

## Reliable dynamics in Boolean and continuous networks

Eva Ackermann<sup>1</sup>, Tiago P Peixoto<sup>2,3</sup> and Barbara Drossel<sup>1</sup>

<sup>1</sup> Institut für Festkörperphysik, TU Darmstadt, Hochschulstraße 6,  
D-64289 Darmstadt, Germany

<sup>2</sup> Institut für Theoretische Physik, Universität Bremen, Hochschulring 18,  
D-28359 Bremen, Germany

E-mail: [tiago@itp.uni-bremen.de](mailto:tiago@itp.uni-bremen.de)

*New Journal of Physics* **14** (2012) 123029 (17pp)

Received 2 October 2012

Published 18 December 2012

Online at <http://www.njp.org/>

doi:10.1088/1367-2630/14/12/123029

**Abstract.** We investigate the dynamical behavior of a model of robust gene regulatory networks which possess ‘entirely reliable’ trajectories. In a Boolean representation, these trajectories are characterized by being insensitive to the order in which the nodes are updated, i.e. they always go through the same sequence of states. The Boolean model for gene activity is compared with a continuous description in terms of differential equations for the concentrations of mRNA and proteins. We found that entirely reliable Boolean trajectories can be reproduced perfectly in the continuous model when realistic Hill coefficients are used. We investigate to what extent this high correspondence between Boolean and continuous trajectories depends on the extent of reliability of the Boolean trajectories, and we identify simple criteria that enable the faithful reproduction of the Boolean dynamics in the continuous description.

<sup>3</sup> Author to whom any correspondence should be addressed.



Content from this work may be used under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike 3.0 licence](https://creativecommons.org/licenses/by-nc-sa/3.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

**Contents**

<b>1. Introduction</b>	<b>2</b>
<b>2. Model</b>	<b>4</b>
2.1. Boolean model and entirely reliable trajectories . . . . .	4
2.2. Translation to continuous dynamics . . . . .	4
2.3. Comparison . . . . .	6
<b>3. Results</b>	<b>8</b>
3.1. Entirely reliable trajectories (Hamming distance $h = 1$ ) . . . . .	8
3.2. ‘Partially reliable’ trajectories (Hamming distance $h > 1$ ) . . . . .	10
<b>4. Discussion and conclusions</b>	<b>13</b>
<b>Acknowledgment</b>	<b>15</b>
<b>References</b>	<b>15</b>

**1. Introduction**

Biological networks display a remarkable degree of robust behavior, the origins of which are not yet fully understood [1]. Despite ubiquitous sources of noise, which tend to drive the system away from deterministic dynamical behavior, many biological networks manage to maintain a predictable behavior under these circumstances, such as a robust cell cycle [2] and circadian clocks [3]. These systems must possess features that guarantee the correct execution of certain dynamical steps in the presence of noise [4, 5]. The general principles that are responsible for these characteristics have not yet been fully identified [6], although dynamical robustness has been studied for a variety of biological systems [7–11]. Most studies focused on two types of perturbations: noise which directly affects the states of individual nodes in the system [12–17], or fluctuations in the update sequence of the nodes [4, 6, 18–20].

Perhaps the simplest approach to tackling this question is to consider minimal, conceptual models which discard certain features of the underlying dynamics. A common approach, especially used for gene regulatory networks, consists in discretizing protein concentrations by using Boolean values and implementing the regulatory dynamics in discrete time steps [2, 21–27]. Such a description is particularly useful when dealing with large networks [28], because it reduces the huge complexity of a continuous system to a problem with a logical structure that is easier to handle and understand. It permits us to study generic features of entire classes of systems [21], or to reproduce biologically meaningful sequences of states in gene regulation networks that must function reliably, such as cell cycle dynamics [2].

Within the Boolean framework, Peixoto and Drossel [4] showed that it is possible to construct networks that have dynamical trajectories that are *entirely reliable* in the presence of perturbations in the update sequence. Such networks produce exactly the same sequence of states independently of the order in which the nodes are updated. This is achieved by designing periodic trajectories where consecutive states differ in the value of only one node, i.e. these states have a Hamming distance  $h = 1$ . Peixoto and Drossel constructed ensembles of minimal networks which reproduce predetermined reliable trajectories, and analyzed their topologies and update functions.

Although it is quite suggestive of a general underlying principle which guarantees robustness, this reliability criterion is strongly tied to the Boolean description of the system. In contrast, real biological systems are more realistically described via *continuous* variables representing the concentrations of the participating macromolecules, e.g. mRNA and proteins. Although much less analyzed than discrete models [29], ordinary differential equations (ODEs)—which model the switch-like dynamics of genes by using sigmoidal regulatory functions—can include more detailed information about transcription and translation processes. Depending on the parameter values, such models can show oscillating behavior or stable fixed points. These dynamical patterns may or may not correspond qualitatively to their Boolean counterparts, and so far little is known about the conditions under which this may happen [30]. Indeed, even for small systems of only two genes there seems to be no simple relation between Boolean and continuous models [31]. It is therefore unclear to what extent Boolean reliability criteria do apply to the continuous variants.

Glass *et al* [19] investigated the behavior of small networks (with  $N = 5$  nodes at most) possessing entirely reliable dynamics based on results and methods to identify networks displaying stable robust oscillations developed in the 1970s [32]. They indeed found that it is possible that both the Boolean and continuous descriptions exhibit compatible dynamics. Perkins *et al* [1] analyzed the entirely reliable trajectory of the yeast cell cycle and constructed a minimal network that contains almost all experimentally confirmed links. They found that by approximating the Boolean functions by appropriate continuous nonlinear sigmoidal functions, the essential features of the dynamics are conserved. However, compared to Peixoto and Drossel [4], they imposed severe restrictions on the dynamics of the networks and the basin of attraction in addition to the reliability condition, by prohibiting self-inputs and by requiring that states that differ by one bit from a state on the trajectory must lead directly to the trajectory. In this paper, we analyze the continuous dynamics of networks that have entirely reliable Boolean trajectories. We translate the Boolean dynamics into continuous ODEs for the concentrations of mRNA and proteins using sigmoidal Hill functions, and we analyze the resulting time series. We do not analyze specific individual networks, but instead investigate an ensemble of thousands of networks with different sizes and reliable trajectories with different lengths. For this type of trajectory, we find perfect agreement between the Boolean and the continuous dynamics for networks with up to 50 nodes, with biologically reasonable values of the Hill coefficients. Hill coefficients are usually assumed to be between 1 and 4 [33], but larger values are also observed, for instance for *Escherichia coli*, where values of the Hill coefficient up to  $n \approx 11$  were reported [34]. In order to investigate how the degree of agreement depends on the reliability criterion, we compare our results to state sequences with larger average Hamming distances  $h$  between consecutive states, and we indeed find that increasing the Hamming distance destroys the good agreement. This demonstrates that the Boolean criterion for reliability is, in fact, a criterion for agreement between continuous and Boolean dynamics. This paper is structured as follows. In section 2, we introduce the modeling of gene regulatory networks in the Boolean case in general, and reliable trajectories in particular, and explain the translation to continuous dynamics (section 2.2). In section 3.1, the results of the comparison between Boolean and continuous dynamics for entirely reliable trajectories with Hamming distance  $h = 1$  are presented. We then investigate trajectories with larger Hamming distances  $h$  between consecutive states in section 3.2. Finally, in section 4 we discuss and compare our findings with previous studies.

