BIOINFORMATICS

Vol. 25 ISMB 2009, pages i101–i109
doi:10.1093/bioinformatics/btp214

Modeling stochasticity and robustness in gene regulatory networks

Abhishek Garg[1], Kartik Mohanram[2], Alessandro Di Cara[3], Giovanni De Micheli[1] and Ioannis Xenarios[4,*]

[1]Ecole Polytechnique Federale de Lausanne, Station 14, 1015 Lausanne, Switzerland, [2]Rice University, Houston, TX 77005, USA, [3]Merck Serono, 1202 Geneva and [4]Swiss Institute of Bioinformatics, Vital-IT Group, 1015 Lausanne, Switzerland

ABSTRACT

Motivation: Understanding gene regulation in biological processes and modeling the robustness of underlying regulatory networks is an important problem that is currently being addressed by computational systems biologists. Lately, there has been a renewed interest in Boolean modeling techniques for gene regulatory networks (GRNs). However, due to their deterministic nature, it is often difficult to identify whether these modeling approaches are robust to the addition of stochastic noise that is widespread in gene regulatory processes. Stochasticity in Boolean models of GRNs has been addressed relatively sparingly in the past, mainly by flipping the expression of genes between different expression levels with a predefined probability. This stochasticity in nodes (SIN) model leads to over representation of noise in GRNs and hence non-correspondence with biological observations.

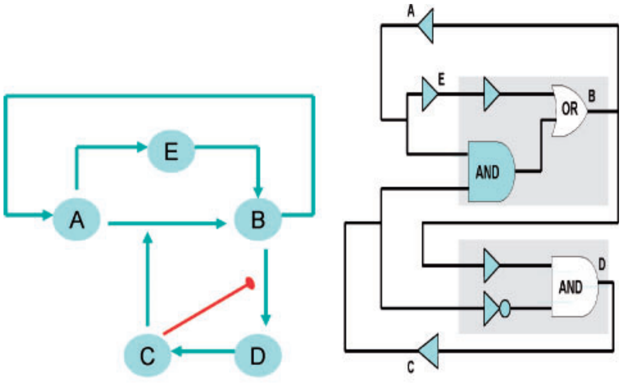
Results: In this article, we introduce the stochasticity in functions (SIF) model for simulating stochasticity in Boolean models of GRNs. By providing biological motivation behind the use of the SIF model and applying it to the T-helper and T-cell activation networks, we show that the SIF model provides more biologically robust results than the existing SIN model of stochasticity in GRNs.

Availability: Algorithms are made available under our Boolean modeling toolbox, *GenYsis*. The software binaries can be downloaded from http://si2.epfl.ch/~garg/genysis.html.

Contact: abhishek.garg@epfl.ch

1 INTRODUCTION

With the advent of high-throughput technologies and advanced measurement techniques, biologists are now able to study gene regulation processes in significant detail. An improved understanding of functions that control gene expression facilitates representation of biological interactions inside a cell using *gene regulatory networks* (GRNs). A GRN represents interactions between a gene/protein and its regulators (such as proteins, transcriptional factors and mRNA). A small synthetic GRN is shown in Figure 1a. *Nodes* in Figure 1a represent the genes/proteins and *edges* represent the biological interactions among the genes/proteins. As the complexity of these networks grows, a need for tools to model these networks becomes more evident (Bernot *et al*., 2004; Chabrier *et al*., 2004; Devloo *et al*., 2003; Garg *et al*., 2008; Karlebach and Shamir, 2008; Klamt *et al*., 2006; Naldi *et al*., 2007; Remy *et al*., 2006).



(a) A Gene Regulatory Network (b) Mapped GRN

Fig. 1. (**a**) A GRN. (**b**) The GRN mapped to Boolean functions (gates). The labels next to the gates represent the output genes/proteins. The gates susceptible to stochasticity are colored dark.

Prior work on modeling the dynamics of GRNs can be broadly categorized into continuous and discrete modeling approaches. The continuous modeling approach uses coupled *ordinary differential equations* (ODEs) to model evolution of gene expression over time (Chen *et al*., 2004; Goodwin, 1963; Li *et al*., 2008). Modeling at the ODE level requires knowledge of kinetic rate constants that are either scarce or are difficult to estimate for each interaction in large networks, thereby limiting the application of ODEs to small well-studied networks. On the other hand, in discrete modeling approaches such as Boolean networks (Albert and Othmer, 2003; Fauré *et al*., 2006; Kauffman, 1969; Kauffman *et al*., 2003; Mendoza and Xenarios, 2006; Thomas, 1991), the concentration of reacting species has a discrete value of 1 or 0 representing the protein/gene being ‘present’ or ‘absent’. Interactions between the genes are modeled using Boolean functions. An example of a GRN mapped to a Boolean network using logic gates is shown in Figure 1b. A logic gate in Figure 1b represents the biological interaction among different genes/proteins. By raising the level of abstraction at which the gene interactions are modeled, Boolean modeling can simulate GRNs with thousands of proteins and compute steady states in a few minutes (Garg *et al*., 2008) on a regular desktop computer.

1.1 Stochasticity in GRNs

Gene regulation processes have been shown to be inherently stochastic (Becskei and Serrano, 2000; Elowitz *et al*., 2002;

*To whom correspondence should be addressed.

Kaern *et al*., 2005; Losick and Desplan, 2008; McAdams and Arkin, 1999; Pedraza and Oudenaarden, 2005). At low concentrations of reacting species, the probability of two molecules undergoing a biochemical reaction decreases, thereby adding a stochastic effect on the reaction product concentration. This approach of noisy gene regulation can be efficiently simulated in continuous modeling approaches by using *chemical master equations* (CMEs) and Gillespie's algorithm (Gillespie, 1976, 1977; Gonze and Goldbeter, 2006; Rao *et al*., 2002; Schultz *et al*., 2007). However, it is still required that all the kinetic rate constants are known so as to simulate the stochastic behavior of GRNs, thereby restricting the application to only small well-studied networks.

To simulate the stochastic effects due to low concentrations of reacting species in Boolean models, methods are proposed in the literature where the nodes in the GRN are flipped from 0 to 1 or vice versa with some predefined flip probability (Álvarez-Buylla *et al*., 2008; Davidich and Bornholdt, 2008; Ribeiro and Kauffman, 2007; Willadsena and Wiles, 2007). With this model of generating stochasticity in Boolean models, differentiation into multiple steady states can be simulated. We refer to this model of stochasticity by the term *stochasticity in nodes* (or SIN). Unlike the CME approach, the SIN model of stochasticity does not take into consideration the correlation between the expression values of reacting species and the probability of flipping the expression of a node due to noise. Further, the SIN approach models the stochasticity at a node regardless of the susceptibility to stochasticity of the underlying biological function that leads to its activation.

Biological functions can have varying levels of complexity and hence, show varying levels of stochasticity. Although it is experimentally difficult to quantify the measure of stochasticity involved in different biological functions, it is a well-known fact that some functions, such as proteasome degradation, are least prone to stochasticity while functions, such as scaffolding complexes that integrate signals arising from different pathways, are likely to behave most stochastically. In practice, most of the biological functions lie somewhere between the above two extremes. Keeping this in mind, we can broadly divide the probability of stochasticity into three different classes, namely: low probability of error ($\epsilon \approx 0$), medium probability of error ($\epsilon \approx 0.5$) and high probability of error ($\epsilon \approx 1$). Figure 2 gives an example of few biological functions divided into these different classes of stochasticity.

