the stochastic dynamics of gene networks. *Transactions in Computational Systems Biology*, 3939:99–122, 2006.

[4] J. M. Bower and H. Bolouri. *Computational Modeling of Genetic and Biochemical Networks*. The MIT Press, 2001.

[5] P. Buchholz and P. Kemper. A toolbox for the analysis of discrete event dynamic systems. In *Proceedings of the 11th International Conference on Computer Aided Verification*, pages 483–486, London, UK, 1999. Springer.

[6] H. Busch, W. Sandmann, and V. Wolf. A numerical aggregation algorithm for the enzyme-catalyzed substrate conversion. In *Computational Methods in Systems Biology (CMSB)*, volume LNCS 4210, pages 298–311. Springer Verlag, 2006.

[7] M. Calder, S. Gilmore, and J. Hillston. Modelling the influence of RKIP on the ERK signalling pathway using the stochastic process algebra PEPA. In *Transactions on Computational Systems Biology VII*, number 4230 in LNCS. Springer, 1–23 2006.

[8] G. Ciardo and A. Miner. SMART: The stochastic model checking analyzer for reliability and timing. In *Proceedings of the 1st International Conference on Quantitative Evaluation of Systems*, pages 338–339, 2004.

[9] M. A. Gibson and J. Bruck. A probabilistic model of a prokaryotic gene and its regulation. In *In Bower and Bolouri[4]*, chapter 2, pages 49–71.

[10] D. T. Gillespie. A general method for numerically simulating the time evolution of coupled chemical reactions. *Journal of Computational Physics*, 22:403–434, 1976.

[11] D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *Journal of Physical Chemistry*, 81(25):2340–2361, 1977.

[12] P. Goss and J. Peccoud. Quantitative modeling of stochastic systems in molecular biology by using stochastic petri nets. In *Proceedings of the National Academy of Science USA*, pages 6750–6755, 1998.

[13] J. Hillston. The nature of synchronisation. In U. Herzog and M. Rettelbach, editors, *Proceedings of the Second International Workshop on Process Algebras and Performance Modelling*, pages 51–70, Erlangen, November 1994.

[14] J. Hillston. A compositional approach to performance modelling, 1996.

[15] C. Kuttler and J. Niehren. Gene regulation in the pi calculus: simulating cooperativity at the lambda switch. *Transactions on Computational Systems Biology VII*, 4230:24–55, 2006.

[16] M. Kwiatkowska, G. Norman, D. Parker, O. Tymchyshyn, J. Heath, and E. Gaffney. Simulation and verification for computational modelling of signalling pathways. In *Proceedings of the Winter Simulation Conference*, 2006. To appear.

[17] A. Phillips and L. Cardelli. A correct abstract machine for the stochastic pi-calculus. In *Bioconcur'04*. ENTCS, August 2004.

[18] C. Priami. Stochastic π-calculus. *Comput. J.*, 38(7):578–589, 1995.

[19] C. Priami, A. Regev, E. Shapiro, and W. Silverman. Application of a stochastic name-passing calculus to representation and simulation of molecular processes. *Inf. Process. Lett.*, 80(1):25–31, 2001.

[20] T. Turner, S. Schnell, and K. Burrage. Stochastic approaches for modelling in vivo reactions. *Computational biology and chemistry*, 28(3):165–178, 2004.

[21] N. G. van Kampen. *Stochastic Processes in Physics and Chemistry*. North-Holland, Amsterdam, 1992.

[22] V. Wolf. Modelling of biochemical reactions by stochastic automata networks. In *Proceedings of MecBic'06*, ENTCS, 2006. To appear.

[6] In PEPA multi-way cooperation is possible but the corresponding rate semantics is not appropriate in case of molecular interactions; the overall synchronisation rate stays constant in the number of synchronisation participants [13].

APPENDIX

In this section, we construct an interaction system for parts of the genetic network in bacteriophage λ. This phage infects Escherichia coli (E. coli) cells and is one of the best-understood organisms for that probabilistic models may be strictly required. Infected host cells either end up in lysis which means the dissolution of the E. coli cell and the release of new phages or they end up in lysogeny which means that the viral DNA is integrated into the host's DNA for passive replication. The decision between lysis and lysogeny is probabilistic and several probabilistic models for parts of this system have been proposed recently [1], [9], [15].

