Stochasticity in reactions: a probabilistic Boolean modeling approach

Sven Twardziok*
Charité Berlin
Institute for Microbiology and Bioinformatics
Berlin, Germany
sven.twardziok@charite.de

Heike Siebert
DFG-FZ MATHEON
Freie Universität Berlin
FB Mathematik und Informatik
Berlin, Germany
siebert@mi.fu-berlin.de

Alexander Heyl
Freie Universität Berlin
Institute for Applied Genetics
Berlin, Germany
heyl@zedat.fu-berlin.de

ABSTRACT

Boolean modeling frameworks have long since proved their worth for capturing and analyzing essential characteristics of complex systems. Hybrid approaches aim at exploiting the advantages of Boolean formalisms while refining expressiveness. In this paper, we present a formalism that augments Boolean models with stochastic aspects. More specifically, biological reactions effecting a system in a given state are associated with probabilities, resulting in dynamical behavior represented as a Markov chain. Using this approach, we model and analyze the cytokinin response network of *Arabidopsis thaliana* with a focus on clarifying the character of an important feedback mechanism.

Categories and Subject Descriptors

I.6 [**Simulation and Modeling**]: Model Development—*Modeling methodologies*; J [**Computer Applications**]: Life and Medical Sciences; G [**Mathematics of Computing**]: Probability and Statistics—*Markov processes*

1. INTRODUCTION

Today, mathematical modeling is an integral part of systems biology research. Mathematical formalisms not only offer a rigorous way of collecting information on a given system, but also allow for comprehensive analysis of structural aspects and dynamical behavior of biological networks, testing of hypotheses and a focused approach to experimental design. As a first step in mathematical modeling, we need to choose a modeling formalism capable of representing the system. If enough data of sufficient precision is available, differential equation modeling is often well suited to the task. If quantitative information is lacking and kinetic mechanisms are unknown, or if low copy numbers of biological entities have to be modeled, discrete modeling methods are a good choice.

The decision which modeling formalism to choose is often not clear-cut. Discrete methods might not be able to capture some important aspects of a biological system, a fully quantitative model might be out of reach. Hybrid systems allow to keep the simplicity of discrete models while expanding the expressiveness of the formalism for some aspects. For example, a continuous time flow might be added to a discrete state space in order to evaluate the effects of different time delays on the system's behavior (see e. g. [3, 11]). In this paper we introduce a Boolean modeling framework extended to include stochastic aspects. In the following, we shortly recall basic concepts concerning Boolean modeling and stochastic extensions of Boolean frameworks. In Sect. 2, we introduce the *Stochasticity in Reactions* framework, which is based on a local Boolean modeling approach and introduces probabilities for the processes generating the network dynamics. We shortly discuss several analysis approaches, before we illustrate the method by modeling and analyzing the cytokinin signal transduction network in *Arabidopsis thaliana* in Sect. 3. Joint work of plant biologists, mathematicians and computer scientists on this signaling network, with a focus on identifying the mechanism behind an inhibitory feedback effect involved in regulation of the signal transduction, led to the formalism as it is presented in this paper. We close with a short discussion and perspectives for future work.

1.1 Boolean Models

A Boolean model is a very abstract representation of a system. Assume we consider a system consisting of $N \in \mathbb{N}$ components. Each component is represented by a Boolean variable. Interpretation of the variable values differs considerably depending on the nature of the components modeled. A component having value one might signify that the concentration of the corresponding substance exceeds a given threshold, it might indicate that a gene is being expressed or that a receptor receives a signal and so on. In any case, the state space of the system is the set $\mathcal{S} := \{0,1\}^N$.

The dynamics of the system is encoded in a Boolean function $f : \mathcal{S} \to \mathcal{S}$, whose coordinate functions (called *state variable update functions*) f_i, $i \in \{1,\ldots,N\}$, determine the behavior of the corresponding components. Given the function f and the state space $\mathcal{S}$, a state transition graph $G := (V, E)$ representing the system's dynamics can be derived. The vertex set of this graph corresponds to the state

*Corresponding author.

space of the system and edges are defined based on the function f and on a chosen update strategy. The most common update strategies are the so-called synchronous update linking a state with its image under f and the asynchronous update that does not allow for simultaneous update of component values and results in a non-deterministic representation of the dynamics (see e. g. [7, 12] for definitions).

Boolean models often do not capture all important aspects of a given system, however, they still allow for rigorous representation and analysis and often reveal fundamental properties of the system. Hybrid methods are a way to refine the Boolean approach and to incorporate effects that cannot be modeled in the Boolean framework, such as time delays or stochastic aspects.

1.2 Stochastic expansion

Quite some work has been done considering stochastic expansions of Boolean models, motivated by the non-deterministic nature of many biological systems. Substance concentrations might fluctuate over time and depending on environmental conditions, perturbations or damage might occur and so on. Different approaches incorporating stochastic aspects in Boolean frameworks can be distinguished based on the way stochasticity is included. Some examples for such methods are Probabilistic Boolean Networks [10], Stochasticity in Nodes [9], Stochasticity in Functions [4] and Probabilistic Asynchronous Update [16]. These approaches are closely related but emphasize different aspects of modeling and analyzing biological systems. In the following, we focus exemplarily on the Probabilistic Boolean Network approach for illustrative and comparative purposes.

In a probabilistic Boolean model, the stochastic events effecting a system in a state $s \in \mathcal{S}$ are represented as elementary events of a finite *probability space* $\Omega_s := \{\omega_1, \ldots, \omega_k\}$, $k \in \mathbb{N}$. A probability function $P_s : \Omega_s \to [0,1]$ assigns probabilities to the events. By definition the sum of the probabilities of the elementary events $\omega \in \Omega_s$ is one. In a slight abuse of notation we drop the index s and denote by P the probability function for Ω_s for all $s \in \mathcal{S}$. Evolution of the system in state s is then determined by a random experiment in Ω_s. We denote the set of pairs of a state $s \in \mathcal{S}$ and an elementary event $\omega \in \Omega_s$ by $\mathcal{O}$:

$$\mathcal{O} := \{(s, \omega) : s \in \mathcal{S}, \omega \in \Omega_s\}.$$

The *probabilistic transition function* $\mathsf{f} : \mathcal{O} \to \mathcal{S}$ calculates the successor state for a state s and an elementary event $\omega \in \Omega_s$. Each elementary event $\omega \in \Omega_s$ with $\mathsf{f}(s, \omega) = s'$ represents a possible state transition from s to s'. We illustrate this idea in Figure 1. The system starts in a state s with probability space Ω_s. The next state s' is chosen by a random experiment in Ω_s in agreement with $\mathsf{f}(s, \omega)$.

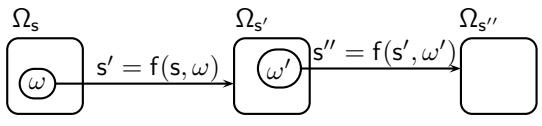


Figure 1: State transitions of a probabilistic Boolean model.

The dynamics of a probabilistic Boolean model is described by a *probabilistic state transition graph*, i. e. a directed weighted graph with a state space as its vertex set. Edges represent possible state transitions while edge weights represent transition probabilities.

Definition 1.1. *We call the graph $\mathcal{G} = (V, \rho, E)$ with vertex set $V := \mathcal{S}$, a function $\rho : \mathcal{S} \times \mathcal{S} \to [0,1]$*

$$\rho(s, s') = \sum_{\substack{\omega \in \Omega_s, \\ \mathsf{f}(s,\omega) = s'}} P(\omega),$$

and edge set $E := \{(s, s') \in V \times V : \rho(s, s') > 0\}$ a probabilistic state transition graph.