We show in this article that the SIN model of stochasticity often leads to overrepresentation of noise in GRNs by making all the genes/proteins equally likely to flip, independent of the expression of the input genes and complexity of the underlying biological function. If a state of the network is defined as a snapshot of the expression of all the genes/proteins, this SIN model can potentially predict transitions among any two states of the network. We will see in Section 2 that overrepresentation of noise is a limitation of the SIN model. We propose an alternative stochastic model, called the *stochasticity in functions* (SIF) model to address the shortcomings of the SIN model of stochasticity in GRNs. In the SIF model, stochasticity is induced at the level of biological functions rather than at the level of expression of a protein/gene. SIF associates a probability of failure with different biological functions and models stochasticity in these functions depending upon the expression of the input nodes (similar to concentration of reactant species in CME). With the above two constraints in the SIF model, the probability of flipping a node at a given time instant depends upon the probability of function failure and the activity of other nodes in the network at that instant in time, thereby making it possible to integrate the stochasticity due to complexity of a biological function with the dynamics of the GRN.

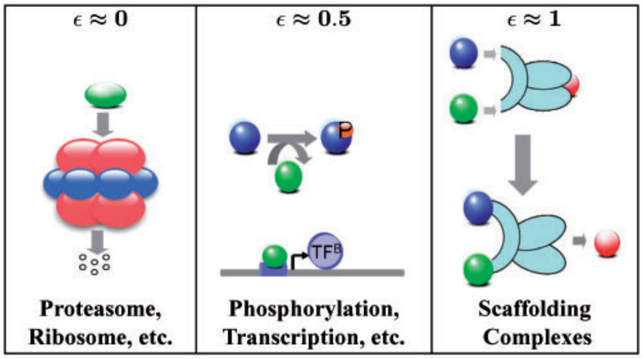


Fig. 2. Biological functions categorized into three different classes of stochasticity and error probability. From left to right, we can broadly classify different biological processes from very stable structures to highly stochastic systems involving scaffold proteins.

Another method proposed in the literature for modeling the probabilistic dynamics in GRNs is based on *probabilistic Boolean networks* (PBNs) (Shmulevich *et al*., 2002). However, PBNs model the choice between alternate biological functions for activating a given gene/protein in a GRN, rather than modeling the stochasticity of the function failure itself. Further, the robustness in PBNs is modeled either by using the SIN type of stochastic models or by modifying Boolean functions (Xiao and Dougherty, 2007). In an overly constrained GRN, where only a single biological function exists for every gene/protein, a corresponding PBN model of the GRN can be constructed such that each node in the PBN will have two Boolean functions. One Boolean function would correspond to the true value of the biological function and the other Boolean function will assign the noisy biological function to the given gene/protein. Such a PBN could be compared with the corresponding SIF model of the GRN. However, the requirement of having only a single biological function for every node in the GRN significantly constraints the type of GRNs that can be modeled for stochasticity due to function failure using PBNs. On the other hand, SIF with some modifications can be used to compute the robustness in PBNs. However, this is not addressed in the present article where we focus on the stochasticity in Boolean models of GRNs.

1.2 Impact of stochasticity

In modeling GRNs, one is often more interested in knowing the steady-state behavior of the network as compared with that of the transient states. This is because in biological experiments, the measurement and comparision of the transient states across multiple experiments is difficult as the dynamics may vary in each experiment. Steady-state behavior, which corresponds to the end point of an experiment when all cells stabilize, is easier to measure and compare with similar experiments. It is also experimentally easier to understand the impact of stochasticity on the steady-state behavior (by measuring the fraction of different cell phenotypes in an experiment). In the dynamic simulation of GRNs, a state of the network evolves over time and stabilizes in an *attractor* (or the steady state). Hence, an attractor represents the long-term behavior

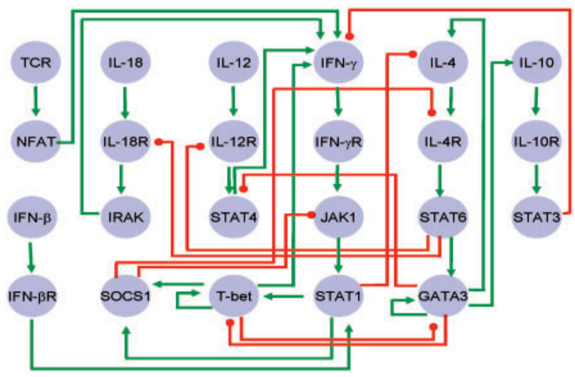


Fig. 3. T-helper GRN (Mendoza and Xenarios, 2006).

of the genes/proteins in the regulatory networks. Attractors (or the steady states) of Boolean networks are hypothesized to correspond to the cellular steady states (or phenotypes) (Huang *et al.*, 2005; Kauffman, 1969; Maamar *et al.*, 2007). In this article, we compare the results obtained using SIN and SIF stochasticity models with respect to two properties of steady states: (a) cellular differentiation in response to an external stimulus and (b) robustness of attractors.

1.2.1 Cellular differentiation In the absence of stochasticity, all biological functions behave as per their description and an initial state of the network differentiates into a specific steady state. However, in the presence of stochasticity in the functions, two similar networks starting from the same initial state may stabilize as different steady states. The probability of differentiating into one steady state can be different from the probability of differentiating into another steady state. This simulation behavior can be used to explain the well-known biological observation of emergence of phenotypically distinct subgroups within an isogenic (Kaern *et al.*, 2005) cell population in response to an input stimuli (such as on exposure to external ligands). A sample simulation experiment on a T-helper Differentiation network (Mendoza and Xenarios, 2006), as in Figure 3, can be effectively used to describe the stochastic differentiation of naïve T-helper cells (i.e. Th0) in response to a pulse of IFNγ, a key cytokine known to play an important role in Th0 to Th1 differentiation. In Figure 4a, cells are initially in a naïve undifferentiated cell state (i.e. Th0). On receiving an input stimulus on IFNγ, cells must differentiate into Th1 cell state in the absence of any stochasticity. This is shown in Figure 4b. Biologically it is known that, while most of the cells should differentiate into Th1 state in response to an IFNγ dosage, a few cells can revert to the Th0 state (Bergmann *et al.*, 2001; Murphy and Reiner, 2002). This difference in response across the cell population is often attributed to inherent stochasticity in biological functions. The stochasticity in cellular differentiation processes can be simulated with different noise models and can give completely different results under SIN and SIF modeling approaches as will be explained through Figures 4c and d in Section 2.1. By computing the differentiation of cell population under the more biologically motivated SIF stochasticity model, one can study if the simulation results of their GRNs correspond to biological observations and then further refine/research the networks accordingly.

1.2.2 Robustness of attractors Robustness of attractors of a GRN can be defined as the probability of an attractor reverting back to itself when the expression of one or more nodes is perturbed from its original expression value. The perturbation method depends upon the stochasticity model used and can give significantly different results based on the model. In the absence of any stochasticity in the biological functions, there should be no transition among two different attractors. If a perturbation changes the state of an attractor, it is possible that the new perturbed state may transition to a different attractor. The perturbed state may be generated in response to external stimuli such as ligands, inhibitors or due to internal stochasticity of the cell. Biologically, cellular steady states are highly robust to internal stochasticity due to redundancy of critical biological functions. Redundant alternative biological pathways to control the expression of genes/proteins is nature's solution to the short term stochastic behavior of subsections of the pathway and are known to exist in abundance in any biological system. To associate high confidence in a GRN, it is imperative that the robustness of cellular steady states is reflected by the robustness of attractors under the stochastic simulations of the corresponding GRNs. Hence, a biologically motivated stochastic model for quantifying the robustness properties of a GRN is essential to compare multiple network configurations for the same biological problem.