## 2. Model

### 2.1. Boolean model and entirely reliable trajectories

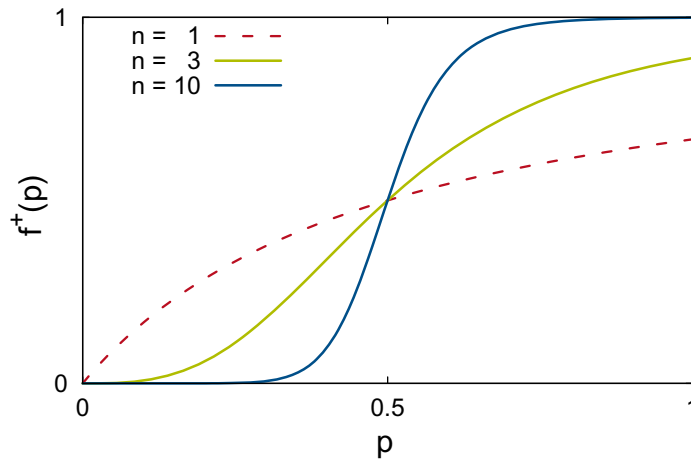
In Boolean networks, the state of each node or ‘gene’ is either ‘on’ or ‘off’ [21, 22]. Each node receives input from other nodes. The Boolean variables  $\sigma \in \{1, 0\}^N$ , which describe the states of the  $N$  nodes, are subject to a dynamical update rule,  $\sigma_i(t+1) = F_i(\sigma(t))u_i(t) + \sigma_i(t)(1 - u_i(t))$ , where  $F_i$  is the update function assigned to node  $i$ , which depends exclusively on the states of its inputs. The binary vector  $u(t)$  represents the update schedule and has components  $u_i(t) = 1$  if node  $i$  is updated at time  $t$ , and  $u_i(t) = 0$  otherwise. The update schedule can be chosen, in general, in three different ways: (a) synchronous (parallel update), where all  $u_i(t) = 1$ , i.e. all nodes are updated simultaneously at every time step; (b) asynchronous and deterministic, where, for instance, each node  $i$  is updated at time intervals  $t_i$ , starting at a time  $t_i^0$ ; the period  $t_i$  and phase  $t_i^0$  can be different for different nodes; and (c) asynchronous and stochastic, where at each time step the node that is updated is chosen at random.

Networks can respond in different ways to stochasticity in the update sequence: the system may lose memory of its past states, or the system may have some degree of order and stay in the vicinity of a well-defined trajectory in state space despite these perturbations. In particular, one may require that the system possesses an entirely reliable dynamics where the sequence of states is always the same, no matter in which order the nodes are updated [4]. Consecutive states of such a reliable attractor trajectory differ only in the value of one node, which means a Hamming distance  $h = 1$ . In state space, such attractors are represented as simple loops even under stochastic update. The number of different states on the attractor, i.e. the length of this trajectory, is  $L = \sum_i l_i$ , where  $l_i$  is the number of times node  $i$  changes its state during a full period. If we require that each node is updated at least once, the shortest trajectory length is  $L = 2N$  with  $l_i = 2$  for each node, and the longest trajectory length is  $L = 2^N$  and comprises all states of the network.

Peixoto and Drossel [4] obtained networks with entirely reliable trajectories by solving the inverse problem, i.e. by first defining the reliable trajectory and afterwards determining the topology and the update functions of the corresponding network. Briefly, the trajectories were generated by performing self-avoiding and closed random walks on the Hamming hypercube, where all nodes flip an even number of times. The minimal networks were obtained by searching through all possible input combinations for each node, until one with the minimal number of inputs is found (if more than one combination is possible, one is randomly chosen). Furthermore, since the trajectory does not necessarily fully specify the complete truth table of all the functions, an additional minimality condition was imposed, where the unspecified output states in the truth table correspond to the majority of the specified ones (or a uniform random value if there is no majority). The ensemble of minimal networks generated in this way is characterized by two parameters: the number of nodes  $N$  and the trajectory length  $L$ .

### 2.2. Translation to continuous dynamics

In the more detailed continuous model, the number of variables becomes  $2N$ , where  $N$  is the number of genes, as the concentrations of mRNA and proteins of one gene are considered separately. The protein is generated from the mRNA through translation with a rate that depends linearly on the mRNA concentration. The degradation of mRNA and proteins is supposed to occur with a constant rate as well. Transcription of mRNA is regulated by the proteins of those



**Figure 1.** Activating the Hill function for different values of the Hill coefficient  $n$  and  $\theta = 0.5$ .

genes that act as the input in the Boolean model. These proteins can have an activating or inhibiting effect. The ODE system describing the reaction kinetics can be written as [35]

$$\begin{aligned} \dot{r}_i &= \epsilon(m_i F_i - r_i), \\ \dot{p}_i &= r_i - p_i, \end{aligned} \quad (1)$$

where the concentration of mRNA produced by gene  $i$  is denoted by  $r_i$  and the corresponding protein concentration is denoted by  $p_i$ , for  $i = \{0, \dots, N\}$ .  $m_i$  is the maximum transcription rate and  $\epsilon$  determines the time scale ratio between mRNA and protein dynamics.