We focus on the early life cycle of phage λ and a specific gene, namely the N gene, which is a temporal regulator having great impact on the decision between the two different cell fates.

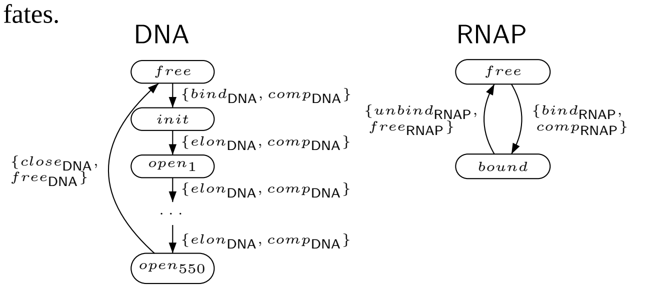

Fig. 3. The local transition systems of component DNA and RNAP.

a) Transcription: We start with component DNA which represents gene N and assume a constant concentration of transcription factors as it is the case early in the life cycle. The population of RNA polymerase (RNAP) molecules is given by component RNAP (see Figure 3). Here, DNA and RNAP are sufficient for transcription into messenger RNA (mRNA) (represented by component mRNA; see Figure 4). More precisely, RNAP and DNA form a complex to initiate transcription and the double-stranded DNA is unwinded in order for RNA polymerase to access free unravelled DNA strands. Thus, component DNA is equipped with input port $bind_{\mathsf{DNA}}$ and output port $comp_{\mathsf{DNA}}$. Similarly, component RNAP is able to perform $bind_{\mathsf{RNAP}}$ and $comp_{\mathsf{RNAP}}$. The corresponding interaction is given by

$$Init_{tcr} : \{bind_{\mathsf{DNA}}\}+\{bind_{\mathsf{RNAP}}\} \xrightarrow{c_1} \{comp_{\mathsf{DNA}}, comp_{\mathsf{RNAP}}\}.$$

Recall that this means that $Init_{tcr}$ equals

$$(c_1, \{\{bind_{\mathsf{DNA}}\}, \{bind_{\mathsf{RNAP}}\}\}, \{\{comp_{\mathsf{DNA}}, comp_{\mathsf{RNAP}}\}\}).$$

The transcription process proceeds with elongation which means the DNA is scanned nucleotide by nucleotide (input port $elon_{\mathsf{DNA}}$). We assume the number of transcribed nucleotides is 550 (compare [1]) and the duration of each single scan step has the same stochastic rate. Thus, component DNA has $550 + 2$ different states. As depicted in Figure 3, the state labelled $free$ is the initial state of DNA in which the DNA has not (yet) bound to RNAP. State $init$ refers to the transcription initiation and $open_1, \ldots, open_{550}$ to the elongation steps. All weights of component DNA equal one because there is only one DNA molecule present in the system. We set

$$Transcr : \{elon_{\mathsf{DNA}}\} \xrightarrow{c_2} \{comp_{\mathsf{DNA}}\}.$$

Transcription is terminated after the 550 nucleotides of the DNA have been scanned (input port $close_{\mathsf{DNA}}$ and output port $free_{\mathsf{DNA}}$) e.g. DNA has reached state $open_{550}$. Then the production of a new mRNA molecule is finished. This last transcription step requires that

- the RNAP molecule separates from the DNA (input port $unbind_{\mathsf{RNAP}}$, output port $free_{\mathsf{RNAP}}$),
- the DNA strands are closed (input port $close_{\mathsf{DNA}}$, output port $free_{\mathsf{DNA}}$),
- a state change in the mRNA component happens because a new mRNA molecule is released. The associated port is $free_{\mathsf{mRNA}} \in O_{\mathsf{mRNA}}$ (compare Figure 4).