2 RESULTS AND DISCUSSION

We studied the SIN and SIF models of stochasticity on two real GRNs proposed in the literature, namely: T-helper network (Mendoza and Xenarios, 2006) and T-cell activation network (Klamt *et al.*, 2006). We modeled these GRNs under increasing number of faults in the network, where a *fault* is defined as the stochastic behavior of a node (for SIN) or the logic gate (for SIF) in the GRN. The results and discussion is organized under the earlier mentioned two properties of steady states, i.e. cellular differentiation and robustness of attractors.

2.1 Cellular Differentiation

2.1.1 T-helper network On simulating the Th0 to Th1 cellular differentiation in response to external IFNγ stimulus under the SIN model of stochasticity, we found that an almost equal number of cells differentiate into Th1 and Th2 from the Th0 cell state and a few cells revert to Th0 (Fig. 4c). Biologically it is known that Th0 cells cannot differentiate to Th2 state in response to an IFNγ stimuli (Murphy and Reiner, 2002). The difference in the simulation results from the known biological observation can be a shortcoming of the GRN, of Boolean modeling, or of the model of stochasticity. The equally likely cellular differentiation of steady states under the SIN model of stochasticity has been observed earlier in Kadanoff *et al.* (2003) and was tagged as a shortcoming of Boolean models. However, in our opinion, SIN model of stochasticity is the main reason behind this discrepancy in simulation results. If we use the more biologically motivated SIF model of stochasticity, where the stochasticity in a biological function is tightly linked to activity of other nodes in the network, we see that a major sub-population of Th0 cells differentiate into Th1 cellular state and a few cells revert to Th0 in response to IFNγ dosage (Fig. 4d). This is consistent with the expected biological behavior of T-helper cells (Murphy and Reiner, 2002) and thereby makes a strong case for the refined SIF model.

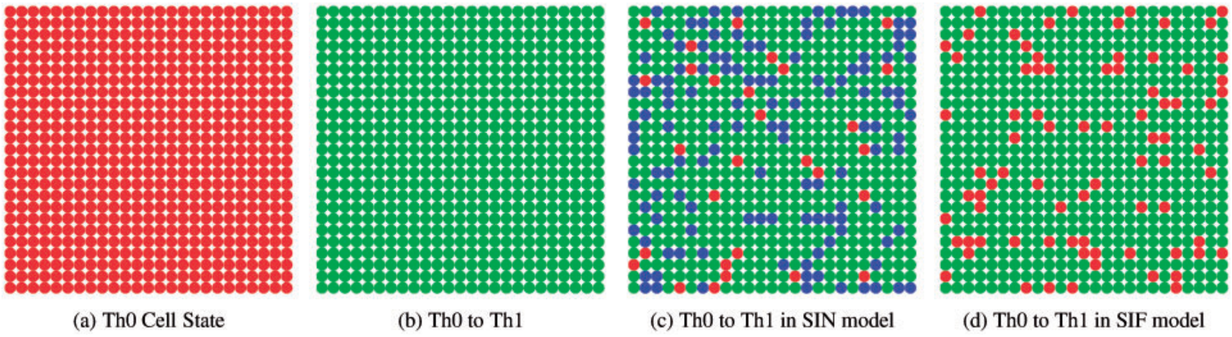


Fig. 4. Simulation results showing the effect of noise on T-helper cell differentiation process with an external stimulus of IFNγ. Each small circle is representative of a T-helper cell and each cell is modeled to behave independent of the neighboring cells. Red cells represent the naïve undifferentiated Th0 cells, green cells represent Th1 cell state and blue cells represent Th2 cell state. Ratio of number of red (green or blue) cells to total number of cells in a panel is representative of the probability of differentiating into Th0 (Th1 or Th2) cell state. (**a**) Cell culture maintained in Th0 state. (**b**) In absence of any stochasticity all Th0 cells differentiate to Th1 cell state on receiving IFNγ. (**c**) Th0 cells differentiate into Th1 and Th2 under the SIN model of stochasticity. Few cells revert to Th0 state as seen by the few patches of red color. (**d**) SIF model of stochasticity shows that Th0 cells differentiate into Th1 cells while some cells cannot differentiate on receiving IFNγ and revert to Th0 cell state. None of the cells differentiate into Th2 cell state. The probability of failure (i.e. ϵ_i) is 0.5 for all the nodes (functions) in the SIN model (SIF model).

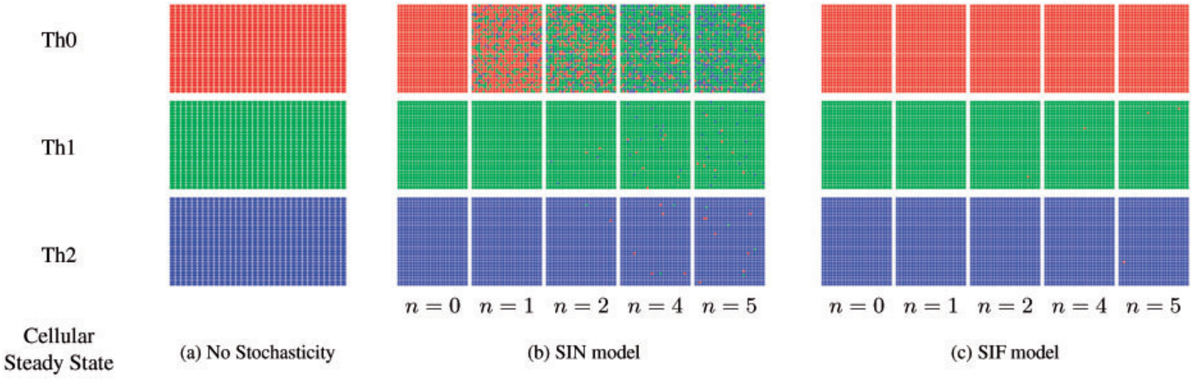
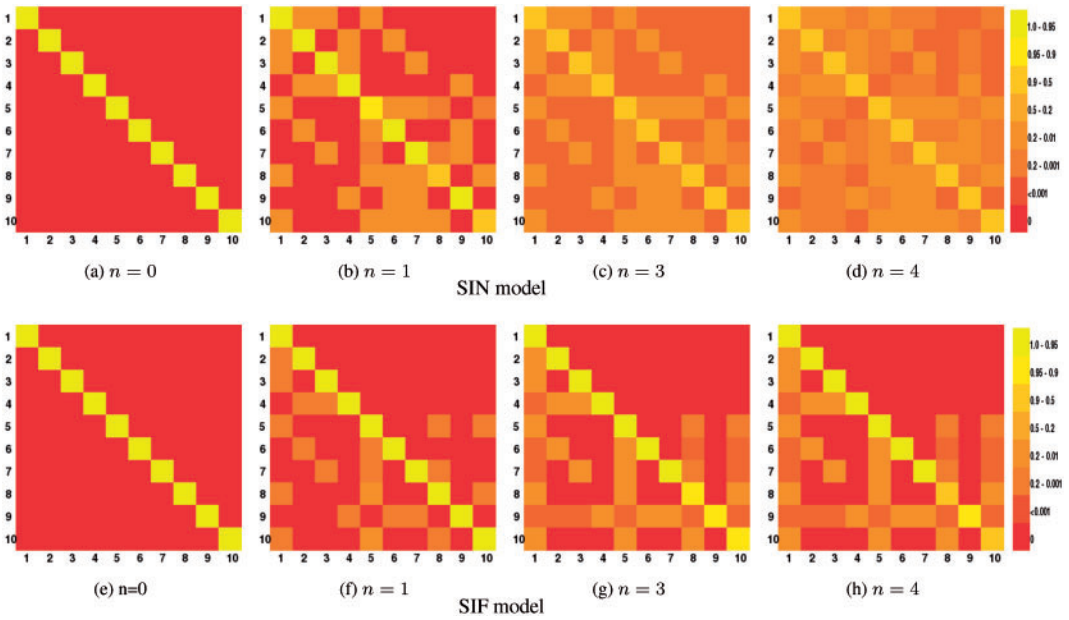