The regulatory functions  $F_i$  depend on the concentrations of the respective regulatory proteins. Regulation by only one protein is usually modelled by a monotonically increasing sigmoidal function when the protein is an activator and by a decreasing function when the protein is an inhibitor. Experimental evidence [36] suggests the usage of *Hill functions* [37], which are therefore commonly employed in models of gene regulation. The Hill function for activation is

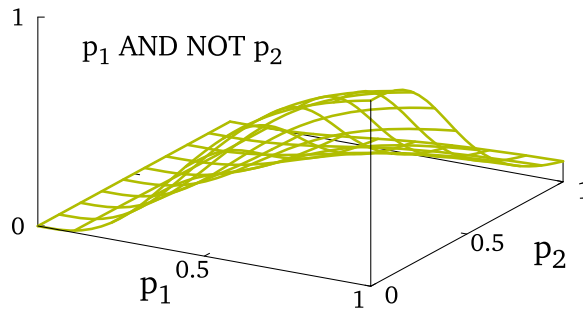
$$f^+(p_i, k_i, n_i) = \frac{p_i^{n_i}}{p_i^{n_i} + k_i^{n_i}}, \quad (2)$$

and the Hill function for inhibition is

$$f^-(p_i, k_i, n_i) = 1 - f^+(p_i, k_i, n_i) = \frac{k_i^{n_i}}{k_i^{n_i} + p_i^{n_i}}, \quad (3)$$

where  $k_i$  is the *activation coefficient* or *expression threshold* that defines the concentration of protein  $i$  needed to significantly activate expression. The parameter  $n_i$ , called the *Hill coefficient*, controls the steepness of the Hill function. The larger the  $n_i$ , the more step-like is the regulatory function (see figure 1). Biologically,  $n_i$  reflects the molecular binding mechanism: it describes the number of proteins that are required for saturation of binding to the DNA [38] and is therefore also called the *cooperativity coefficient*. The Hill function can be considered as the probability that the promoter region is bound, averaged over many binding and release events of proteins  $i$  [33].

The combined effect of multiple transcription factors is described by using multi-dimensional input functions. An example for deriving such a function from gene expression



**Figure 2.** HillCube of the input function  $F = p_1$  AND NOT  $p_2$ .

measurements was given by Setty *et al* [39], who used the *lacZYA* operon of *E. coli*. The function  $F_i(p_1, p_2)$  can, for instance, integrate an activator  $p_1$  and a repressor  $p_2$  [33]. If the activator and the inhibitor bind to the promoter independently, there are four binding states of the promoter: unbound, bound to either  $p_1$  or  $p_2$ , or bound to both proteins. Transcription may occur mainly in the case where the activator binds the promoter and the repressor does not, resulting in a  $p_1$  AND NOT  $p_2$  input function. Already in 1978, Glass and Pasternack [40] translated multi-dimensional Boolean functions into continuous analogues and nearly 30 years later, Wittmann *et al* [41] introduced a general method for constructing continuous counterparts of Boolean functions. They defined the so-called HillCubes, by taking sums over the appropriate products of Hill functions with a weight 1 or 0, according to the output value of the Boolean function. For instance, the AND NOT function (see figure 2) is written as

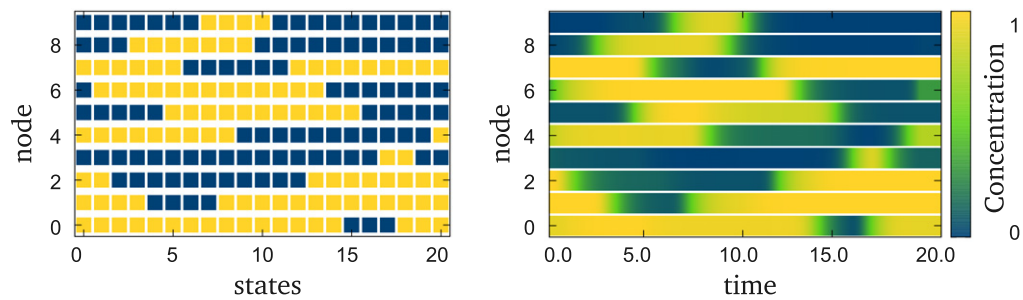
$$F_i^{\text{AND NOT}}(p_1, p_2) = 0 (f^+(p_1) \cdot f^+(p_2)) + 1 (f^+(p_1) \cdot f^-(p_2)) + 0 (f^-(p_1) \cdot f^+(p_2)) \\ + 0 (f^-(p_1) \cdot f^-(p_2)).$$

In the following, we compare the reliable dynamics of minimal Boolean networks [4] with the dynamics of their continuous counterparts. For this purpose, we translate a Boolean network into a continuous model by constructing the associated ODEs (1) for the mRNA and proteins using HillCubes for the update functions  $F_i$ , as explained above. These ODEs then can be solved, e.g., with the Runge–Kutta method, and the time series can be analyzed and compared with the Boolean trajectory. In the simulations, for which the results are shown in the following section, the parameters were set to  $m_i = \epsilon_i = 1$  and  $k_i = 0.5$ .

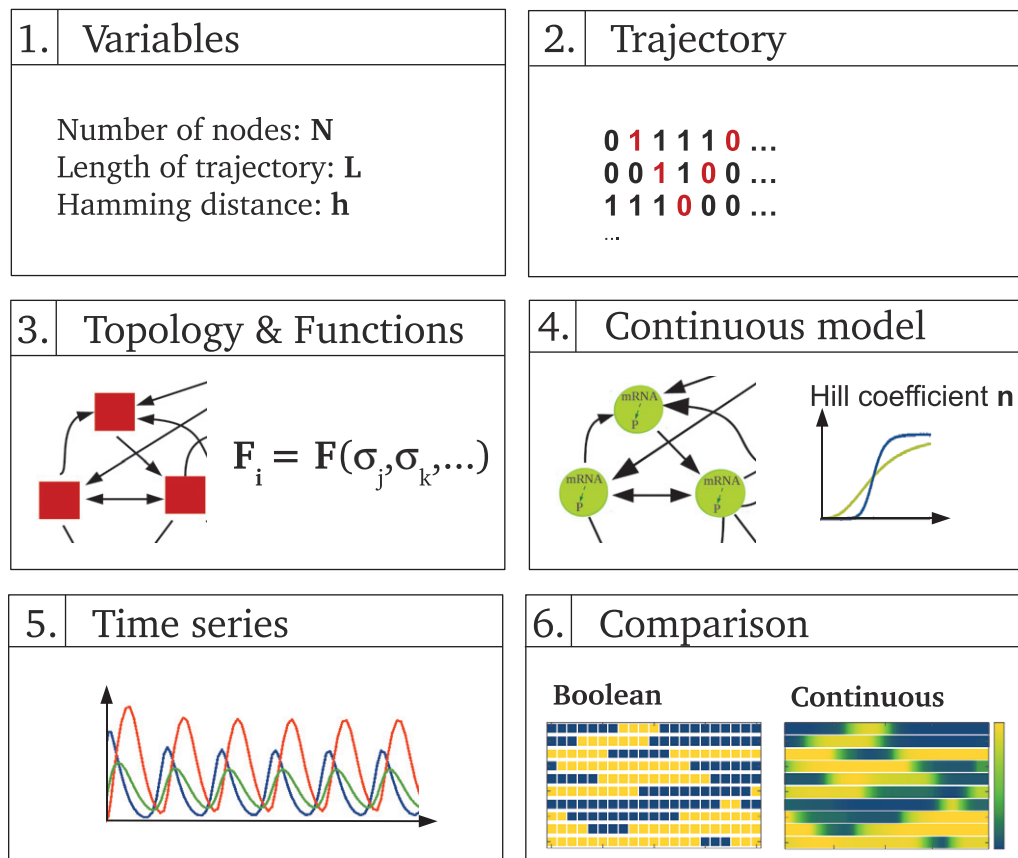
### 2.3. Comparison

As an illustration, we first present one example of a network with an entirely reliable trajectory and visually compare it with the time series of the continuous version of the network. In figure 3, on the right, the time series of the concentrations of the nodes in the continuous system are plotted versus time. The different colors refer to the concentration values from 0.0 (blue) through 0.5 (green) to 1.0 (yellow). This can be compared with the Boolean state sequence (figure 3, left), which is discrete in time and only has two states for the concentration: 0 and 1 (blue and yellow, respectively). Figure 3 shows very good agreement of the different states of the trajectory after a short transient time in the continuous case.