We let

$$\begin{aligned} A_{\mathsf{DNA}} &= (\{bind_{\mathsf{DNA}}, elon_{\mathsf{DNA}}, close_{\mathsf{DNA}}\} \\ &\quad\ \{comp_{\mathsf{DNA}}, free_{\mathsf{DNA}}\}), \\ A_{\mathsf{RNAP}} &= (\{bind_{\mathsf{RNAP}}, unbind_{\mathsf{RNAP}}\}, \\ &\quad\ \{comp_{\mathsf{RNAP}}, free_{\mathsf{RNAP}}\}) \end{aligned}$$

and

$$\begin{aligned} Term_{tcr} : \ \{close, unbind_{\mathsf{RNAP}}\} &\xrightarrow{c_3} \\ \{free_{\mathsf{DNA}}\} + \{free_{\mathsf{RNAP}}\} &+ \{free_{\mathsf{mRNA}}\}. \end{aligned}$$

Since in our model we assume that RNAP molecules can only bind to the promoter region of gene N and transcribe gene N, component RNAP has only two different states[7].

Ports $bind_{\mathsf{RNAP}}$ and $comp_{\mathsf{RNAP}}$ represent complex formation with the DNA molecule and $w_{\mathsf{RNAP}}(q, bind_{\mathsf{RNAP}})$ for $q = free$ equals the number of free RNAP molecules. The position of RNAP while scanning gene N is recorded by the local state of component DNA.

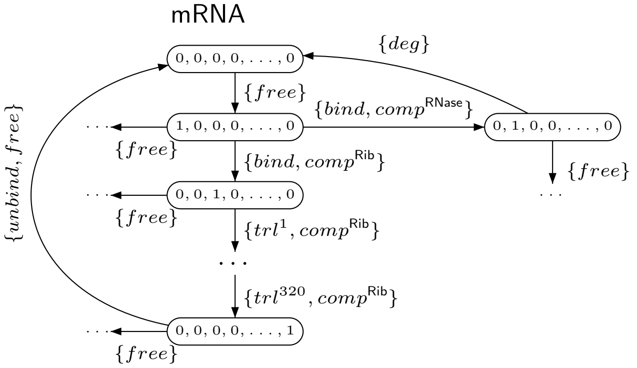

Fig. 4. Parts of component mRNA's local transition system; subscripts are omitted.

b) Translation: Similar to the transcription process translation is done in several elongation steps beginning with initiation. As opposed to transcription, we have to keep track of the current elongation step by using different input ports. This is necessary because we may have several mRNA molecules

[7] Note that if we model RNAP production RNAP would count the current number of RNAP molecules. Furthermore, if we model that RNAP can interact with other parts of the DNA, there might be further possible states and ports in the local transition system of component RNAP.

being in different states. In Figure 4 the mRNA component is illustrated (subscript mRNA of ports is omitted for simplicity). A ribosome molecule is necessary to start translation of mRNA. The ribosome population is represented by component Rib. We put

$$\begin{aligned}
A_{\mathsf{mRNA}} = \quad & (\{bind_{\mathsf{mRNA}}, deg_{\mathsf{mRNA}}, trl^{1}_{\mathsf{mRNA}}, \ldots \\
& \qquad trl^{320}_{\mathsf{mRNA}}, unbind_{\mathsf{mRNA}}\}, \\
& \{comp^{\mathsf{RNase}}_{\mathsf{mRNA}}, comp^{\mathsf{Rib}}_{\mathsf{mRNA}}, free_{\mathsf{mRNA}}\}), \\
A_{\mathsf{Rib}} = \quad & (\{bind_{\mathsf{Rib}}, unbind_{\mathsf{Rib}}\}, \{comp_{\mathsf{Rib}}, free_{\mathsf{Rib}}\}), \\
Init_{trl}: \quad & \{bind_{\mathsf{mRNA}}\} + \{bind_{\mathsf{Rib}}\} \xrightarrow{c_4} \\
& \qquad \{comp^{\mathsf{Rib}}_{\mathsf{mRNA}}, comp_{\mathsf{Rib}}\}.
\end{aligned}$$

The states of mRNA in Figure 4 are distinguished as follows. The first entry counts the number of free mRNA molecules. The second entry counts the number of mRNA molecules forming a complex with component RNase (see below). Similarly, the third entry counts the number of mRNA-ribosome complex molecules. The remaining entries record the number of mRNA molecules that are in the respective translation elongation steps. Note that output port $free_{\mathsf{mRNA}}$ always leads to a state in which the first entry is incremented by one. Figure 4 shows all those states in which there is at most one mRNA molecule. Of course, the component may reach states with several mRNA molecules.