Fig. 5. Simulation results showing the robustness of cellular steady states of T-helper cell differentiation network under SIN and SIF stochasticity models. Red cells represent the naïve undifferentiated Th0 cells, green cells represent Th1 cell state and blue cells represent Th2 cell state. In the first row, ratio of number of red cells to total number of cells in a panel represent the robustness of Th0 cell state. Similarly, second and third row represent the robustness of Th1 and Th2. Number of faults n in the network are modeled from $n=0$ to 5. Multiple faults are injected sequentially (as further discussed in Section 3.2). (**a**) Robustness under no stochasticity. All the cells remain in their steady states. (**b**) Robustness under SIN model of stochasticity. (**c**) Robustness under SIF model of stochasticity. The probability of failure (i.e. ϵ_i) is 0.5 for all the nodes (functions) in the SIN model (SIF model).

2.1.2 T-cell activation Network Unlike the T-helper network, the T-cell activation network does not differentiate into different cell types. The attractors of the T-cell GRN just represent the gene expression profiles of differentially activated T-cells. Hence, the cellular differentiation property is not applicable to the study of the T-cell activation GRN.

2.2 Robustness of attractors

2.2.1 T-helper network We simulated the robustness of attractors of T-helper network under the SIN and the SIF models of stochasticity. Under the SIN model (Fig. 5b), all the three attractors are found to make a transition into each other with a significant probability (represented by the ratio of cells in different cellular states). Robustness of attractors is measured as the number of faults in the network are increased from 0 (i.e. no stochasticity) to 5 faults. One can see from Figure 5b that robustness decreases (i.e. more cells can differentiate to different attractors) if the number of faults in the network is increased. This observation could be specific to the GRN of T-helper network but a similar observation was made on the T-cell activation network. As any GRN is not robust under the SIN model, it may not be a good model of stochasticity for comparing different GRNs. Under the SIF model, Th0 cell state is found to be robust to stochasticity (first row of Fig. 5c). Th1 and Th2 cellular states are very robust as most of the cells stay in the original attractor state even with five sequential faults in the network. Moreover, Th1 and Th2 cells do not show transition among each other as the number of faults is increased. This observation further increases our confidence in the SIF model as biologically Th0, Th1 and Th2 cell states are known to be robust and the underlying T-helper network of Figure 3 is a well-established GRN in the literature.



(a) $n=0$ (b) $n=1$ (c) $n=3$ (d) $n=4$

SIN model

(e) n=0 (f) $n=1$ (g) $n=3$ (h) $n=4$

SIF model

Fig. 6. Simulation results showing the transition probability among the 10 attractors of the T-cell activation network. (**a**)–(**d**) Transition probabilities in the SIN model as the number of faults n in the network is increased from $n=0$ to 4. (**e**)–(**h**) Transition probabilities in the SIF model. In each figure, the intensity of yellow color in the entry $i-j$ corresponds to the probability of transition from the attractor i to the attractor j. The colorbar in the rightmost column indicates the color-probability encoding. The probability of failure (i.e. ϵ_i) is 0.5 for all the nodes (functions) in the SIN model (SIF model).

2.2.2 T-cell activation network We next applied the SIN and SIF models of stochasticity on T-cell activation network from Klamt *et al*. (2006). We measured the probability of reachability among the attractors with an increasing number of faults in the network. The T-cell activation network has 10 attractors in the absence of stochasticity. Since, in the absence of any stochasticity, an attractor cannot make a transition to another attractor, Figure 6a has non-red entries only along the diagonal. As the number of faults is increased, the number of red entries in the heatmaps of Figure 6 decreases showing the decreasing robustness of different attractors. The intensity of yellow color in a cell is proportional to the probability of transition among the corresponding attractors labeled on the X and Y axes. For single fault injection, one can already see that the probability of transitions among these attractors in the SIN model is more widespread than in SIF model on the same network. To test if SIN and SIF models show the same reachability among the attractors with an increasing number of faults, we simulated the injection of four faults sequentially. Just after three faults, almost all the attractors show transitions among each other in the SIN model (Fig. 6c). The results are similar to the those seen earlier for the T-helper network, where a population of cells had a mixture of Th0, Th1 and Th2 cell states in response to internal stochasticity (Fig. 5). Under the SIF model, transitions among the attractors is sparse and a saturation in transition probabilities is observed as the number of faults in the network increases. Since the SIF model is closer to the biological phenomenon of inducing faults in biological functions and does not always give low robustness measure, it can provide an effective way to compare the robustness of two different configurations of GRNs in response to internal stochasticity.

3 METHODS AND TECHNIQUES

3.1 Boolean mapping of GRNs

In Boolean modeling of GRNs, the interaction between genes/proteins is represented using Boolean logic functions such as AND, OR, BUFF and NOT. Truth tables defining the characteristic function of these Boolean functions are shown in Table 1. Boolean logic function IAND corresponds to the biological functionality where the absence of a gene/protein is required for the activity of another gene/protein. In the presence of stochasticity in Boolean functions, output of these logic gates can be different from those specified by their characteristic functions.

Given a GRN (such as in Fig. 1a), the state of a node (or gene) i at time t is represented by a Boolean variable x_i^t. Then the Boolean functions BUFF, NOT, AND, OR, IAND in Figure 1 can be mathematically defined by Equations (1–5), respectively. Symbols $\leftrightarrow, \neg, \vee$ and $\wedge$ in Equations (1–5) stand for logic equivalence, logic negation, logic min and logic max, respectively.

$$\text{BUFF}: \quad f^B(x_a) = (x_c \leftrightarrow x_a) \tag{1}$$

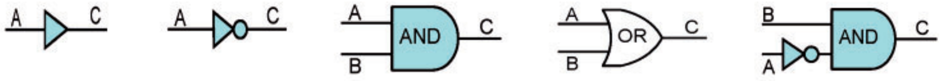
$$\text{NOT}: \quad f^N(x_a) = (x_c \leftrightarrow \neg x_a) \tag{2}$$

$$\text{OR}: \quad f^O(x_1,..,x_p) = (x_c \leftrightarrow \bigvee_{i=1}^{p} x_i) \tag{3}$$

$$\text{AND}: \quad f^A(x_1,..,x_p) = (x_c \leftrightarrow \bigwedge_{i=1}^{p} x_i) \tag{4}$$

$$\text{IAND}: \quad f^{IA}(x_1,..,x_p) = \{x_c \leftrightarrow (\bigwedge_{i=1}^{p^{in}} \neg x_i^{in} \wedge \bigwedge_{j=1}^{p^a} x_j^a)\} \tag{5}$$

Table 1. Truth tables representing the transfer function of different Boolean logic gates. *A* and *B* are the input genes, *C* represents the output gene expression in the absence of stochasticity and $\tilde{C}$ represents the output gene expression in the presence of stochasticity under the SIF model.