Probabilistic Boolean Networks

In a Probabilistic Boolean Network (PBN), there exist $l(n) \in \mathbb{N}$ alternative state variable update functions $f_n^1, \ldots, f_n^{l(n)}$ for each state variable $n \in \{1, \ldots, N\}$. A function $f_n^{k_n}$ with $k_n \in \{1, \ldots, l(n)\}$ is chosen for the next update of state variable n with a predefined probability $c_n^{k_n} \in [0,1]$. Naturally, the sum of the update function probabilities associated with a state variable n must be equal to one:

$$\sum_{k=1}^{l(n)} c_n^k = 1\,.$$

The choices of state variable update functions are independent of each other as well as of the current state of the system. A state transition is derived using synchronous update, i. e., all state variables are updated simultaneously according to the chosen update functions. The probability space for an arbitrary state $s \in \mathcal{S}$, is then defined as

$$\Omega_s = \Big\{(k_1, \ldots, k_N) : k_n \in \{1, \ldots, l(n)\}, n \in \{1, \ldots, N\}\Big\}\,.$$

It does not depend on the choice of s. The probability for a single elementary event $\omega = (k_1, \ldots, k_N)$ is given by

$$P(\omega) = \prod_{n=1}^{N} c_n^{k_n}\,.$$

The probabilistic transition function is then defined by

$$\mathsf{f}(s, \omega) = \big(f_1^{k_1}(s), \ldots, f_N^{k_N}(s)\big)$$

for all $s \in \mathcal{S}$ and $\omega = (k_1, \ldots, k_N) \in \Omega_s$.

2. STOCHASTICITY IN REACTIONS

2.1 Model description

The Boolean modeling formalism underlying our framework is based on a description of the biological processes making up the system's behavior as a whole. Rather than giving a global description of the dynamics via a function f, this approach allows us to model in a local fashion often more suited to translate the available knowledge into a model. Extending the Boolean approach, we associate probabilities with the different reactions possible between network components. Referring to the approaches [9, 4] where stochasticity is introduced by considering different global update functions, we call our approach *Stochasticity in Reactions* (SIR).

As a first modeling step, we describe possible reactions in the system qualitatively. Here, a reaction has a local effect on the system in the sense that it influences a subset of system components. For example, we can model substance degradation as a reaction that only influences the substance

in question, while a biochemical reaction consuming and producing some substances will affect every network component representing one of the substances involved. To describe a reaction we represent its effect on the state of the system as a vector. Furthermore, we assign a probability for the reaction to occur which naturally depends on the state of the system.

DEFINITION 2.1. *A* reaction *is a pair* $r := (e, p)$ *consisting of an effect vector* $e \in \{-1, 0, 1\}^N$ *and a function* $p : S \to [0, 1]$ *satisfying* $p(s) = 0$ *for all* $s \in \mathcal{S}$ *with* $s + e \notin \mathcal{S}$. *We say that* r *is* valid in a state $s \in S$ *if* $p(s) \neq 0$.

The vector e specifies the way certain state variables are affected by the reaction. The components with value zero represent network components unaffected by the reaction. Note that different reactions may have the same effect vector. The function p calculates the probability of the reaction taking place in a given state. In the following, we represent p as $p = b^e(s) \cdot pr$, where $b^e : \mathcal{S} \to \{0, 1\}$ is a Boolean function and $pr \in [0, 1]$ is a probability value. Thus, the probability for a reaction is either zero or pr. The Boolean function b^e indicates whether the execution of a reaction in a given state is possible at all. That is, while the effect vector encodes the impact of a reaction, the function b^e encodes the conditions necessary for the reaction to occur. The probability value then is the same for all states that satisfy those conditions. This modeling assumption has proved sufficient for our purposes so far. Nevertheless, it is certainly worth thinking about modeling the probability with greater dependence on the current state, i. e., considering a function $pr : \mathcal{S} \to [0, 1]$ instead of a constant value pr.

Given a set of reactions $\mathcal{R}$, the set of all valid reactions in a state s is denoted with $R_s := \{r \in \mathcal{R} : p(s) \neq 0\}$.

The following example will be used to illustrate the formalism during the first part of our paper.

EXAMPLE 2.1. *Consider a toy network with three state variables, i. e.* $\mathcal{S} = \{0, 1\}^3$, *and a set of four reactions* $R := \{r_1, r_2, r_3, r_4\}$. *Each reaction is defined by its effect vector and probability function:*

$$\begin{aligned}
r_1 &: \ e_1 = (1, -1, 0), \ p_1(s_1, s_2, s_3) = (1 - s_1) \cdot s_2 \cdot pr_1 \,, \\
r_2 &: \ e_2 = (0, 1, -1), \ p_2(s_1, s_2, s_3) = (1 - s_2) \cdot s_3 \cdot pr_2 \,, \\
r_3 &: \ e_3 = (-1, 0, 0), \ p_3(s_1, s_2, s_3) = s_1 \cdot pr_3 \,, \\
r_4 &: \ e_4 = (1, 0, 1), \ p_4(s_1, s_2, s_3) \\
&\qquad = (1 - s_1) \cdot (1 - s_3) \cdot pr_4 \,.
\end{aligned}$$

Here, r_1 *could represent a transformation of a substance represented by the second network component to a substance represented by the first component, which is encoded in the first two components of the effect vector. We want to model that the transformation can only occur, if there is no substance* 1 *present yet and substance* 2 *is available, which translates to the Boolean function* $b^1(s_1, s_2, s_3) = (1 - s_1) \cdot s_2$. *Reaction* r_3 *could be used to model degradation of the substance represented by the first component.*

The sets of valid reactions are $R_{(0,0,0)} = \{r_4\}$, $R_{(0,0,1)} = \{r_2\}$, $R_{(0,1,0)} = \{r_1, r_4\}$, $R_{(0,1,1)} = \{r_1\}$, $R_{(1,0,0)} = \{r_3\}$, $R_{(1,0,1)} = \{r_2, r_3\}$, $R_{(1,1,0)} = \{r_3\}$, $R_{(1,1,1)} = \{r_3\}$.

We have not yet defined probability spaces Ω_s, $s \in \mathcal{S}$, for our model. In agreement with Sect. 1.2, we assume that, given a state s, an elementary event in Ω_s basically describes a possible state transition depending on a given transition function. We will give a more specific definition later. In our approach, state transitions depend on the reactions that may be executed in s. In general, it is possible for more than one reaction to occur in a given state, and even for a set of reactions to be executed simultaneously. Thus a reaction might be involved in several of the possible state transitions. Mathematically speaking, this amounts to a valid reaction $r = (e, p)$ with $r \in R_s$ describing a probability event $A_r \subseteq \Omega_s$ (*reaction event*) from the probability space Ω_s. That is, reaction events are not necessarily elementary events. Thus, given a state s it is possible that two (or more) reactions $r_1, r_2 \in R_s$ exist with $A_{r_1} \cap A_{r_2} \neq \emptyset$. Therefore, we need to define the sets of reactions which might occur in the same state transition and with this the joint probability for the reaction events.

We call two reactions r_i and r_j *compatible* and denote it with $r_i \sim r_j$ if they satisfy the following two conditions:

$$\begin{aligned}
&\exists s \in \mathcal{S} : r_i, r_j \in R_s \,, \\
&\forall s \in \mathcal{S} \text{ with } r_i, r_j \in R_s : \ (s + e_i + e_j) \in \mathcal{S} \,.
\end{aligned}$$

The effect of a set of reactions occurring in a state is calculated as the sum of the single reaction effects. Therefore, we need the compatibility definition of reactions to assure, that this effect does not leave the Boolean state space. This means, reactions which consume or produce the same state variable are not compatible. Note in addition that validity of r_i and r_j in s ensures that corresponding components of the effect vectors have the same value if they are non-zero. Due to this observation, we can easily see that we preserve pairwise compatibility, if we generalize the concept to sets of more than two reactions.