In order to systematically make this comparison for different values of Hill coefficients and for thousands of networks, we applied the algorithm described in the following.



**Figure 3.** Comparison of the Boolean reliable trajectory with  $L = 20$  states of a network with  $N = 10$  nodes (left) with the oscillations of the continuous model with the Hill coefficient  $n = 4$  (right).



**Figure 4.** Different steps of the procedure to compare the Boolean and continuous dynamics.

For each combination of  $N$  (node number) and  $L$  (trajectory length), we first generated an ensemble of Boolean reliable trajectories. The topology and the Boolean update functions of the minimal networks that display these trajectories were determined according to the algorithm described in [4] (compare figure 4, steps 1–3). The minimal Boolean networks were translated into continuous ODEs using Hill functions and HillCubes, and a simulation was run starting with a large Hill coefficient  $n = 10$ . As the initial condition we chose a Boolean state of the

reliable trajectory, i.e. the concentration values of mRNA and proteins are 0 or 1, depending on the chosen initial state. Then we run the system of ODEs (compare figure 4, steps 4–5).

After a transient time, we translated the continuous dynamics into Boolean states in two steps: firstly, we searched for the maximum and minimum values of each node and chose the midpoint to be the individual threshold for this node. Secondly, we translated the entire continuous time series back into a logical state sequence by scanning it in small time steps. We defined a value above 1.1 times the threshold value to be ‘on’ (1), and a value below 0.9 times the threshold value to be ‘off’ (0). If the concentration value is in between, the node retains its state. Introducing in this way a finite width of the threshold prevents minor oscillations around the threshold being counted as a switching event. The second step of the translation stops when the first network state of the sequence is reached again. In this way, a closed periodic trajectory can be found. In the time series obtained by this procedure, each state occurs many times in a row until it is replaced by a state that differs in one node. We eliminated these repetitions of the same Boolean state from the final trajectory, thus discarding the information on how long a certain state appears in the continuous model, but keeping the information about the sequence of states. The sequence of states was then compared with the Boolean trajectory, which should be identical if the continuous model is able to reproduce the Boolean dynamics (and vice versa). For this comparison, firstly we checked the length of the trajectory, i.e. the number of states of the attractor, secondly the states themselves and thirdly the order in which they occur. Only if all these criteria are fulfilled, we consider the Boolean and continuous dynamics to be in agreement. We emphasize that if a trajectory is only partially in agreement, we count this as a failure.

Next, we decreased the value of the Hill coefficient in steps of size 0.2. As the initial condition for the continuous simulation we chose a state of the trajectory calculated with the previous value of  $n$ , because the Boolean values of the concentrations, 0 and 1, are not close to actual concentration values on attractors when  $n$  is not large. Using this procedure, the simulation can be started in the basin of attraction of the reliable trajectory if it is still an attractor and if its basin is not too small. This permitted us to explore the dynamics of the continuous model version for values of the Hill coefficient  $n \in \{0, \dots, 10\}$ .

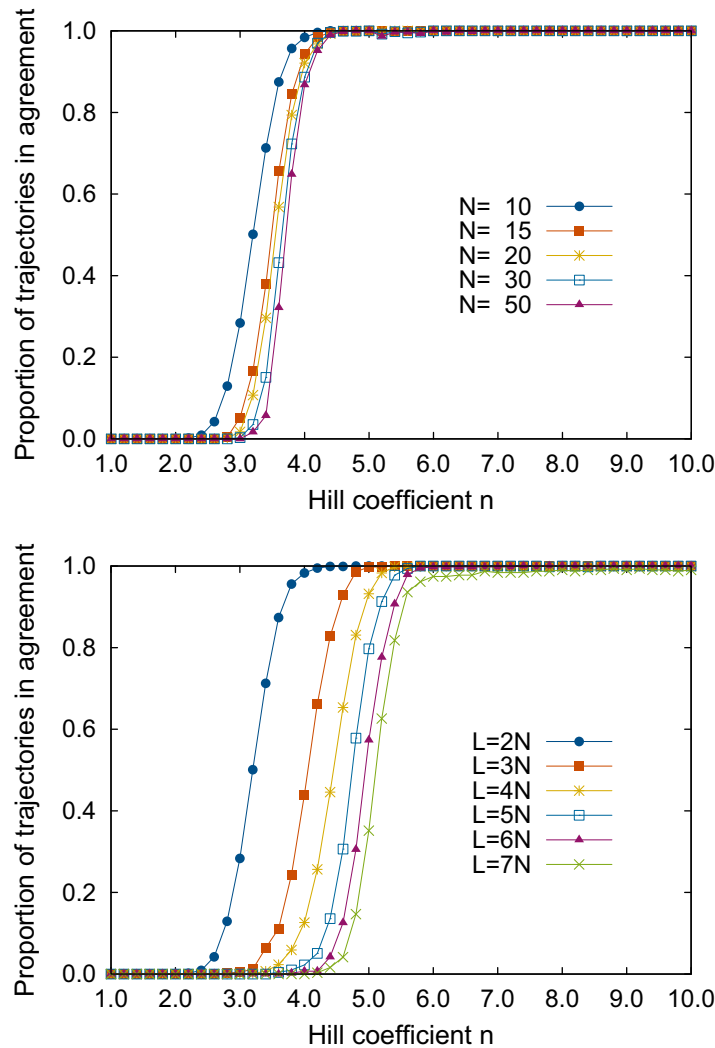
We repeated this procedure with thousands of reliable trajectories and the associated minimal networks and evaluated the number of instances for which the Boolean and the continuous case agree, as a function of the Hill coefficient  $n$ , the number of nodes in the network  $N$  and the length of the trajectory  $L$ .