A	C	$\tilde{C}$
0	0	0
1	1	0

(a) BUFF

A	C	$\tilde{C}$
0	1	1
1	0	1

(b) NOT

A	B	C	$\tilde{C}$
0	0	0	0
0	1	0	0
1	0	0	0
1	1	1	0

(c) AND

A	B	C	$\tilde{C}$
0	0	0	0
0	1	1	1
1	0	1	1
1	1	1	1

(d) OR

A	B	C	$\tilde{C}$
0	0	0	0
0	1	1	0
1	0	0	0
1	1	0	1

(e) IAND

The expression of each gene i at time $t+1$ can be written as a function $x_i(t+1)$ of the state of the genes acting as its input at time t. The function $x_i(t+1)$ can be formed by composing Boolean gates as in Figure 1b and using the corresponding Equations (1–5). For example, $x_B(t+1)$ for the node B in Figure 1 is defined in Equation (6), where f^B, f^A and f^O are defined in Equations (1), (4) and (3) respectively, and x_C^t and x_E^t represent the expression of nodes C and E at the time instant t.

$$x_B(t+1) = f^O(f^B(x_E^t), f^A(x_A^t, x_C^t)) \quad (6)$$

A snapshot of the activity level of all the genes in the network at time t is called the state of the network. The state of the network at time t can be represented by a Boolean vector, $\mathbf{x}_t$, of size N (number of genes in the network) and is called the *present state* vector. Each bit of this vector represents whether the gene is active or inactive. Another Boolean vector, $\mathbf{x}_{t+1}$, of size N is used to represent the state of the network in the next step and is called the *next state* vector. Assuming a synchronous model of transition, the transition function from the present state to the next state of the network is given by the Boolean function $T_i(\mathbf{x}_t, \mathbf{x}_{t+1})$ in Equations (7) and (8).

$$T_i(\mathbf{x}_t, \mathbf{x}_{t+1}) = \left(x_i^{t+1} \leftrightarrow x_i(t+1)\right) \quad (7)$$

$$T(\mathbf{x}_t, \mathbf{x}_{t+1}) = T_0(\mathbf{x}_t, \mathbf{x}_{t+1}) \wedge \cdots \wedge T_N(\mathbf{x}_t, \mathbf{x}_{t+1}) \quad (8)$$

Equation (7) gives the transition function for a single gene i and states that the value of a gene in the next time step, x_i^{t+1}, is equal to the value of the function $x_i(t+1)$. Equation (8) states that all the genes in the network make a simultaneous transition from the present state $\mathbf{x}_t$ to the next state $\mathbf{x}_{t+1}$.

3.2 Fault model

A *fault* in a GRN is defined as the stochastic behavior of a node or the logic gate in the GRN. In this article, we assume that at most one gene or one function can have a fault at a given instant in time and that multiple faults are spread over different time instants. A sequence of network states from a given starting state to an attractor is called the *trajectory* of the state. If n faults in the network exist then at most n faults can lie on any trajectoy. However, multiple faults cannot exist on a trajectory at the same time instant. We refer to this fault injection model as the *single fault model*. Further, under the single-fault model, given a state of the network, all the possible single faults are independent of each other and can exist with equal probability. This leads to multiple outgoing trajectories from a single state. The assumption of a single fault at a time has been widely used in the literature for stochastic Boolean modeling of GRNs under the SIN model (Álvarez-Buylla *et al.*, 2008; Davidich and Bornholdt, 2008; Ribeiro and Kauffman, 2007; Willadsena and Wiles, 2007). The single-fault model corresponds to a small probability of two distinct biological functions behaving stochastically at the same instant of time.

3.3 Stochasticity in nodes

In the SIN model, any node can flip its expression due to internal stochasticity. Let us represent the internal stochasticity in gene i using a Boolean variable Δ_i, such that if $\Delta_i = 1$ then gene i takes the faulty value and $\Delta_i = 0$ represents the normal expression value for gene i. The transition function for a single gene, $T_i(\mathbf{x}_t, \mathbf{x}_{t+1})$, in Equation (7) can be modified such that x_i^{t+1} is equal to the value of the function $x_i(t+1)$ if there is no fault (i.e. $\Delta_i = 0$). Otherwise x_i^{t+1} takes the value opposite to current value of the function [i.e. $\neg x_i(t+1)$]. Equation (9) represents the modified transition function $T_i(\mathbf{x}_t, \mathbf{x}_{t+1})$ in the presence of a fault in the SIN model.

$$T_i(\mathbf{x}_t, \mathbf{x}_{t+1}) = \left[\left(x_i^{t+1} \leftrightarrow x_i(t+1)\right) \wedge \neg \Delta_i\right] \vee \left[\left(x_i^{t+1} \leftrightarrow \neg x_i(t+1)\right) \wedge \Delta_i\right] \quad (9)$$

Since any node can flip its expression in the SIN model, there are exactly N possible faults in the network at a given instance of time, where N is the number of genes in the network. Under the single-fault model, if the faults in the network are represented by a Boolean vector $\boldsymbol{\Delta}$ of size N, at most one gene x_i has a fault (i.e. $\Delta_i = 1$ for at most 1 bit). Since all the faults are independent of each other, given the state of the network $\boldsymbol{y}_t$, a set of independent and equiprobable fault configuration vector $\boldsymbol{\Delta}$s can exist in the network. That is, if we represent the set of possible fault configuration vectors by a set $D = \{\boldsymbol{\Delta}^1, \boldsymbol{\Delta}^2, \ldots, \boldsymbol{\Delta}^N\}$, the probability of selecting the fault vector $\boldsymbol{\Delta}^i$ is given by Equation (10).

$$P(\boldsymbol{\Delta} = \boldsymbol{\Delta}^i) = \frac{1}{N} \quad (10)$$

If the probability of flipping a node i is given by ϵ^i, then the probability that the gene i has a fault [i.e. $P(\Delta_i = 1)$] in the fault vector $\boldsymbol{\Delta}$ is given by Equation (11).

$$P(\Delta_i = 1) = \Delta_i \cdot \epsilon^i \quad (11)$$

3.4 Stochasticity in functions

SIF models the stochasticity in biological functions that are represented using Boolean gates AND, OR, BUFF, NOT and IAND in Figure 1b. The noisy output values are given in the last column of Tables 1a–e. In Tables 1a–d, noise has an impact on the function only when all the positive inputs are 'active' or 1. This constraint biologically corresponds to the fact that a biological function can behave stochastically only when it is functionally active. For example, transcription of a gene can take place only when the transcription factor is present and there is a natural stochasticity involved in the process of transcription. On the other hand, if the transcription factor is absent, there can be no stochasticity in the transcription process and the gene would never be expressed. Boolean gates BUFF, NOT and AND have all the input ports as positive inputs. Boolean gate IAND have some ports which go through the NOT gate and act as negative inputs. Boolean OR gate is modeled to have no stochasticity because it just represents that the two alternate biological functions can have an impact on the same gene/protein. Note that the noise in these alternate biological functions is already modeled with the remaining stochastic gates (i.e. AND, NOT and BUFF).