Compatible reactions may occur in the same state transition. We assume that reaction events of compatible reactions in a state s are *stochastically independent*. This is a reasonable assumption from a modeling perspective, since the framework allows to model dependent processes as a single reaction. Furthermore, compatibility ensures that two reactions do not use the same resources or generate the same product. To assign a probability to a reaction event, we define a probability function P on a subset of 2^{Ω_s} for $s \in \mathcal{S}$. The joint probability for the occurrence of two reactions $r_i, r_j \in R_s$ with $r_i = (e_i, p_i)$ and $r_j = (e_j, p_j)$ can then be set as

$$P(A_{r_i} \cap A_{r_j}) = \begin{cases} p_i(s) \cdot p_j(s) & \text{if } r_i \sim r_j \\ 0 & \text{else} \end{cases} .$$

We define the term for more than two compatible reactions accordingly.

If it is essential for modeling purposes to include the possibility of non-compatible reactions being valid in the same state, we have to add one more condition for the choice of reaction probabilities. They need to be defined such that the union of all reaction events in this state is lower or equal one:

$$\forall s \in \mathcal{S} : P\Big(\bigcup_{r \in R_s} A_r\Big) \leq 1.$$

The condition allows for consistent calculation of the probability of unions of reaction events, which is then calculated in the compatible as well as the non-compatible case with

the addition formula of probabilities [2]

$$P\Big(\bigcup_{i=1}^{n} A_i\Big) = \sum_{k=1}^{n} (-1)^{k+1} \sum_{\substack{I \subseteq \{1,\dots,n\},\\ |I|=k}} P\Big(\bigcap_{i \in I} A_i\Big).$$

EXAMPLE 2.2. *Consider again the running example. It is easy to see that the reactions r_1 and r_4 are not compatible. However, both reactions might occur in the state $s = (0,1,0)$. The union of the two reaction events in the state $s = (0,1,0)$ is calculated as:*

$$P(A_1 \cup A_4) = P(A_1) + P(A_4) - P(A_1 \cap A_4) = P(A_1) + P(A_4).$$

Since the reactions are not compatible, we know that the disjunction of their reaction events is empty. That is, only one or the other reaction can occur and the probabilities for the two reactions must be defined such that $p_1(s) + p_4(s) \leq 1$.

As already mentioned, reactions are local processes, and sets of reactions might be involved in state transitions. To capture all possible sets of reactions in a given state s we basically have to consider all subsets M of the set of reactions valid in s such that M contains only compatible reactions. This leads to the following definition.

DEFINITION 2.2. *Let $\mathcal{R}$ be a set of reactions. A* combination *C is a subset of $\mathcal{R}$ such that there exists $s \in \mathcal{S}$ with $r \in R_s$ for all $r \in C$ and $r_i \sim r_j$ for all $r_i, r_j \in C$ with $i \neq j$.*

It is sufficient to check the compatibility for pairs of reactions in C to ensure that the joint effect of the reactions in C results in a well-defined state. Note furthermore that the empty set $\emptyset$ is always a combination. It has probability greater than zero in a state, if the union of all reaction events in this state is lower than one. This basically amounts to the system remaining in its current state.

We set $C_s := \{C : C \text{ combination}, \forall r \in C\, r \in R_s\}$ the set of valid combinations in a state s. Each combination represents a possible state transition. In other word, each combination in a state s represents an elementary event ω from the probability space Ω_s. This leads to the definition

$$\Omega_s := C_s. \tag{1}$$

We already discussed how to determine the probability of compatible reactions occurring. We can now define a probability function on Ω_s, $s \in \mathcal{S}$, which we call again P in a slight abuse of notation. The probability of the execution of a combination C, that is, the probability of an elementary event in Ω_s, can then be calculated as:

$$P(C) \;=\; P\Big(\bigcap_{r \in C} A_r\Big) \cdot \Big(1 - P\Big(\bigcup_{r \in C_s^c} A_r\Big)\Big). \tag{2}$$

Here, the set C_s^c represents the reactions in R_s which are compatible to all reactions in C:

$$C_s^c := \{r : r \notin C, r \in R_s, \forall r' \in C : r \sim r'\}. \tag{3}$$

We summarize the key elements of the modeling approach in the following definition.

DEFINITION 2.3. *Let $N \in \mathbb{N}$. The pair $\mathfrak{S} = (\mathcal{S}, \mathcal{R})$ with state space $\mathcal{S} := \{0,1\}^N$ and a set of $K \in \mathbb{N}$ reactions $\mathcal{R} := \{r_1, \dots, r_K\}$ is called a* Stochasticity in Reactions (SIR) *model.*

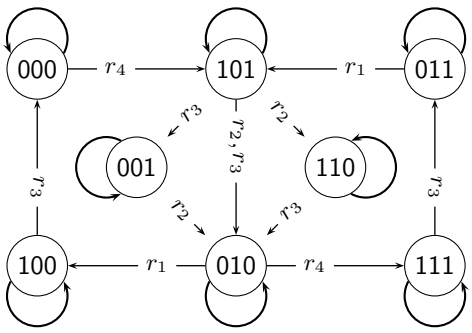


Figure 2: Probabilistic state transition graph for the running example.

2.2 Dynamics

We represent the dynamics of a SIR model as a probabilistic state transition graph. The transition function $\mathsf{f} : \mathcal{O} \to \mathcal{S}$ calculating the successor of a given state s with respect to a combination C of reactions valid in s is defined as follows:

$$\mathsf{f}(s, C) := s + \sum_{e:(e,p) \in C} e\,.$$

Thus, we consider the graph with vertex set $\mathcal{S}$ and edge set derived from the function $\rho : \mathcal{S} \times \mathcal{S} \to [0,1]$,

$$\rho(s, s') = \sum_{\substack{C \in \Omega_s,\\ \mathsf{f}(s,C) = s'}} P(C)\,,$$

where $P(C)$ is calculated according to Equation 2.

The probabilistic state transition graph for the running example is illustrated in Figure 2. For clarity, the edge labels in the figure do not represent the transition probabilities, but the respective combinations corresponding to a transition. The corresponding probabilities are listed in Table 1.

A probabilistic state transition graph $G = (V, \rho, E)$ defines a stochastic state transition matrix $\mathcal{P}$ such that $\mathcal{P} := (\rho_{s,s'})$ with $s, s' \in \mathcal{S}$. A *state distribution* π is a vector of length N with:

$$\pi \in [0,1]^N \text{ and } \sum_{i=1}^{N} \pi_i = 1\,.$$

An initial state distribution π^σ assigns each state the probability of this state being the initial state of the system. If we want to start in a specific initial state $s^0 \in \mathcal{S}$, then we choose the distribution π^σ with $\pi^\sigma_{s^0} = 1$ and $\pi^\sigma_s = 0$ for all $s \neq s^0$.