### 3. Results

#### 3.1. Entirely reliable trajectories (Hamming distance $h = 1$ )

Figure 5 shows the proportion of networks for which the reliable trajectory of the Boolean and the continuous model are in agreement, depending on the Hill coefficient  $n$ , for different sizes of the networks (top) and different lengths of the trajectories (bottom). Nearly all trajectories with Hamming distance  $h = 1$  are in perfect agreement with the continuous trajectory for systems with the Hill coefficient around  $n = 4.5$ . This good agreement occurs both for small networks with only 10 nodes and  $L = 20$ , i.e. trajectories with 20 states, as well as large networks with up to 50 nodes. For  $n = 3$ , around 30% of the networks with  $N = 10$  reproduce the correct trajectory, and 5% of the networks with  $N = 15$  (all with  $L = 2N$ ), while for larger networks this fraction becomes vanishingly small. We additionally analyzed networks with longer





**Figure 5.** Proportion of trajectories in agreement between Boolean and continuous dynamics depending on the Hill coefficient  $n$  for different numbers of nodes  $N$  in the network, with  $L = 2N$  (top) and different lengths  $L$  of the trajectories, with  $N = 10$  (bottom).

trajectories: here, the proportion of trajectories with perfect agreement is close to 100% with Hill coefficients between  $n = 4$  and 6.5 for  $N = 10$  and state sequences with a length between  $L = 2N$  and  $7N$ . When the continuous trajectory does not agree with the Boolean one, although the Hill coefficients are quite high, the disagreement is due to the occurrence of Boolean states in the wrong order. This often happens because a given node switches before its target node, the concentration of which is still too far away from its threshold. These mistakes disappear with larger Hill coefficients, as the concentrations pass the threshold much faster in this case. For small Hill coefficients, the dynamics of the systems rather end up at fixed points.

In the simulations underlying figure 5, the time scales of the mRNA and protein dynamics were chosen to be equal ( $\epsilon = 1$  in equation (1)). The case where the mRNA dynamics are much faster ( $\epsilon \approx 100$ ) effectively reduces the system to  $N$  equations instead of  $2N$ , since the mRNA

concentrations always have enough time to reach the quasi-steady state that corresponds to the present protein concentrations. Because this reduces the delay between subsequent changes of the states of nodes, larger values of the Hill coefficient  $n$  are required for perfect agreement of the continuous oscillations with the Boolean trajectories (results not shown). Additionally, the maximum transcription rate  $m_i$  was varied in the range 1–5 but no significant difference was perceived, as only the maxima and the threshold values changed. Thus, the variation of  $m_i$  only changes the range of the concentrations.

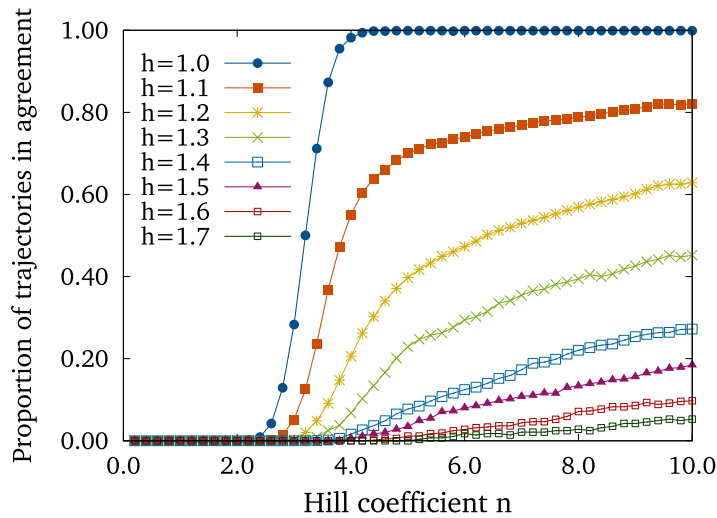
We expect that the Hill coefficient needed to reproduce most reliable trajectories in the continuous model increases when either  $N$  or  $L$  becomes even larger, because the likelihood of errors increases with the number of states in the trajectory. Thus, the Hill coefficient, needed in order to obtain good agreement between Boolean and continuous dynamics, increases for larger networks. This is in contrast to the results for simple feedback loops, where the Hopf bifurcation takes place at lower values of the Hill coefficient for higher dimensions [19]. The reason for this is, on the one hand, the less defined state space and, on the other hand, the fact that in simple feedback loops the time between two changes of the node states increases when there are more nodes in the system. Thereby, the Hill function may be less steep, i.e. the Hill coefficient may be smaller. This relationship between the time between two flips of one node and the size of the network is not given for the more complex networks investigated here.

### 3.2. ‘Partially reliable’ trajectories (Hamming distance $h > 1$ )

In order to investigate whether this good agreement between the Boolean and the continuous dynamics is indeed due to the Hamming distance  $h = 1$  between consecutive states, we increased the average Hamming distance  $h$  of the trajectories in small steps, resulting in ‘partially reliable’ trajectories. The value  $h$  is defined as the sum of the individual Hamming distances  $h$  between two subsequent states (under parallel update), divided by the length of the trajectory  $L$ , i.e. the number of states. Again, we generated an ensemble of trajectories and the corresponding minimal networks, using the same algorithm as before. We investigated again the continuous model version of these networks and evaluated the proportion of trajectories in agreement between the Boolean and the continuous model, for different sizes of the networks and trajectory lengths. An average Hamming distance  $h > 1$  means that in some state transitions more than one node flip at the same time in the Boolean model. In the continuous model, this cannot happen, and we therefore checked whether those nodes that flip at the same time in the Boolean system with parallel update cross their thresholds directly one after the other in the continuous case. The order in which these nodes change their state in the continuous model does not matter, because any order can occur in the Boolean model when stochastic update is used.

Figure 6 shows the results of this investigation. Again, the proportion of networks that reproduce the Boolean trajectory in the continuous model version increases with  $n$ ; however, it appears to approach an asymptotic value that is smaller than 1. This asymptotic value decreases with increasing  $h$ , and also with increasing  $N$  or  $L$  (not shown).

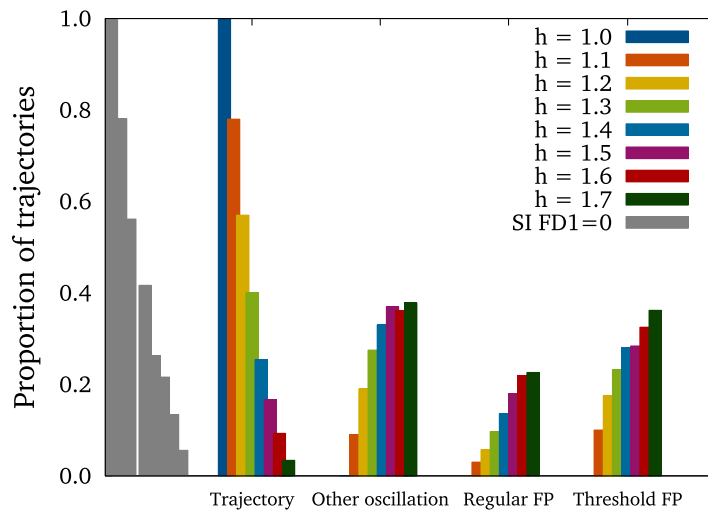
In the following, we investigate the reasons why part of the trajectories with  $h > 1$  cannot be reproduced in the continuous model even for large  $n$ . When observing the failure networks in detail, we found two classes. (i) The continuous model shows oscillations that include states that do not occur in the Boolean trajectory or that are in the wrong order. Such differences are small and might even vanish if the parameters of the continuous model version were chosen