Similar to the SIN model, a Boolean variable Δ is used to represent stochasticity in each Boolean function. In the SIN model Δ represents stochasticity in the expression of a gene whereas in the SIF model, Δ represents stochasticity in Boolean functions. Since not all Boolean functions in a GRN behave stochastically, let us define $G=\{G_1,G_2,\ldots,G_p\}$ as a set of stochastic functions in the mapped GRN of Figure 1b. If $\Delta_i=1$, then Boolean function G_i behaves stochastically and take the expression value as defined by the last column of truth Tables 1a–e. Otherwise, if $\Delta_i=0$, Boolean function G_i behaves per its original description. Equations (12–16) formally describe the stochastic Boolean functions.

$$BUFF: f^{B}(x_a)=[(x_c\leftrightarrow\mathbf{0})\wedge\Delta]\vee[(x_c\leftrightarrow x_a)\wedge\neg\Delta] \tag{12}$$

$$NOT: f^{N}(x_a)=[(x_c\leftrightarrow\mathbf{1})\wedge\Delta]\vee[(x_c\leftrightarrow\neg x_a)\wedge\neg\Delta] \tag{13}$$

$$OR: f^{O}(x_1,\ldots,x_p)=(x_c\leftrightarrow\bigvee_{i=1}^{p}x_i) \tag{14}$$

$$AND: f^{A}(x_1,\ldots,x_p)=[(x_c\leftrightarrow\mathbf{0})\wedge\Delta]\vee \left[(x_c\leftrightarrow\bigwedge_{i=1}^{p}x_i)\wedge\neg\Delta\right] \tag{15}$$

$$IAND: f^{IA}(x_1,\ldots,x_p)=\left[\{x_c\leftrightarrow\bigwedge_{j=1}^{p^a+p^{in}}x_j\}\wedge\Delta\right]\vee \left[\{x_c\leftrightarrow(\bigwedge_{i=1}^{p^{in}}\neg x_i^{in}\wedge\bigwedge_{j=1}^{p^a}x_j^{a})\}\wedge\neg\Delta\right] \tag{16}$$

The transition function $T_i(\mathbf{x}_t,\mathbf{x}_{t+1})$ of a gene i for SIF model is the same as Equation (7) for the deterministic Boolean networks. Note that whereas the SIN model modifies the transition function of a gene, the description of Boolean functions is modified in the SIF model.

Given a state of the network $\mathbf{y}_t$, not all Boolean functions in the set G behave stochastically. We use a Boolean vector $\mathbf{\Delta}$ of size $|G|$ to represent the faulty Boolean functions in the network. In a given fault vector $\mathbf{\Delta}$, the bit $\Delta_i=1$ only if all the positive inputs to the function G_i are active or 1. Hence, the number of faults in the network at a given time instant t depends upon the current state of the network. Again, assuming the single-fault model, a set $D=\{\mathbf{\Delta}^1,\mathbf{\Delta}^2,\ldots,\mathbf{\Delta}^{|G|}\}$ of independent fault vectors $\mathbf{\Delta}^i$ may exist such that in each fault vector, at most one Boolean function G_i has a fault (i.e. $\Delta_i=1$ for at most 1 bit). The probability of selecting the fault vector $\mathbf{\Delta}^i$ is given by Equation (17).

$$P(\mathbf{\Delta}=\mathbf{\Delta}^i)=\frac{1}{|D|} \tag{17}$$

Boolean functions in the set G correspond to Biological functions and have a probability of failure ϵ^i associated with each G_i. The probability of failure ϵ^i is independent of the state of the network and solely depends upon the complexity of the biological function that it represents. The probability that Boolean function i has a fault (i.e. $P(\Delta_i=1)$) in a given fault vector is given by Equation (18).

$$P(\Delta_i=1)=\Delta_i\cdot\epsilon^i \tag{18}$$

Whereas the set D of fault configuration vectors does not depend upon the state of the network and is always the same in the SIN model, both the size of the set D and its elements depend upon the current state of the network $\mathbf{y}_t$ in the SIF model. Further, the size of fault configuration vector $\mathbf{\Delta}$ is different for SIF and SIN model. Whereas the size of the vector $\mathbf{\Delta}$ is equal to the number of genes in the network in the SIN model, it is equal to the number of stochastic gates in the SIF model of stochasticity.

With these two models of stochasticity in GRNs, we provide, in the next two sections, algorithms to compute the probability of cellular differentiation and robustness.

3.5 Probability of cellular differentiation

Given a state of the network $\mathbf{y}_t$, a fault vector $\mathbf{\Delta}$ and the transition function $T_i(\mathbf{x}_t,\mathbf{x}_{t+1})$, the state of the network in the next time step, $\mathbf{y}_{t+1}$, can be computed by using Equation (19) where the symbol $\exists$ stands for existential quantification.

$$\mathbf{y}_{t+1}=\exists_{\Delta}\exists_{x_{t+1}}\{T(\mathbf{x}_t,\mathbf{x}_{t+1})\wedge\mathbf{y}_t\wedge\Delta\} \tag{19}$$

Given a fault vector $\mathbf{\Delta}$ of length n and a current state of the network $\mathbf{y}_t$, the probability that the network would exist in the faulty state $\mathbf{y}_{t+1}^{\Delta}$ and the fault-free state $\mathbf{y}_{t+1}$ is given by Equations (20) and (21), respectively.

$$P(\mathbf{y}_{t+1}^{\Delta}|(\mathbf{y}_t,\mathbf{\Delta}))=\sum_{i=1}^{n}P(\Delta_i=1) \tag{20}$$

$$P(\mathbf{y}_{t+1}|(\mathbf{y}_t,\mathbf{\Delta}))=\sum_{i=1}^{n}(1-P(\Delta_i=1)) \tag{21}$$

By applying Bayes' rule on Equations (20) and (21), the probability of the network being in state $\mathbf{y}_{t+1}$ at the next time instant is given by Equations (22–26). In Equation (22), we marginalize the probability over all possible fault configuration vector $\mathbf{\Delta}$s. By using Equation (17) in Equation (22), we get the probability of generating the faulty next state $\mathbf{y}_{t+1}^{\Delta}$ from the current state $\mathbf{y}_t$ in Equation (23). If the network can be in only one starting state $\mathbf{y}_t$, the probability of

Algorithm 1 Algorithm for computing probability of faulty next states.

```
1  stochastic_next_states($T, S_t, P_t, G$)
2  begin
3      $S_{t+1} = \emptyset$
4      for $i = 0$ to $|S_t|$ do
5          $D = $construct_fault_config($S_t^i, G$)
6          $\mathbf{\Delta} = \mathbf{0}$
7          $s_{tmp} = \exists_{\Delta} \exists_{x_{t+1}} \{T(\mathbf{x}_t, \mathbf{x}_{t+1}) \wedge S_t^i \wedge \mathbf{\Delta}\}$
8          for $j = 0$ to $|D|$ do
9              $s^{\Delta}_{tmp} = \exists_{\Delta} \exists_{x_{t+1}} \{T(\mathbf{x}_t, \mathbf{x}_{t+1}) \wedge S_t^i \wedge D^j\}$
10             $P(s^{\Delta}_{tmp}) = P(s^{\Delta}_{tmp} \mid S_t^i) \cdot P_t^{S_t^i}$
11             $P(s_{tmp}) = P(s_{tmp} \mid S_t^i) \cdot P_t^{S_t^i}$
12             $P_{t+1}^{s^{\Delta}_{tmp}} = P_{t+1}^{s^{\Delta}_{tmp}} + P(s^{\Delta}_{tmp})$
13             $S_{t+1} = S_{t+1} \cup s^{\Delta}_{tmp}$
14             $P_{t+1}^{s_{tmp}} = P_{t+1}^{s_{tmp}} + P(s_{tmp})$
15             $S_{t+1} = S_{t+1} \cup s_{tmp}$
16     return $(S_{t+1}, P_{t+1})$
17 end
```

generating the state $\mathbf{y}^{\Delta}_{t+1}$ is given by Equation (24).