We are now interested in the sequence of random variables $S(t)$, $t \in \{0, \dots, T\}$ with $T \in \mathbb{N}$, describing the probability that the system is in state s at time $t \in \{0, \dots, T\}$. We calculate the probabilities as follows:

$$\begin{aligned} P(S(0) = s) &= \pi^\sigma_s\,, \\ P(S(t+1) = s' \mid S(t) = s) &= \rho_{s,s'}\,. \end{aligned}$$

Such a sequence $S(t)$ is called a *Markov chain* [2]. A realization of the Markov chain is a sequence of states $\boldsymbol{\xi} = (\xi_0, \dots, \xi_T)$, $\xi_t \in \mathcal{S}$, $t \in \{0, \dots, T\}$, such that the probability with respect to $\mathcal{P}$ and π^σ

$$P\big((S(0), \dots, S(T)) = \boldsymbol{\xi} \mid \pi^\sigma\big) = \pi^\sigma_{\xi_0} \cdot \rho_{\xi_1,\xi_2} \cdot \ldots \cdot \rho_{\xi_{T-1},\xi_T}$$

$s = (0,0,0)$	$\Omega_s = \{\emptyset, \{r_4\}\}$	$P(\emptyset) = 1 - p_4(s)$	$P(\{r_4\}) = p_4(s)$
$s = (0,0,1)$	$\Omega_s = \{\emptyset, \{r_2\}\}$	$P(\emptyset) = 1 - p_2(s)$	$P(\{r_2\}) = p_2(s)$
$s = (0,1,0)$	$\Omega_s = \{\emptyset, \{r_1\}, \{r_4\}\}$	$P(\emptyset) = 1 - p_1(s) - p_4(s)$	$P(\{r_1\}) = p_1(s)$ $P(\{r_4\}) = p_4(s)$
$s = (0,1,1)$	$\Omega_s = \{\emptyset, \{r_1\}\}$	$P(\emptyset) = 1 - p_1(s)$	$P(\{r_4\}) = p_1(s)$
$s = (1,0,0)$	$\Omega_s = \{\emptyset, \{r_3\}\}$	$P(\emptyset) = 1 - p_3(s)$	$P(\{r_3\}) = p_3(s)$
$s = (1,0,1)$	$\Omega_s = \{\emptyset, \{r_2\}, \{r_3\}, \{r_2, r_3\}\}$	$P(\emptyset) = (1 - p_2(s)) \cdot (1 - p_3(s))$ $P(\{r_2, r_3\}) = p_2(s) \cdot p_3(s))$	$P(\{r_2\}) = p_2(s) \cdot (1 - p_3(s))$ $P(\{r_3\}) = (1 - p_2(s)) \cdot p_3(s)$
$s = (1,1,0)$	$\Omega_s = \{\emptyset, \{r_3\}\}$	$P(\emptyset) = 1 - p_3(s)$	$P(\{r_3\}) = p_3(s)$
$s = (1,1,1)$	$\Omega_s = \{\emptyset, \{r_3\}\}$	$P(\emptyset) = 1 - p_3(s)$	$P(\{r_3\}) = p_3(s)$

Table 1: Probability space definitions for each state of Example 2.1.

is larger than zero. Such a sequence of states is called trajectory.

Comparison with other stochastic Boolean frameworks

In the following we show how our formalism relates to the PBN framework described in Sect. 1.2, before we end this section with some general comments. In a PBN model, regardless of the particular state variable update function chosen in a state s, the corresponding state variable remains unchanged, increases or decreases in a state transition. For each state variable $n \in N$ we can define an activation reaction $r_n^+ = (e_n^+, p_n^+)$ and a deactivation reaction $r_n^- = (e_n^-, p_n^-)$ where e_n^+ is the n-th unit vector and $e_n^- = -e_n^+$.

From the PBN definition, activation of a state variable n in a state s may occur, if the state variable is zero and if an update function f_n^k with $k \in 1, \ldots, l(n)$ and $f_n(s) = 1$ exists. A deactivation may occur, if the state variable is zero and if an update function f_n^k with $k \in 1, \ldots, l(n)$ and $f_n(s) = 1$ exists. We can now simply transfer the probabilities in the following way:

$$p_n^+ := (1 - s_n) \cdot \sum_{\substack{k=1, \\ f_n^k(s)=1}}^{l(n)} c_n^k ,$$

$$p_n^- := (s_n) \cdot \sum_{\substack{k=1, \\ f_n^k(s)=0}}^{l(n)} c_n^k .$$

This leads to the set of valid reactions in a state s

$$\begin{aligned} R_s := \; & \{r_n^+ : n \in \{1, \ldots, N\}, s_n = 0, \\ & \exists k \in \{1, \ldots, n\} : f_n^k(s) = 1\} \\ & \cup \{r_n^- : n \in \{1, \ldots, N\}, s_n = 1, \\ & \exists k \in \{1, \ldots, n\} : f_n^k(s) = 0\} . \end{aligned}$$

Compatibility is clearly not a problem, thus we can consider the power set of R_s, $s \in \mathcal{S}$, as probability space Ω_s. It is easy to see that the PBN and the SIR model generate the same probabilistic state transition graph.

The same reasoning applies when transferring a SIR model into a PBN model, as long as there are no reaction which modify values of two or more state variables. If that is not the case, problems arise since in the PBN formalism the state variable update functions are considered completely independent of each other.

A slight generalization of our formalism allows a much more general observation. Given an arbitrary probabilistic state transition graph, we can construct a SIR model that generates this graph. To obtain this result, we allow for the possibility to artificially declare reactions to be non-compatible. Such a declaration poses no problem from a theoretical point of view, we omitted it for the sake of clarity and due to the page restriction. It leads to smaller sets of valid combinations. Given an arbitrary probabilistic state transition graph, we define a reaction r for each transition from a state s to a state s' in the graph. The effect vector e of r is chosen such that $s + e = s'$, the probability function is chosen such that r is only valid in s and the probability in s matches the transition probability given by the graph. With all reactions being defined to be not compatible, the combinations are just the sets containing only one reaction. Obviously the state transition graph of this model matches the one we started with. In this sense, it is possible to express any model framed in one of the formalisms mentioned in the beginning of Sect. 1.2 as a SIR model, but the idea of reactions as local independent events might get lost.

2.3 Analysis

When analyzing a SIR model, we can employ the usual techniques available for Markov chains. Often, we are interested in the state distributions corresponding to certain trajectories. The goal is to identify sequences of state transitions, that is, trajectories in the Boolean state space, that are associated with high probabilities, and thus can be interpreted as most likely behavior of the system. Focussing on trajectories derived from Markov chains, we analyze the corresponding state distributions over time.

Let $\pi(t)$ denote the state distribution of the Markov chain after t time steps, that is, $P(S(t) = s) = \pi(t)_s$. Given an initial state distribution π^σ, it is calculated as:

$$\begin{aligned} & \pi(t) = \pi(t-1)\mathcal{P} \\ & \text{with } \pi(0) := \pi^\sigma . \end{aligned}$$

For more details see e. g. [2]. The probability for the system to be in a certain subset of state space is then calculated as the sum of the corresponding single state probabilities. For example, such a subset might signify an attractor of the system or the set of states where a certain component has value 1.

In addition, existence and reachability of steady states are often points of interest. In the context of Markov chains we consider behavior stable if the state distribution over time converges to a stationary state distribution. If existent the stationary state distribution of a Markov chain is unique. We denote such a stationary state with π^*. Using the stochastic state transition matrix $\mathcal{P}$, a stationary state

satisfies:

$$\pi^* := \pi^* \mathcal{P}\,.$$

Every Markov chain converges to a stationary state distribution, if such a distribution exists. Thus it can be seen as a representation of the long term behavior of a system.

In the remainder of this section we want to focus on analysis of SIR models, where not all or none of the parameters, i. e., the reaction probabilities, are specified. In particular, we are interested in methods to identify parameters. To give an idea of how to approach such a problem we focus on the following very concrete question motivated by the application described in the following section.

Suppose we identified a state s' of particular interest for the functionality of the system. For example, it might be the initial state for some essential process. We now want to choose parameters such that the system favors trajectories reaching the state s'. In particular, given a state s we want to maximize the probability of the system to quickly move from s to s'. In our formalism, that translates to making sure that certain reactions occur while others are omitted.