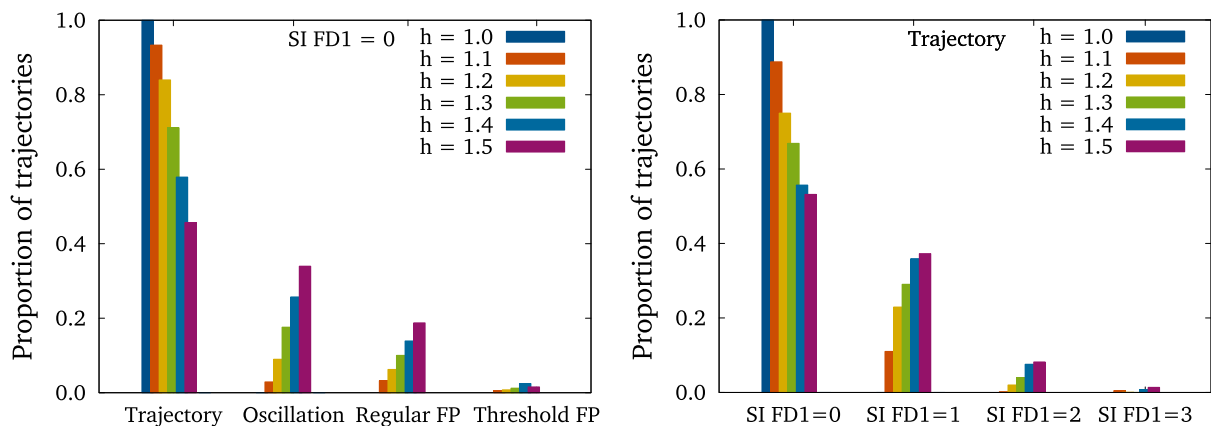
**Figure 6.** Proportion of trajectories in agreement between the Boolean and continuous dynamics when the average Hamming distance  $h$  is varied (for networks with  $N = 10$  and  $L = 20$ ).

differently. (ii) The continuous model does not show oscillations at all, and the dynamics end up on a fixed point. These fixed points can be further divided into two classes: ‘regular’ fixed points, with concentration values near 0 and 1 for sufficiently large values of the Hill coefficient  $n$ , and fixed points where one or more nodes have concentration values near their threshold. In previous studies [42–45], these were referred to as singular steady states; we will refer to them as threshold fixed points. We analyzed in which cases these threshold fixed points occur and found that one criterion is especially relevant, which affects both the topology of the networks and their dynamic properties: whenever there is a node that flips twice in two consecutive time steps (*dynamics*) due to a self-input (*topology*), the continuous dynamics might not be able to generate a sufficient degree of activation and inhibition, and therefore the concentration is trapped at values near its threshold. We checked the hypothesis that short flip durations due to self-inputs may cause disagreement between the Boolean and the continuous systems by evaluating the proportion of networks with each of the four mentioned types of dynamical behavior. Figure 7 shows for  $n = 10$  and for different average Hamming distances  $h$  the proportions of networks with  $N = 10$  and  $L = 20$  that reproduce the Boolean trajectory, that show a different oscillation, that have a regular fixed point, or a threshold fixed point. The gray bars on the left additionally show the proportion of networks that do not have a node with a state duration of only one time step due to a self-input in the Boolean model. This information is not based on simulations, but only depends on the Boolean trajectory and the resulting network structure. Interestingly, the proportion of networks where no node shows a flip duration of only one time step coupled with a self-input (referred to as (SI\_FD1 = 0)) is very similar to the proportion of networks showing agreement with the Boolean trajectory.

All networks with trajectories with Hamming distances  $h = 1.0$  have SI\_FD1 = 0, because the network would return to the same state if the same node flipped twice in two consecutive time steps. Only when more than one node flips per time step can the system be in a different state after the same node flipped twice in a row.



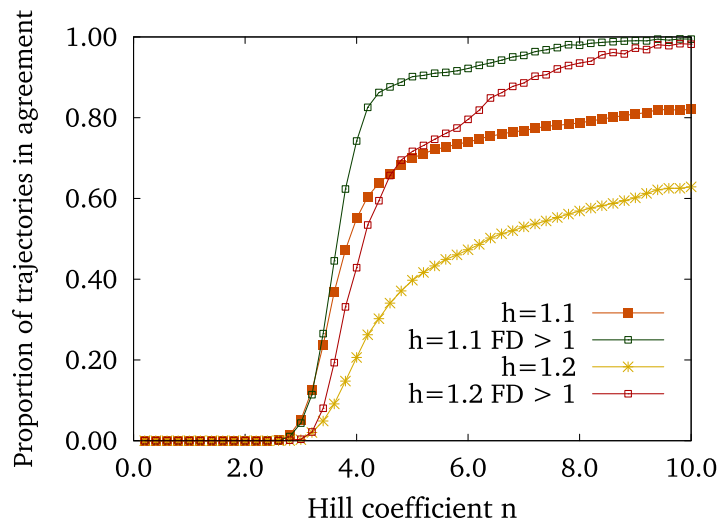
**Figure 7.** Proportion of networks falling into the four distinct groups—trajectory reproduced, other oscillation, regular fixed point and threshold fixed point—for simulations performed with  $n = 10$ . The gray bars on the left indicate how many of the networks in the Boolean model do not have nodes with a state duration of only one time step due to a self-input.



**Figure 8.** Proportion of networks showing the different classes while  $SI\_FD1 = 0$  (left) and the proportion of networks showing different values of  $SI\_FD1$  while reproducing the trajectory (right). Parameters are  $N = 10$ ,  $L = 20$  and  $n = 10$ .

Figure 8 shows that for  $N = 10$  the proportion of networks that reproduce the Boolean trajectory increases if only those networks are considered that have no node with a self-input coupled with a flip duration of one time step ( $SI\_FD1 = 0$ ). For example, for an average Hamming distance  $h = 1.1$  the proportion of trajectories in agreement increases from 78 to 93% and for  $h = 1.4$  the proportion increases from 25 to 51%. Those networks with  $SI\_FD1 = 0$  that do not reproduce the Boolean trajectory have other oscillations or go to a fixed point, but the number of networks showing threshold fixed points is reduced noticeably.