Weights of mRNA are such that for input port $bind_{\mathsf{mRNA}}$ the current number of free mRNA molecules is chosen, for trl^{i}_{mRNA} the weight equals the current number of mRNA-ribosome complex molecules after the $(i-1)$-th elongation step, i.e. the $i+2$ entry of the state vector.

The i-th elongation step of translation is given by

$$Transl_i \quad : \quad \{trl^{i}_{\mathsf{mRNA}}\} \xrightarrow{c_5} \{comp_{\mathsf{mRNA}}\}.$$

Component Rib (see Figure 5) counts the number of ribosomes being free or being a part of a mRNA-ribosome complex molecule. The initial state is given by $s_{\mathsf{Rib}} = (x^{(0)}_{\mathsf{Rib}}, 0)$ where $x^{(0)}_{\mathsf{Rib}}$ is the initial population of ribosomes. The respective weights of Rib are such that $w_{\mathsf{Rib}}(q, bind_{\mathsf{Rib}})$ equals the number of free ribosomes of state q and $w_{\mathsf{Rib}}(q, unbind_{\mathsf{Rib}})$ the number of complex molecules.

The last translation step produces a new protein molecule of type N (output port $free_{\mathsf{Prot}}$) which means a state change in component Prot. Hence, we put

$$\begin{aligned}
Term_{trl}: \quad & \{unbind_{\mathsf{mRNA}}, unbind_{\mathsf{Rib}}\} \xrightarrow{c_8} \\
& \{free_{\mathsf{mRNA}}\} + \{free_{\mathsf{Rib}}\} + \{free_{\mathsf{Prot}}\}.
\end{aligned}$$

c) Degradation: The mRNA and the protein population will grow infinitely if we do not limit the state space of the local transition systems of these two components. Thus, a maximal size for these components has to be specified. Furthermore, we assume degradation of mRNA and protein molecules. In case of mRNA this occurs when a ribonuclease (RNase) binds to the mRNA (input port $bind_{\mathsf{RNase}}$) which is a competitive binding from the mRNA perspective because either a mRNA-RNase or a mRNA-ribosome complex is formed. Thus, input port $bind_{\mathsf{mRNA}}$ is used in both cases but the different outcomes are distinguished by output port $comp^{\mathsf{RNase}}_{\mathsf{mRNA}}$ and $comp^{\mathsf{Rib}}_{\mathsf{mRNA}}$. Note that, in general, the explicit statement of the outcome is necessary to distinguish between transitions having the same input ports and to update further components that are involved by output ports only. Of course, both bindings have weights that are determined by the number of RNase and ribosome molecules, respectively. The weight of component mRNA in case of complex formation equals the number of free mRNA molecules. Thus, $w_{\mathsf{mRNA}}(q, bind_{\mathsf{mRNA}}) = k$ where k is the first entry of the state vector $q \in Q_{\mathsf{mRNA}}$. Similarly, $w_{\mathsf{RNase}}(q, bind_{\mathsf{RNase}})$ and $w_{\mathsf{Rib}}(q, bind_{\mathsf{Rib}})$ equals the number of free RNase and ribosome molecules for $q \in Q_{\mathsf{RNase}}$ and $q \in Q_{\mathsf{Rib}}$, respectively.