$$P(\mathbf{y}^{\Delta}_{t+1}|\mathbf{y}_t) = \sum_{\delta \in D} \{P(\mathbf{y}^{\Delta}_{t+1}|(\mathbf{y}_t, \Delta = \delta)) \cdot P(\Delta = \delta)\} \tag{22}$$

$$= \frac{1}{|D|} \sum_{\delta \in D} P(\mathbf{y}^{\Delta}_{t+1}|(\mathbf{y}_t, \Delta = \delta)) \tag{23}$$

$$P(\mathbf{y}^{\Delta}_{t+1}) = P(\mathbf{y}^{\Delta}_{t+1}|\mathbf{y}_t) \cdot P(\mathbf{y}_t) \tag{24}$$

A similar set of equations exist for the fault-free next state $\mathbf{y}_{t+1}$ [Equations (25) and (26)].

$$P(\mathbf{y}_{t+1}|\mathbf{y}_t) = \frac{1}{|D|} \sum_{\delta \in D} P(\mathbf{y}_{t+1}|(\mathbf{y}_t, \Delta = \delta)) \tag{25}$$

$$P(\mathbf{y}_{t+1}) = P(\mathbf{y}_{t+1}|\mathbf{y}_t) \cdot P(\mathbf{y}_t) \tag{26}$$

If the network may exist in a set of initial states S and the probability of each initial state is specified, then the probability of being in the next states $\mathbf{y}^{\Delta}_{t+1}$ and $\mathbf{y}_{t+1}$ is given by Equations (27) and (28), respectively.

$$P(\mathbf{y}^{\Delta}_{t+1}) = \sum_{y_t \in S} P(\mathbf{y}^{\Delta}_{t+1}|\mathbf{y}_t) \cdot P(\mathbf{y}_t) \tag{27}$$

$$P(\mathbf{y}_{t+1}) = \sum_{y_t \in S} P(\mathbf{y}_{t+1}|\mathbf{y}_t) \cdot P(\mathbf{y}_t) \tag{28}$$

Algorithm 1 describes how the probability of the set of next states S_{t+1} is computed from a given set of initial states S_t. In line 5 of Algorithm 1, the possible fault configuration vectors are computed from an initial state S_t^i $(i = 1, 2, \ldots, |S_t|)$. For each fault configuration vector in the set D, the next states are generated in line 9 and the probability of each next state is computed in lines 10–14 by using Equations (22–28). In Algorithm 1, P_t and P_{t+1} represents the probability of states in the set S_t and S_{t+1}, respectively.

Algorithm 2 describes how the probability of transition into different steady states can be computed from a given set of initial states. In Algorithm 2, given a set of states S we define backward reachable set $BR(S)$ as a set of all the states of the network that can make a transition into the states in S in one or more time steps under the no-stochasticity condition (i.e. $\mathbf{\Delta} = \mathbf{0}$). If SS_{a_i} represents the set of states in an attractor a_i, one can test if the current state of the network $\mathbf{y}_t$ can differentiate into the attractor a_i given a fault vector $\mathbf{\Delta}$ by testing if $BR(SS_{a_i}) \bigcap \mathbf{y}_{t+1} \neq \emptyset$, where $\mathbf{y}_{t+1}$ is computed using Equation (19). The probability of making a transition to an attractor a_i is then given by the sum of $P(\mathbf{y}_{t+1})$ for all the states $\mathbf{y}_{t+1}$ that can make a transition to a_i. The function stochastic_differentiation() in lines 15–23 of Algorithm 2 computes the probability of differentiation into all the cellular steady states from an initial set of states S. In line 17, we compute the probability of all the faulty states that may exist by injecting a single fault in the network. Multiple faults may exist in the network that we model using the function stochastic_differentiation_k_faults() in lines 1–14. This function models k sequential faults in the network. However, each fault is injected under the single-fault model and multiple faults exist only in consecutive time steps.

Algorithm 2 Algorithm for computing probability of differentiation into various attractors in the presence of upto k faults.

```
1  stochastic_differentiation_k_faults($T, S, G, k, SS$)
2  begin
3      for $i = 1$ to $|S|$ do
4          $P_t^{S^i} = 1/S.size()$
5      for $i = 1$ to $|SS|$ do
6          $P_{SS_i} = 0$
7      $S_t = S$
8      for $i = 1$ to $k$ do
9          $(S_{t+1}, P_{t+1}, \tilde{P}_{SS}) = $stochastic_differentiation($T, S_t, G, P_t, SS$)
10         for $j = 1$ to $|SS|$ do
11             $P_{SS_j} = P_{SS_j} + \tilde{P}_{SS_j}/k$
12         $t = t + 1$
13     return $P_{SS}$
14 end
15 stochastic_differentiation($T, S, G, P, SS$)
16 begin
17     $(S_{t+1}, P_{t+1}) = $stochastic_next_states($T, S, P, G$)
18     for $j = 1$ to $|SS|$ do
19         for $i = 1$ to $|S_{t+1}|$ do
20             if $BR(SS_j) \bigcap S^i_{t+1} \neq \emptyset$ then
21                 $P_{SS_j} = P_{SS_j} + P_{t+1}^{S^i_{t+1}}$
22     return $(S_{t+1}, P_{t+1}, P_{SS})$
23 end
```

3.6 Robustness computation

In the absence of any stochasticity, the attractors in a GRN are first computed using algorithms we have proposed in the past (Garg *et al.*, 2008) for the deterministic Boolean modeling of GRNs. The details of these algorithms are omitted for brevity and can be found in Garg *et al.* (2008). By the definition of an attractor, there is no path among the attractors in a deterministic model. If SS_{a_i} represents the states in the attractor a_i, then by using the function stochastic_differentiation_k_faults(), defined in Algorithm 2, the probability of differentiation into various attractors can be computed

Algorithm 3 Algorithm for computing Robustness of Attractors in the presence of upto k faults.

```
robust_attractors_k_faults($T, S, G, k$)
begin
    $T' = \exists_\Delta \{T(\mathbf{x}_t, \mathbf{x}_{t+1}) \wedge \Delta\}$
    $SS = \text{all\_attractors}(T')$
    for $i = 1$ to $|SS|$ do
        $P_{SS}[i] = \text{stochastic\_differentiation\_k\_faults}(T, SS_i, G, k, SS)$
end
```

for every attractor a_i . The probability to differentiate into various attractors in turn represents the robustness of an attractor a_i. Algorithm 3 formally describes the procedure to compute the robustness of attractors.