To state the problem more precisely, we introduce the following notions. A *direct path* from state s to state s' is a path $\boldsymbol{\tau}_{s,s'} = (\tau_0, \ldots, \tau_l)$, $l \in \mathbb{N}$, in the probabilistic state transition graph such that $\tau_0 = s$, $\tau_l = s'$ and $\tau_i \neq \tau_j$ if $i \neq j$. Note that $\boldsymbol{\tau}_{s,s'}$ then represents a trajectory. Trajectories are called *direct trajectories* if we obtain a direct path by merging consecutive identical states. That is, we allow self-loops in the trajectory, and thus there might exist an infinite number of direct trajectories from s to s' that correspond to the same direct path. We define the probability of a direct path $\boldsymbol{\tau}$ as the sum of the probabilities over all direct trajectories that generate $\boldsymbol{\tau}$ by merging consecutive identical states. We now represent a direct path as the corresponding sequence of non empty combinations. We call such a sequence a *combination sequence*.

DEFINITION 2.4. *Given a SIR model $\mathfrak{S} = (\mathcal{S}, \mathcal{R})$, a combination sequence $\mathfrak{r}$ is a finite sequence of non-empty combinations. We say, a combination sequence $\mathfrak{r} := (C^1, \ldots, C^l)$ is valid in a state $s \in \mathcal{S}$ if there exist states $s^1, \ldots, s^l$ with $s = s^1$ such that*

$$C^k \in C_{s^k} \text{ for all } k \in \{1, \ldots, l\} \text{ and}$$
$$s^{k+1} = \mathfrak{f}(s^k, C^k) \text{ for all } k \in \{1, \ldots, l-1\}.$$

We now can use combination sequences to calculate the probability of direct paths.

To calculate the probability of a combination sequence, we introduce a random variable ν representing the combination leading from a state s to a successor s' with $s \neq s'$. In other words, the state variable ν accepts values in the probability space $\Omega_s^\nu := \Omega_s \backslash \{\emptyset\}$. The probability that a specific combination $C \in C_s \backslash \{\emptyset\}$ is chosen for the next state transition is calculated as:

$$P(\nu = C \mid s) := \frac{P(C)}{P(\bigcup_{C \in C_s \backslash \{\emptyset\}} C)}. \tag{4}$$

EXAMPLE 2.3. *Consider the running example in Figure 2. Given the state $s = (0, 1, 0)$ possible combinations in s are $C_s = \{\emptyset, \{r_1\}, \{r_4\}\}$. The probability that the next state transition is due to the combination $C = \{r_1\}$ is:*

$$P(\nu = \{r_1\} \mid (0,1,0)) = \frac{P(\{A_1\})}{P(\{A_1\}) + P(\{A_4\})}.$$

Equation (4) allows us to calculate the probability for a specific state transition. The probability of a combination sequence is then calculated as the product of the corresponding transition probabilities.

We now determine all direct paths, and thus combination sequences, from s to s'. The sum of their probabilities represents the probability to go from s to s' using a direct path. This leads to a function F which calculates this probability in terms of the (non-specified) reaction probabilities of the included reactions. Maximizing this function leads to a parameter set which maximizes the probability to go from s directly to s', which represents a desirable trait of our system.

When modeling biological systems, we often do not know the exact particulars of all processes involved. The optimization approach introduced above can also be used to compare the impact of different reactions on the system. That is, we basically consider several models that differ in only a small set of reactions and compare the impact of these reactions under parameter constraints derived from optimization for effectiveness of the reactions in question. We illustrate this idea in the following section.

Lastly, we want to mention that it is often reasonable to restrict analysis to suitable subgraphs of the probabilistic state transition graph. Here, we can often exploit the underlying discrete network to identify regions of state space of particular interest.

3. THE CYTOKININ SIGNAL TRANSDUCTION NETWORK

3.1 Biological background

Cytokinin is a plant hormone playing an important role in many developmental and physiological processes in the plant, such as regulation of shoot and root growth, leaf senescence and pathogen resistance. The core of the cytokinin signaling system in *Arabidopsis thaliana* is a multistep phospho-relay system that influences the expression of a group of target genes. Some of these genes in turn code for regulators of the signaling pathway activity resulting in a negative feedback effect. However, the nature of the feedback mechanism is still unknown. We used the SIR formalism to model the signaling system and test several hypotheses concerning the nature of the feedback mechanism. For a detailed biological background see [5].

The components involved in the signal transduction can be grouped in families of proteins. Proteins from the same family show very similar behavior. To reduce the complexity of the network, we consider the overall behavior of these families, modeling a family as a single network component. In the following we shortly present the different network components.

Cytokinin
Cytokinins are a class of plant hormones that influence different important plant functions [8]. Cytokinin initiates the expression of several genes, which are called the cytokinin primary response genes [1].

Arabidopsis Histidin Kinases (AHKs)
The AHKs transmit the cytokinin signal from the apoplast into the cell. Cytokinin binds to the ligand binding domain in the extracellular space. This binding causes the canon-

ical histidine residue of the histidine kinase domain to autophosphorylate. After an intramolecular phosphotransfer, the phosphate group can be transferred to an *Arabidopsis* histidin phosphotransfer protein (AHP) via the receiver domain [5, 15].

Arabidopsis Histidin Phosphotransfer Proteins (AHPs)
The AHPs act as kinases. AHPs bind to the receiver domain of AHKs and receive their phosphate group. After the phosphorylation of the AHPs, they translocate into the nucleus and transfer their phosphate group to *Arabidopsis* response regulators (ARR) [5, 15]. These can be grouped into two different families.

Type-B Arabidopsis response regulators (type-B ARRs)
The type-B ARRs expression seems to be independent of the cytokinin signal. However their activity as transcription factor is directly related to the cytokinin concentration. The phosphorylated type-B ARRs activate the transcription of most cytokinin primary response genes [6].

Type-A Arabidopsis response regulators (type-A ARRs)
The type-A ARRs are part of the cytokinin primary response genes. Their transcription is activated by the phosphorylation of the type-B ARR. It has been shown that the type-A ARRs have a (direct or indirect) negative effect on the expression of the cytokinin response genes. Additionally, the phosphorylation of the type-A ARRs results in a higher protein stability.[5, 15]

We focused in our analysis mainly on the character of the feedback mechanism involving type-A ARRs. There exist two theories explaining the negative influence of the type-A ARRs on the cytokinin response. The first theory assumes that the type-A ARRs and the type-B ARRs compete for the phosphate group of AHPs. This is called *AB-competition*. The second theory assumes that active type-A ARRs inhibit the activation of type-B ARRs directly with some unknown mechanism (*AB-inhibition*) [14, 13]. In the following, both theories will be considered.

3.2 SIR Model

State space

We model the components corresponding to cytokinin, AHK, AHP and type-B ARR as Boolean variables. Cytokinin is either *absent* or *present* and the three components AHK, AHP and type-B ARR are either *inactive* or *active*. We need to be more precise in the case of type-A ARRs. Type-A ARRs are either *absent*, *present*/*inactive*, or *present*/*active*. This is modeled using two Boolean variables. The first state variable indicates presence or absence of type-A ARRs and the second specifies their activity status. Type-A ARRs will not be present in their active and their inactive form at the same time. The network component are listed in Table 2 with the associated variable name in parenthesis.

Reactions

Since we want to consider different manifestations of the negative feedback mechanism involving type-A ARRs, we have to consider different definitions for the corresponding reactions. However, several reactions are independent of the assumptions concerning type-ARRs negative feedback. For illustrative purposes, we explain the modeling of some reactions in detail. All specifications are listed in Table 3.

$\mathcal{S}_1$	:	Cytokinin (C)
$\mathcal{S}_2$	:	AHK (K)
$\mathcal{S}_3$	:	AHP (P)
$\mathcal{S}_4$	:	Type-B (B)
$\mathcal{S}_5$	:	Type-A ARR inactive (A)
$\mathcal{S}_6$	:	Type-A ARR active (Aa)

Table 2: Components of the cytokinin signaling network.