On the other hand, among the networks that reproduce the Boolean trajectory, 10–40% for average Hamming distances from  $h = 1.1$  to 1.5 do have a node that flips twice in a row due



**Figure 9.** If the duration of nodes' states is limited, that they must hold at least for two time steps, it cannot happen that a node flips due to a self-input twice in a row. For average Hamming distances  $h = 1.1$  or  $1.2$ , the proportion of networks showing agreement between the Boolean and continuous trajectories is increased obviously.

to a self-input (see the right column of figure 8). We observed similar trends for larger  $N$  (not shown).

One can conclude that a general prohibition for a node to change its state twice within two time steps increases the proportion of reproduced trajectories significantly for average Hamming distances  $h > 1$ . A 'time step' in this case relates to a change of the state in the Boolean network with a parallel update. We tested this assumption for networks with  $N = 10$ ,  $L = 20$ ,  $h = 1.1$  and  $1.2$  and found that with this restriction on the trajectories the plateau does not exist (see figure 9). Instead we found oscillations corresponding to the Boolean trajectories for all networks, as long as the Hill coefficient  $n$  was large enough ( $n \approx 10$ ). The proportion of networks is even larger than suggested by figure 8 because now not only short flip durations due to a self-input cannot occur, but all cases where a node changes its state twice within two subsequent updates are prohibited. We tested furthermore whether the Hill coefficient needed to show agreement between the Boolean and continuous dynamics can be decreased with an even more stringent restriction on the flip duration to at least three time steps, but no better results were obtained.

#### 4. Discussion and conclusions

We showed that for networks possessing 'entirely reliable trajectories', i.e. trajectories with Hamming distance  $h = 1$  between consecutive states, there is very good agreement between Boolean and continuous dynamics. We tested the importance of Hamming distance 1 by investigating trajectories with average Hamming distances  $h > 1$  and found that for those 'partially reliable trajectories' the degree of reproducibility decreases. The fact that in the continuous model version both the concentrations and the time are continuous leads to interesting phenomena: like other authors [42–45], we found in the continuous system fixed

points with concentrations near the threshold. The occurrence of these fixed points can be prevented if the trajectories are subject to certain restrictions: the trajectory may not include nodes that—in the Boolean representation—flip back and forth between two states in successive time steps. With this limitation, the Boolean and continuous systems fit together surprisingly well. In the special case with  $h = 1$  the excellent agreement was mostly due to not only the fact that these flips back and forth are impossible, but also that on average the nodes stay a longer time in the same state if only one node can flip per step in the Boolean version.

Trajectories with Hamming distance  $h = 1$  have been investigated before by Wilds and Glass [19], who also note their good reproducibility with continuous equations. In contrast to our work they examined only a few networks with up to five nodes, and included more restrictions on the networks, as self-inputs are forbidden. From our work, we can conclude that self-inputs are not necessarily harmful, which is compatible with the fact that they are observed in real systems. Wilds and Glass [19] applied an additional restriction that neighboring states must lead to the trajectory, i.e. they belong to the basin of attraction of the given trajectory. This restriction does not apply to the trajectories of the networks analyzed in our work, and thereby a larger set of possible trajectories was considered. Nevertheless, we found this restriction not to be relevant for the comparison of Boolean and continuous dynamics if the initial state is chosen to be part of the trajectory.

Inspired by the work of Wilds and Glass, we performed an additional study in order to analyze the impact of the basin of attraction. For networks with five nodes and entirely reliable trajectories with 20 states, we compared two extreme cases: on the one hand, minimal networks, which have the smallest possible number of inputs per node and the simplest functions possible, and on the other hand, fully connected networks, where the state space is completely determined and each state leads to the trajectory, so that there is only one attractor, i.e. the entirely reliable trajectory. However, this precise prescription of the state space did not decrease the Hill coefficient required for 100% agreement between Boolean and continuous dynamics in this network ensemble. The reason for this is that for Hamming distances  $h = 1$  there are no intermediate states needed in the continuous model to reproduce the trajectory; thus no intermediate states can coincide with regular fixed points of the Boolean network, which would then trap the dynamics of the continuous system. Therefore no improvement of the reproducibility with continuous dynamics can be observed when the state space is further specified. It is certainly worthwhile to further investigate the difference between the minimal and fully connected networks for larger Hamming distances, because in this case the state space could be defined such that those states lead to the trajectory, which are needed as intermediate states in the continuous model to overcome Hamming distances  $h > 1$ . This detailed prescription of the state space could presumably prevent the continuous trajectory from ending up on regular fixed points and therefore increase the proportion of trajectories in agreement between the Boolean and the continuous model version.

It is interesting to compare our results for general networks with those found by Braunewell and Bornholdt [6], who studied the network of the yeast cell cycle, which has properties similar to the networks investigated by us: the cell cycle network contains 11 nodes, and the trajectory has 13 states. In contrast to us, they do not consider the correlation between Boolean and continuous dynamics, but the robustness of the trajectory to fluctuations in the processing times of protein concentration build-up and decay. They check whether the system reproduces the

sequence of states under fluctuations in the update time and find that its very robust dynamic is based on three criteria: the occurrence of catcher states that have a Hamming distance  $h = 1$  to the next state, a large basin of attraction and a sufficiently long duration of the state of a node before it flips again.

In our study, the importance of catcher states for robust dynamics shows in the fact that networks with many catcher states correspond to trajectories with low average Hamming distances  $h$ . The criterion that a node must stay in the same state for a sufficiently long time agrees with our finding that the proportion of networks that reproduce the Boolean trajectory is much larger when there are no nodes that change their state in two consecutive time steps due to a self-input. Sevim *et al* [5] point out the importance of filtering out quick back-and-forth flips, as well, and justify this restriction with the fact that those short flip durations would represent a very rapid build-up and decay of transcription products, which is not realistic.

According to our results, the dynamics between the Boolean and continuous models are in agreement for those networks that are robust against fluctuations in the update sequence. This means that for systems such as the cell cycle, which need to be robust against noise, the simplified Boolean approach is a faithful, albeit conceptual, representation of the real biological system.

## Acknowledgment

EA and BD acknowledge support from the Deutsche Forschungsgemeinschaft under grant number Dr300/4-2.