4 CONCLUSION

We have proposed the SIF model for the simulation of stochasticity in GRNs. Unlike the traditional SIN model that simulates stochasticity by flipping gene expression values, SIF models the stochasticity induced at the level of biological functions. SIF associates a probability of failure with different biological functions and models stochasticity in these functions depending upon the expression of the input nodes. By applying the SIN and SIF models on the T-helper network, we show that whereas the SIN model predicts biologically implausible behavior, the SIF model correctly predicts the Th0 to Th1 cellular differentiation process. Further, when the robustness of steady states of the T-helper and T-cell activation networks is analyzed, we show that the SIN model predicts low robustness properties in both cases whereas the SIF model predicts more biologically relevant behavior with higher robustness. With the improved SIF model of stochasticity, we hope to simulate biological phenomena such as gene perturbation experiments more accurately and to construct GRNs that exhibit strong robustness properties. Algorithms for stochastic simulation of GRNs have been integrated in our Boolean modeling toolbox GenYsis, which is available for download at http://si2.epfl.ch/∼garg/genYsis.html.

Conflict of Interest: none declared.

REFERENCES

Albert,R. and Othmer,H.G. (2003) The topology of the regulatory interactions predicts the expression pattern of the Drosophila segment polarity genes. *J. Theor. Biol.*, **223**, 1–18.

Álvarez-Buylla,E.R. *et al.* (2008) Floral Morphogenesis: Stochastic explorations of a gene network epigenetic landscape. *PLoS ONE*, **3**, e3626.

Becskei,A. and Serrano,L. (2000) Engineering stability in gene networks by autoregulation. *Nature*, **405**, 590–593.

Bernot,G. *et al.* (2004) Application of formal methods to biological regulatory networks: extending Thomas' asynchronous logical approach with temporal logic. *J. Theor. Biol.*, **229**, 339–347.

Bergmann,C. and van Hemmen,J.L. (2001) Th1 or Th2: how an appropriate T helper response can be made. *Bull. Math. Biol.*, **63**, 405–430.

Chabrier,N. *et al.* (2004) The Biochemical Abstract Machine BIOCHAM. In *Proceedings of International conference on Computational Methods in Systems Biology*, Vol. 3082 of *LNBI*. Springer, Berlin, Heidelberg, pp. 172–191.

Chen,K.C. *et al.* (2004) Integrative analysis of cell cycle control in budding yeast. *Mol. Biol. Cell*, **15**, 3841–3862.

Davidich,M.I. and Bornholdt,S. (2008) Boolean network model predicts cell cycle sequence of fission yeast. *PLoS ONE*, **3**, e1672.

Devloo,V. *et al.* (2003) Identification of all steady states in large biological systems by logical analysis. *Bull. Math. Biol.*, **65**, 1025–1051.

Elowitz,M.B. *et al.* (2002) Stochastic gene expression in a single cell. *Science*, **297**, 1183–1186.

Fauré,A. *et al.* (2006) Dynamical analysis of a generic Boolean model for the control of the mammalian cell cycle. *Bioinformatics*, **22**, e124–e131.

Garg,A. *et al.* (2008) Synchronous versus asynchronous modeling of gene regulatory networks. *Bioinformatics*, **24**, 1917–1925.

Gillespie,D.T. (1976) A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.*, **22**, 403–434.

Gillespie,D.T. (1977) Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.*, **81**, 2340–2361.

Gonze,D. and Goldbeter,A. (2006) Circadian rhythms and molecular noise. *Chaos*, **16**, 26–110.

Goodwin,B.C. (1963) *Temporal Organization in Cells; A Dynamic Theory of Cellular Control Processes*. Academic Press, New York.

Huang,S. *et al.* (2005) Cell fates as high-dimensional attractor states of a complex gene regulatory network. *Phys. Rev. Lett.*, **94**, 128701:1–128701:4.

Kadanoff,L. *et al.* (2003) Boolean dynamics with random couplings. *Springer Applied Mathematical Sciences Series*, Special volume, Springer, New York, pp. 23–89.

Karlebach,G. and Shamir,R. (2008) Modelling and analysis of gene regulatory networks. *Nat. Rev. Mol. Cell Biol.*, **9**, 770–780.

Kauffman,S.A. (1969) Metabolic stability and epigenesis in randomly constructed genetic nets. *J. Theor. Biol.*, **22**, 437–467.

Kauffman,S. *et al.* (2003) Random Boolean network models and the yeast transcriptional network. *Proc. Natl Acad. Sci. USA*, **100**, 14796–14799.

Klamt,S. *et al.* (2006) A methodology for the structural and functional analysis of signaling and regulatory networks. *BMC Bioinformatics*, **7**, 56.

Kaern,M. *et al.* (2005) Stochasticity in gene expression: from theories to phenotypes. *Nat. Rev. Genet.*, **6**, 451–464.

Li,S. *et al.* (2008) A quantitative study of the division cycle of Caulobacter crescentus stalked cells. *PLoS Comput. Biol.*, **4**, e9.

Losick,R. and Desplan,C. (2008) Stochasticity and cell fate. *Science*, **320**, 65–68.

Maamar,H. *et al.* (2007) Noise in gene expression determines cell fate in Bacillus subtilis. *Science*, **317**, 526–529.

McAdams,H.H. and Arkin,A. (1999) Its a noisy business! Genetic regulation at the nanomolar scale. *Trends Genet.*, **15**, 65–69.

Mendoza,L. and Xenarios,I. (2006) A method for the generation of standardized qualitative dynamical systems of regulatory networks. *Theor. Biol. Med. Model.*, **3**, 13.

Murphy,K.M. and Reiner,S.L. (2002) The lineage decisions on helper T cells. *Nat. Rev. Immunol.*, **2**, 933–944.

Naldi,A. *et al.* (2007) Decision diagrams for the representation and analysis of logical models of genetic networks. Vol. 4695 of *LNCS/LNBI*, Springer, Berlin, Heidelberg, pp. 233–247.

Pedraza,J.M. and Oudenaarden,A.V. (2005) Noise propagation in gene networks. *Science*, **307**, 1965–1969.

Rao,C.V. *et al.* (2002) Control, exploitation and tolerance of intracellular noise. *Nature*, **420**, 231–237.

Remy,E. *et al.* (2006) From logical regulatory graphs to standard petri nets: dynamical roles and functionality of feedback circuits. Vol. 4230 of *LNCS*, Springer, Berlin, Heidelberg, pp. 56–72.

Ribeiro,A.S. and Kauffman,S.A. (2007) Noisy attractors and ergodic sets in models of gene regulatory networks. *J. Theor. Biol.*, **247**, 743–755.

Schultz,D. *et al.* (2007) Molecular level stochastic model for competence cycles in Bacillus subtilis. *Proc. Natl Acad. Sci. USA*, **104**, 17582–17587.

Shmulevich,I. *et al.* (2002) Probabilistic Boolean networks: a rule-based uncertainty model for gene regulatory networks. *Bioinformatics*, **18**, 261–274.

Thomas,R. (1991) Regulatory networks seen as asynchronous automata: a logical description. *J. Theor. Biol.*, **153**, 1–23.

Willadsena,K. and Wiles,J. (2007). Robustness and state-space structure of Boolean gene regulatory models. *J. Theor. Biol.*, **249**, 749–765.

Xiao,Y. and Dougherty,E.R. (2007) The impact of function perturbations in Boolean networks. *Bioinformatics*, **23**, 1265–1273.