Activation of AHK: $K^{(+)}$
AHK autophosphorylates in the presence of cytokinin. This reaction requires the inactive form of the AHK. The two conditions are encoded in the Boolean function $b^{K^{(+)}}(s) = s_1 \cdot (1 - s_2)$ partly defining the probability function. The reaction results in an activation of the AHK. This reaction effects the value of the state variable K only. This is reflected in its effect vector $e_{K^{(+)}} = (0, 1, 0, 0, 0, 0)$.

Activation of AHP: $P^{(+)}$
AHP becomes phosphorylated if AHK is phosphorylated. This reaction requires AHK to be in the active and AHP to be in the inactive form as reflected by the Boolean function $b^{P^{(+)}}(s) = s_2 \cdot (1 - s_3)$. The reaction results in an inactivation of the AHK and an activation of the AHP. This reaction increases the value of state variable P and decreases the value from K as can be seen in its effect vector $(0, -1, 1, 0, 0, 0)$.

Deactivation of type-B ARR: $B^{(-)}$
Type-B ARR tend to be unstable in their phosphorylated form. An explanation for this effect is a dephosphorylation reaction involving no other network components. The resulting effect vector is $(0, 0, 0, -1, 0, 0)$. The reaction requires type-B ARR to be in their active form, thus $b^{B^{(-)}}(s) = s_4$.

Expression of type-A ARR: A_e
The expression of type-A ARR is induced by the phosphorylated type-B ARR. This reaction requires type-B ARR to be in their active form and type-A ARR to be absent.

Degradation of inactive type-A ARR: A_d
Inactive type-A ARR might degrade. This reaction requires inactive type-A ARR.

Activation of type-A ARR: $A^{(+)}$
Type-A ARR become phosphorylated as AHP is phosphorylated. This reaction requires AHP to be in their active form and type-A ARR to be in their inactive form. It causes an activation of type-A ARR.

Degradation of active type-A ARR: Aa_d
Degradation of active type-A ARR might proceed at a different rate than degradation of inactive type-A ARR. Therefore, it is modeled as a distinct reaction. This reaction requires active type-A ARR.

Modeling feedback mechanisms

Activation of type-B ARR:$B^{(+)}$
The two type-A ARR negative feedback assumptions effect type-B ARR activation. Therefore, we consider two different reaction definitions for the two assumptions. As a control we also consider a reaction definition without negative feedback.

Version one: no negative feedback: $B^{(+)1}$

$$
\begin{array}{lcl}
p_{K^{(+)}}(s) & = & pr_{K^{(+)}} \cdot (s_1 \cdot (1 - s_2)) \\
e_{K^{(+)}} & = & (0,1,0,0,0,0) \\
p_{P^{(+)}}(s) & = & pr_{P^{(+)}} \cdot (s_2 \cdot (1 - s_3)) \\
e_{P^{(+)}} & = & (0,-1,1,0,0,0) \\
p_{B^{(+)1}}(s) & = & pr_{B^{(+)1}} \cdot (s_3 \cdot (1 - s_4)) \\
e_{B^{(+)1}} & = & (0,0,-1,1,0,0) \\
p_{B^{(+)2}}(s) & = & pr_{B^{(+)2}} \cdot (s_3 \cdot (1 - s_4) \cdot (1 - s_5)) \\
e_{B^{(+)2}} & = & (0,0,-1,1,0,0) \\
p_{B^{(+)3}}(s) & = & pr_{B^{(+)3}} \cdot (s_3 \cdot (1 - s_4) \cdot (1 - s_6)) \\
e_{B^{(+)3}} & = & (0,0,-1,1,0,0) \\
p_{B^{(-)}}(s) & = & pr_{B^{(-)}} \cdot (s_4) \\
e_{B^{(-)}} & = & (0,0,0,-1,0,0) \\
p_{A_e}(s) & = & pr_{A_e} \cdot (s_4 \cdot (1 - s_5) \cdot (1 - s_6)) \\
e_{A_e} & = & (0,0,0,0,1,0) \\
p_{A_d}(s) & = & pr_{Aa_d} \cdot (s_6) \\
e_{A_d} & = & (0,0,0,0,0,-1) \\
p_{A^{(+)}}(s) & = & pr_{A^{(+)}} \cdot (s_3 \cdot s_5) \\
e_{A^{(+)}} & = & (0,0,-1,0,0,1) \\
p_{Aa_d}(s) & = & pr_{Aa_d} \cdot (s_6) \\
e_{Aa_d} & = & (0,0,0,0,0,-1)
\end{array}
$$

Table 3: Reaction definitions for the cytokinin SIR model.

Type-B ARR become phosphorylated as AHP are active. This reaction requires the active form of AHP and an inactive form of type-B ARR. It results in an inactivation of AHP and an activation of type-B ARR.

Version two: competition: $B^{(+)2}$
Type-A ARR and type-B ARR compete for AHP. Due to this consideration, phosphorylation of the type-B ARR is prevented in the presence of inactive type-A ARR.

Version three: inhibition: $B^{(+)3}$
Type-A ARR inhibits phosphorylation of type-B ARR. Phosphorylation of type-B ARR is prevented in the presence of active type-A ARR.

Due to the different considerations about the type-A ARR negative effect on the cytokinin response, we consider three different reaction sets, which in turn signify three different models:

$$
\begin{array}{lcl}
\mathcal{R}_1 & = & \{K^{(+)}, P^{(+)}, B^{(+)1}, B^{(-)}, A_e, A_d, A^{(+)}, Aa_d\}, \\
\mathcal{R}_2 & = & \{K^{(+)}, P^{(+)}, B^{(+)2}, B^{(-)}, A_e, A_d, A^{(+)}, Aa_d\}, \\
\mathcal{R}_3 & = & \{K^{(+)}, P^{(+)}, B^{(+)3}, B^{(-)}, A_e, A_d, A^{(+)}, Aa_d\}.
\end{array}
$$

We do not yet specify the reaction probabilities. Rather, considerations concerning the stochastic parameters are integrated in the system analysis presented in the following.

3.3 Analysis and Results

We want to compare the different assumptions of competition and inhibition in this part. Therefore, we determine the states in which the two effects influence the type-B ARR activation and analyze the influence of the reaction probabilities for reaching these states. We choose the state $s^0 = (1,0,0,0,0,0)$ as initial state, representing a cytokinin signal and a quiescent state for all components of the signaling pathway. It is easy to see that the first transitions are fully determined by this state. That is, the first reaction that occurs is reaction $K^{(+)}$ followed by $P^{(+)}$. Stochastic aspects come into play following the execution of these reactions.

We proceed as follows. First, we focus on the model representing competition. We consider the set of states for which the reaction modeling the competition is valid, determine the impact of competition on state transitions starting in such states and analyze the influence of reaction probabilities on the probability that the system reaches such states. Maximizing the probability for reaching states where competition may come into play yield optimal parameters for effective competition. We apply the methods utilizing direct paths, which we introduced in Sect. 2.3.

Competition effects the behavior of the system if the system is in a state with type-A ARR and type-B ARR inactive. If cytokinin is present, as we assume throughout this analysis, there is a total number of four such states. All of them are reached from states with active type-B ARR and absent type-A ARR by executing the expression reaction of type-A ARR and deactivation of type-B ARR. Furthermore, the expression of type-A need to occur before or at the same time as the type-B inactivation. Type-B ARR deactivation should occur before type-A activation and before type-A degradation. We assume that $K^{(+)}$ and $P^{(+)}$ are very fast reactions, which is in agreement with biological observations. So whenever $K^{(+)}$ or $P^{(+)}$ are valid reactions they should be immediately executed. This can be modeled by assigning them probability 1. This assumption allows us to determine the influence of the reactions A_e, A_d $A^{(+)}$ and $B^{(-)}$ on the accessibility of the competition state by focussing on a subgraph of the state transition graph.