## References

- [1] Perkins T J, Wilds R and Glass L 2010 Robust dynamics in minimal hybrid models of genetic networks *Phil. Trans. R. Soc. A* **368** 4961–75
- [2] Li F *et al* 2004 The yeast cell-cycle network is robustly designed *Proc. Natl Acad. Sci. USA* **101** 4781
- [3] Ueda H R, Hagiwara M and Kitano H 2001 Robust oscillations within the interlocked feedback model of *Drosophila* circadian rhythm *J. Theor. Biol.* **210** 401–6
- [4] Peixoto T P and Drossel B 2009 Boolean networks with reliable dynamics *Phys. Rev. E* **80** 056102
- [5] Sevim V, Gong X and Socolar J E S 2010 Reliability of transcriptional cycles and the yeast cell-cycle oscillator *PLoS Comput. Biol.* **6** e1000842
- [6] Braunewell S and Bornholdt S 2007 Superstability of the yeast cell-cycle dynamics: ensuring causality in the presence of biochemical stochasticity *J. Theor. Biol.* **245** 638–43
- [7] Albert R and Othmer H G 2003 The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in *Drosophila melanogaster* *J. Theor. Biol.* **223** 1–18
- [8] Chaves M, Albert R and Sontag E D 2005 Robustness and fragility of Boolean models for genetic regulatory networks *J. Theor. Biol.* **235** 431–49
- [9] Wagner A 2005 Circuit topology and the evolution of robustness in two-gene circadian oscillators *Proc. Natl Acad. Sci. USA* **102** 11775–80
- [10] Ma W, Lai L, Ouyang Q and Tang C 2006 Robustness and modular design of the *Drosophila* segment polarity network *Mol. Syst. Biol.* **2** 70
- [11] Trane C and Jacobsen E W 2008 Network structure and robustness of intracellular oscillators *Proc. 17th Int. Federation of Automatic Control World Congress (Seoul, Korea, 6–11 July 2008)* p 10989
- [12] Aldana M and Cluzel P 2003 A natural class of robust networks *Proc. Natl Acad. Sci. USA* **100** 8710–4

- [13] Peixoto T P and Drossel B 2009 Noise in random Boolean networks *Phys. Rev. E* **79** 036108
- [14] Peixoto T P 2010 Redundancy and error resilience in Boolean networks *Phys. Rev. Lett.* **104** 048701
- [15] Ghanbernejad F and Klemm K 2011 Stability of Boolean and continuous dynamics *Phys. Rev. Lett.* **107** 188701
- [16] Peixoto T P 2012 The behavior of noise-resilient Boolean networks with diverse topologies *J. Stat. Mech.* **2012** P01006
- [17] Peixoto T P 2012 Emergence of robustness against noise: a structural phase transition in evolved models of gene regulatory networks *Phys. Rev. E* **85** 041908
- [18] Klemm K and Bornholdt S 2005 Topology of biological networks and reliability of information processing *Proc. Natl Acad. Sci. USA* **102** 18414–9
- [19] Wilds R and Glass L 2009 An atlas of robust, stable, high-dimensional limit cycles *Int. J. Bifurcation Chaos* **19** 4055–96
- [20] Aracena J, Goles E, Moreira A and Salinas L 2009 On the robustness of update schedules in boolean networks *Biosystems* **97** 1–8
- [21] Kauffman S A 1969 Metabolic stability and epigenesis in randomly constructed genetic nets *J. Theor. Biol.* **22** 437–67
- [22] Thomas R 1973 Boolean formalization of genetic control circuits *J. Theor. Biol.* **42** 563–85
- [23] Somogyi R, Fuhrman S, Askenazi M and Wuensche A 1997 The gene expression matrix: towards the extraction of genetic network architectures *Nonlinear Anal. Theory, Methods Appl.* **30** 1815–24
- [24] Mendoza L, Thieffry D and Alvarez-Buylla E R 1999 Genetic control of flower morphogenesis in *Arabidopsis thaliana*: a logical analysis *Bioinformatics* **15** 593–606
- [25] Albert R and Barabasi A-L 2002 Statistical mechanics of complex networks *Rev. Mod. Phys.* **74** 47–97
- [26] Fauré A, Naldi A, Chaouiya C and Thieffry D 2006 Dynamical analysis of a generic Boolean model for the control of the mammalian cell cycle *Bioinformatics* **22** e124–31
- [27] Giacomantonio C E and Goodhill G J 2010 A Boolean model of the gene regulatory network underlying mammalian cortical area development *PLoS Comput. Biol.* **6** e1000936
- [28] Bornholdt S 2005 Less is more in modeling large genetic networks *Science* **310** 449
- [29] Mochizuki A 2005 An analytical study of the number of steady states in gene regulatory networks *J. Theor. Biol.* **236** 291–310
- [30] Bagley R J and Glass L 1996 Counting and classifying attractors in high dimensional dynamical systems *J. Theor. Biol.* **183** 269
- [31] Gehrmann E and Drossel B 2010 Boolean versus continuous dynamics on simple two-gene modules *Phys. Rev. E* **82** 046120
- [32] Glass L and Pasternack J S 1978 Stable oscillations in mathematical models of biological control systems *J. Math. Biol.* **6** 207–23
- [33] Alon U 2007 *An Introduction to Systems Biology* (London: Chapman and Hall/CRC)
- [34] Sneppen K and Zocchi G 2005 *Physics in Molecular Biology* (Cambridge: Cambridge University Press)
- [35] Polynikis A, Hogan S J and di Bernardo M 2009 Comparing different ODE modelling approaches for gene regulatory networks *J. Theor. Biol.* **261** 511–30
- [36] Yagil G and Yagil E 1971 On the relation between effector concentration and the rate of induced enzyme synthesis *Biophys. J.* **11** 11–27
- [37] Hill A V 1910 The possible effects of the aggregation of the molecules of hemoglobin on its dissociation curves *J. Physiol.* **40** iv–vii
- [38] Widder S, Schicho J and Schuster P 2007 Dynamic patterns of gene regulation. I: simple two-gene systems *J. Theor. Biol.* **246** 395–419
- [39] Setty Y, Mayo A E, Surette M G and Alon U 2003 Detailed map of a cis-regulatory input function *Proc. Natl Acad. Sci. USA* **100** 7702–7
- [40] Glass L and Pasternack J S 1978 Prediction of limit cycles in mathematical models of biological oscillations *Bull. Math. Biol.* **40** 27–44



- [41] Wittmann D, Krumsiek J, Saez-Rodriguez J, Lauffenburger D A, Klamt S and Theis F 2009 Transforming Boolean models to continuous models: methodology and application to t-cell receptor signaling *BMC Syst. Biol.* **3** 98
- [42] El Snoussi H and Thomas R 1993 Logical identification of all steady states: the concept of feedback loop characteristic states *Bull. Math. Biol.* **55** 973–91
- [43] Plahte E and Kjøglum S 2005 Analysis and generic properties of gene regulatory networks with graded response functions *Physica D* **201** 150–76
- [44] Veflingstad S R and Plahte E 2007 Analysis of gene regulatory network models with graded and binary transcriptional responses *Biosystems* **90** 323–39
- [45] Gjuvsland A B, Plahte E and Omholt S W 2007 Threshold-dominated regulation hides genetic variation in gene expression networks *BMC Syst. Biol.* **1** 57