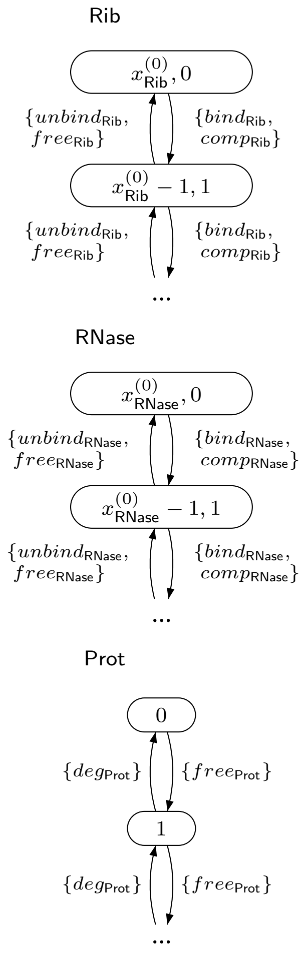
Rib
$x^{(0)}_{\mathsf{Rib}}, 0$
$\{unbind_{\mathsf{Rib}}, free_{\mathsf{Rib}}\}$
$\{bind_{\mathsf{Rib}}, comp_{\mathsf{Rib}}\}$
$x^{(0)}_{\mathsf{Rib}} - 1, 1$
$\{unbind_{\mathsf{Rib}}, free_{\mathsf{Rib}}\}$
$\{bind_{\mathsf{Rib}}, comp_{\mathsf{Rib}}\}$
...
RNase
$x^{(0)}_{\mathsf{RNase}}, 0$
$\{unbind_{\mathsf{RNase}}, free_{\mathsf{RNase}}\}$
$\{bind_{\mathsf{RNase}}, comp_{\mathsf{RNase}}\}$
$x^{(0)}_{\mathsf{RNase}} - 1, 1$
$\{unbind_{\mathsf{RNase}}, free_{\mathsf{RNase}}\}$
$\{bind_{\mathsf{RNase}}, comp_{\mathsf{RNase}}\}$
...
Prot
0
$\{deg_{\mathsf{Prot}}\}$
$\{free_{\mathsf{Prot}}\}$
1
$\{deg_{\mathsf{Prot}}\}$
$\{free_{\mathsf{Prot}}\}$
...

Fig. 5. The local transition systems of components Rib, RNase and Prot.

In case of the protein molecules we assume that the decay rate only depends on the current population size. The corresponding ports and interactions are given by

$$\begin{aligned}
A_{\mathsf{RNase}} \quad & = (\{bind_{\mathsf{RNase}}, unbind_{\mathsf{RNase}}\}, \\
& \quad \{comp_{\mathsf{RNase}}, free_{\mathsf{RNase}}\}), \\
A_{\mathsf{Prot}} \quad & = (\{deg_{\mathsf{Prot}}\}, \{free_{\mathsf{Prot}}\}),
\end{aligned}$$

$$\begin{aligned} Bind_{\mathsf{mRNA-RNase}}: &\quad \{bind_{\mathsf{mRNA}}\} + \{bind_{\mathsf{RNase}}\} \xrightarrow{c_6} \\ &\qquad \{comp^{\mathsf{RNase}}_{\mathsf{mRNA}}, comp_{\mathsf{RNase}}\}, \\ Deg_{\mathsf{mRNA}}: &\quad \{unbind_{\mathsf{RNase}}, deg_{\mathsf{mRNA}}\} \xrightarrow{c_7} \\ &\qquad \{free_{\mathsf{RNase}}\}. \end{aligned}$$

Components Rib, RNase and Prot are simple counters because in each global state the current number of ribosomes, proteins and RNase molecules determines the rate of several possible global transitions.

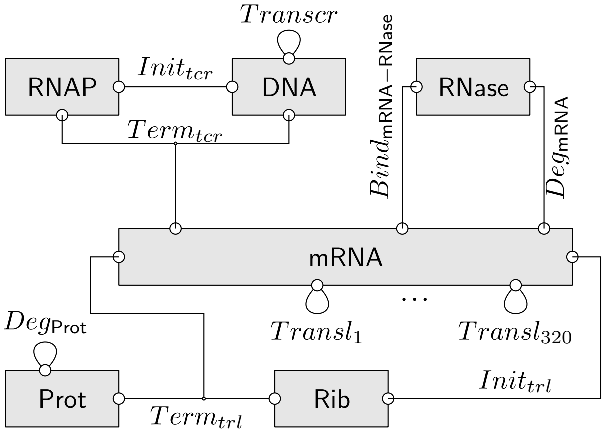

Fig. 6. An overview of the λ phage model.

Finally, we model degradation of proteins by letting

$$Deg_{\mathsf{Prot}} \quad : \quad \{deg_{\mathsf{Prot}}\} \xrightarrow{c_9} \{\}$$

and $w_{\mathsf{Prot}}(q, deg_{\mathsf{Prot}}) = k$ where k equals the current size of the protein population.

Figure 6 gives an overview of the model by illustrating the components as grey boxes. Associated ports are displayed as black dots and the connecting interactions as lines labelled with their respective names.