Figure 3 shows the resulting subgraph of the state transition graph, which is induced by the three states $s^1 = (1,1,1,1,0,0)$, $s^2 = (1,1,1,1,1,0)$ and $s^3 = (1,1,1,0,1,0)$. For clarity, we represent each state only by the components corresponding to variables B, A and Aa. The state $(0,1,0)$ represents the state of interest for competition. Edges represent the possible reaction combinations in the states. Edges without head vertex represent combinations which are valid in the states but do not result in a state of this subgraph. Starting with active type-B ARR and absent type-A ARR, we now want to calculate the probability of the system evolving from $(1,0,0)$ to $(0,1,0)$ without leaving the subgraph. There are two reaction sequences that represent this behavior $\mathfrak{r}_1 = (\{A_e, B^{(-)}\})$ and $\mathfrak{r}_2 = (\{A_e\}, \{B^{(-)}\})$:

$$
\begin{array}{lllll}
(1,0,0) & \xrightarrow{A_e, B^{(-)}} & (0,1,0) & \text{and} & \\
(1,0,0) & \xrightarrow{A_e} & (1,1,0) & \xrightarrow{B^{(-)}} & (0,1,0).
\end{array}
$$

The probability of the reaction sequence $\mathfrak{r}_1$ is calculated using the function $\mathfrak{f}^1 : [0,1]^4 \to [0,1]$ with:

$$
\begin{array}{lcl}
 & & \mathfrak{f}^1(pr_{A_e}, pr_{A_d}, pr_{A^{(+)}}, pr_{B^{(-)}}) \\
 & = & \dfrac{pr_{B^{(-)}} \cdot pr_{Ae}}{pr_{B^{(-)}} + pr_{Ae} - pr_{B^{(-)}} \cdot pr_{Ae}}.
\end{array}
$$

The probability of the reaction sequence $\mathfrak{r}_2$ is calculated using the function $\mathfrak{f}^2 : [0,1]^4 \to [0,1]$ with:

$$
\begin{array}{lcl}
 & & \mathfrak{f}^2(pr_{A_e}, pr_{A_d}, pr_{A^{(+)}}, pr_{B^{(-)}}) \\
 & = & \dfrac{pr_{Ae} \cdot (1 - pr_{B^{(-)}})}{pr_{Ae} + pr_{B^{(-)}} - pr_{B^{(-)}} \cdot pr_{Ae}} \\
 & \times & \dfrac{pr_{B^{(-)}}(1 - pr_{Ad} - pr_{A^{(-)}})}{pr_{B^{(-)}} + pr_{Ad} + pr_{A^{(-)}} - pr_{B^{(-)}} \cdot (pr_{Ad} + pr_{A^{(-)}})}.
\end{array}
$$

We now proceed with an optimization approach as described in Sect. 2.3 to identify parameters that favor competition. This allows us to analyze the maximal effect of the

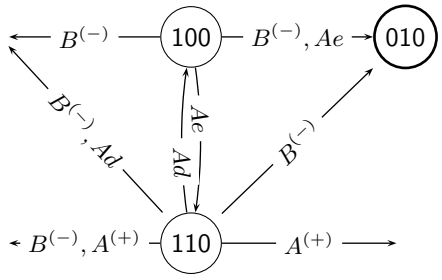


Figure 3: A part of the model state transition graph, for the state variables B, A and Aa with $C = 1$, $K = 1$ and $P = 1$

competition model in terms of negative feedback. To determine optimal conditions for the competition, we optimize the function F with:

$$\begin{aligned} & F(pr_{A_e}, pr_{A_d}, pr_{A^{(+)}}, pr_{B^{(-)}}) \\ = \; & f^1(pr_{A_e}, pr_{A_d}, pr_{A^{(+)}}, pr_{B^{(-)}}) \\ + \; & f^2(pr_{A_e}, pr_{A_d}, pr_{A^{(+)}} pr_{B^{(-)}}) \\ & \text{while } pr_{A_e}, pr_{A_d}, pr_{A^{(+)}}, pr_{B^{(-)}} \in [0.01, 0.5]\,. \end{aligned}$$

Here, we restrict the probability parameter to the interval $[0.01, 0.5]$ in order to allow a minimal reaction probability of 0.01 and to allow a minimal stochastic effect of 0.5. We calculate the maximum of F using a numerical solution algorithm, as implemented for example in MATLAB. Optimization yields a vector $\theta = (0.01, 0.5, 0.01, 0.1429)$ of probability values with $F(\theta) = 0.78$.

Clearly, the calculated probabilities for the reactions A_d and $A^{(+)}$ are at the minimal value of 0.01 while the probability for the reaction A_e is at the maximal possible value with 0.5. This reflects that for effective competition reactions A_d and $A^{(+)}$ should not be executed in the states considered, while reaction A_e should occur rapidly. The probability of $B^{(-)}$ favors the outcome of $B^{(-)}$ occurring before reactions A_d and $A^{(+)}$, but after A_e.

In the following we analyze the behavior of the three models using the parameters favoring competition. A direct comparison is then possible since we use the same parameter set on all three cases. Since we chose the parameters to optimize competition, we would expect that the corresponding model displays strong negative feedback effects, while the other models might show a weaker or no effect at all.

The inhibition mechanism prevents type-B ARR phosphorylation if type-A ARR is active until degradation of the type-A ARR. Therefore the effectiveness of the inhibition is highly dependent on the probability of type-A ARR activation and the probability of the active type-A ARR degradation. We consider three different values for active type-A ARR degradation for the analysis. The parameter specifications are summarized in the following table.

$pr_{K^{(+)}} =$	1	$pr_{A_e} =$	0.5
$pr_{P^{(+)}} =$	1	$pr_{A^{(+)}} =$	0.01
$pr_{B^{(+)}} =$	0.1	$pr_{A_d} =$	0.01
$pr_{B^{(-)}} =$	0.1429	$pr_{Aa_d} \in$	$\{0.001, 0.01, 0.1\}$

For our analysis we focus on the behavior of the components B and Aa since they are directly involved in the different mechanisms. Figure 4 shows state variable activity for the state variables B and Aa for the reaction sets $\mathcal{R}_1$, $\mathcal{R}_2$ and $\mathcal{R}_3$. Here, state variable activity denotes the probability of the system to be in a state where the corresponding state variable has value one. Thus, the curves in the figure illustrate the probability of occurrence of active type-B ARR and active type-A ARR. We used the initial state distribution π^σ specifying $s^0 = (1, 0, 0, 0, 0, 0)$ as initial state, i. e., $\pi^\sigma_{s^0} = 1$, and plotted the probabilities over 500 time steps. The different curves in one coordinate system correspond to different choices for the probability of active type-A degradation.

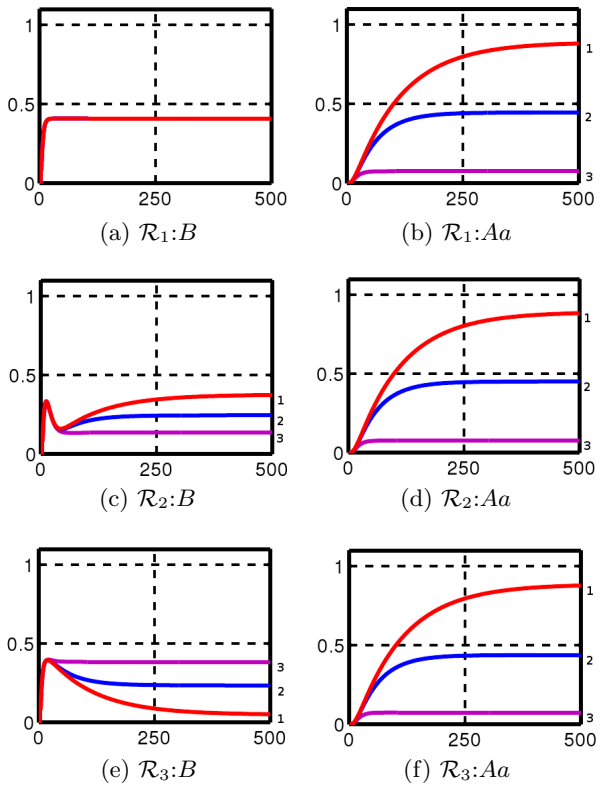
The figure shows that Aa activity hardly varies depending on the choice of model, while a strong impact of the choice of probability value for its degradation can be observed. Activity of B is completely independent of Aa activity for reaction set $\mathcal{R}_1$ due to the absence of negative influence in the control model. For reaction set $\mathcal{R}_2$, modeling competition, higher stability of active type-A ARR (modeled by lower probabilities for Aa degradation) influences B activity in the beginning, yet ongoing observation reveals a stronger inhibition of B activity if we assume low stability of active type-A ARR. Lastly, in the inhibition model effectiveness of the inhibition increases with increasing stability of Aa.

The results show that competition shows a strong inhibitory effect on type-B ARR activity only if stability of type-A ARR is rather low. In contrast, the inhibitory mechanism is the more effective the higher the stability of Aa. Based on these observations, further biological experiments relating stability of active type-A ARR to regulation of the cytokinin response genes might help to clarify the nature of the negative feedback mechanism. Earlier experimental results indicate that type-A ARR stability increases with activation [13]. Together with our calculation, this allows for a tentative hypothesis favoring the inhibitory mechanism.

4. DISCUSSION

In this paper we present a new method to model complex interactions in biological networks using a hybrid framework combining Boolean modeling and stochastic effects. Compared to other probabilistic Boolean models, our method is based on a more local approach, modeling reactions usually only involving a small subset of the system's components. This allows for a very flexible choice with regard to the update strategy determining state transitions in the Boolean state space. Reactions group together changes in component values that are dependent on each other and thus should be executed at the same time. It is furthermore possible to group reactions together in order to consider simultaneous effects of different reactions as well.

Probabilities assigned to reactions allow for modeling of several important aspects influencing the behavior of a biological system that cannot be captured by a purely Boolean approach. Reaction probabilities may represent uncertainties in the execution of processes due to environmental conditions or faulty realization, but they can also be used to distinguish fast and slow processes. Analysis of the resulting probabilistic state transition graph can focus on a variety of aspects, for example determination of trajectories with high probabilities or examination of the importance of a given reaction for the system's dynamics. Since parameter identification is clearly an issue, analysis of unspecified models can be carried out in order to obtain statements relating pa-



(a) $\mathcal{R}_1$:B (b) $\mathcal{R}_1$:Aa

(c) $\mathcal{R}_2$:B (d) $\mathcal{R}_2$:Aa

(e) $\mathcal{R}_3$:B (f) $\mathcal{R}_3$:Aa

Figure 4: State variable activity of B and Aa for the three reaction sets $\mathcal{R}_1$, $\mathcal{R}_2$ and $\mathcal{R}_3$ and with three different active type-A degradation probabilities: $1 : pr_{Aa_d} = 0.001$, $2 : pr_{Aa_d} = 0.01$ and $3 : pr_{Aa_d} = 0.1$

rameter values to possible behaviors. We plan to investigate possible approaches to such questions in future work.

We used the Markov property to simulate the dynamic of the system, which allows for effective simulation and analysis exploiting the rich theory and existing tools for Markov chains. However, the Markov property is a very strong assumption in the context of modeling biological systems, since processes of different time scales might become effective when taking into account accumulation effects. This difficulty can be addressed using additional state variables simulating specific memory effects. However, the addition of state variables should be limited due to the exponential growth of the state space and the probabilistic state transition graph. To balance the effects, methods focussing on analysis of submatrices of the state transition matrix, representing independent modules of the system, need to be studied in more detail. It might also be fruitful to study the properties of the underlying discrete model more closely to exploit available network reduction methods.

The presented formalism already proved useful in application. The mechanism by which the negative feedback loop of cytokinin signaling works has been much discussed but remains obscure. The work presented here marks the protein stability of the type-A ARRs as a decisive factor determining whether inhibition or competition is more likely. Protein stability of the type-A ARRs has not been previously considered in this context. Thus the results of this study provide us with a new experimentally testable hypothesis.

5. REFERENCES

[1] W. G. Brenner, G. A. Romanov, I. Köllmer, L. Bürkle, and T. Schmülling. Immediate-early and delayed cytokinin response genes of arabidopsis thaliana identified by genome-wide expression profiling reveal novel cytokinin-sensitive processes and suggest cytokinin action through transcriptional cascades. *Plant J*, 44(2):314–333, 2005.

[2] K. L. Chung. *Elementary Probability Theory with Stochastic Processes*. Springer-Verlag, 1974.

[3] H. de Jong, J.-L. Gouzé, C. Hernandez, M. Page, T. Sari, and J. Geiselmann. Qualitative simulation of genetic regulatory networks using piecewise-linear models. *Bull. Math. Biol.*, 66(2):301–340, 2004.

[4] A. Garg, K. Mohanram, A. D. Cara, G. D. Micheli, and I. Xenarios. Modeling stochasticity and robustness in gene regulatory networks. *Bioinformatics*, 25(12):i101–i109, 2009.

[5] A. Heyl and T. Schmülling. Cytokinin signal perception and transduction. *Curr Opin Plant Biol*, 6(5):480–488, 2003.

[6] I. Hwang and J. Sheen. Two-component circuitry in arabidopsis cytokinin signal transduction. *Nature*, 413(6854):383–389, 2001.

[7] S. A. Kauffman. *Origins of Order*. Oxford University Press, 1993.

[8] D. W. Mok and M. C. Mok. Cytokinin metabolism and action. *Annu Rev Plant Physiol Plant Mol Biol*, 52:89–118, 2001.

[9] A. Ribeiro, R. Zhu, and S. A. Kauffman. A general modeling strategy for gene regulatory networks with stochastic dynamics. *J Comput Biol*, 13(9):1630–1639, 2006.

[10] I. Shmulevich, E. R. Dougherty, S. Kim, and W. Zhang. Probabilistic boolean networks: a rule-based uncertainty model for gene regulatory networks. *Bioinformatics*, 18(2):261–274, 2002.

[11] H. Siebert and A. Bockmayr. Temporal constraints in the logical analysis of regulatory networks. *Theoretical Computer Science*, 391:258–275, 2008.

[12] R. Thomas and M. Kaufman. Multistationarity, the basis of cell differentiation and memory. ii. logical analysis of regulatory networks in terms of feedback circuits. *Chaos*, 11(1):180–195, 2001.

[13] J. P. C. To, J. Deruère, B. B. Maxwell, V. F. Morris, C. E. Hutchison, F. J. Ferreira, G. E. Schaller, and J. J. Kieber. Cytokinin regulates type-a arabidopsis response regulator activity and protein stability via two-component phosphorelay. *Plant Cell*, 19(12):3901–3914, 2007.

[14] J. P. C. To, G. Haberer, F. J. Ferreira, J. Deruère, M. G. Mason, G. E. Schaller, J. M. Alonso, J. R. Ecker, and J. J. Kieber. Type-a arabidopsis response regulators are partially redundant negative regulators of cytokinin signaling. *Plant Cell*, 16(3):658–671, 2004.

[15] J. P. C. To and J. J. Kieber. Cytokinin signaling: two-components and more. *Trends Plant Sci*, 13(2):85–92, 2008.

[16] L. Tournier and M. Chaves. Uncovering operational interactions in genetic networks using asynchronous boolean dynamics. *J Theor Biol*, 260(2):196–209, 2009.