A Stochastic Modeling Framework for Regulatory Networks Applied to the Yeast Cell Cycle

Stefan Braunewell*
Institute for Theoretical Physics, University of Bremen, D-28359 Bremen, Germany

Brian Chapados†
Department of Molecular Biology, The Scripps Research Institute, La Jolla, CA 92037, USA

Sebastian Fallert‡ and Christoph Neugebauer§
Department of Chemistry, University of Cambridge, Cambridge, UK

The interactions between proteins within cells constitute a network of processes that operate on different time scales, allowing living organisms to respond to their environment. To understand the enormous complexity of these systems, a wide variety of computational models have been introduced at various levels of detail. In this project, we develop an analytical framework based on asynchronous Boolean dynamics to investigate the qualitative dynamical properties of regulatory networks.

As opposed to the common synchronous modeling, we assign rates to the different molecular processes involved in gene regulation and investigate the effects of a stochastic updating mechanism in a simple model of the yeast cell cycle. We model the transitions between different states of the Boolean network as independent Poisson processes. Two different updating schemes are proposed and investigated using both a master equation approach and Monte Carlo simulation.

We find that introducing biologically inspired rates increases the probability of the system to reach the biological G_1 stationary state. We investigate which rate parameters most strongly affect the probability of reaching this state and discuss the underlying mechanisms.

I. INTRODUCTION

A. Gene Regulation

Living cells contain a plethora of molecules that interact to produce commonly observed behaviors such as movement towards a food source, nutrient uptake or cell division. Each cell carries all the information necessary to specify its own replication and multiple functions for survival. This information is encoded in genes, which are the physical and functional units of biological inheritance. Although there is no consensus definition for the term "gene", we use it to refer to the information encoded in DNA that is necessary to synthesize a biologically functional molecule.

Specifically, the information in a gene is encoded as a sequence of chemical bases in a DNA polymer. In order to perform a biological function, a gene encoded in DNA must be transcribed to produce an RNA molecule. Depending on its sequence, the RNA molecule either functions on its own, serves as a template for producing proteins, or forms a complex with other RNA or protein molecules (reviewed in [1]). Genes encoded in DNA have no functional consequence until they are expressed as RNA or other functional molecules. Thus, regulation of gene expression or "gene regulation" links genes to cellular functions.

In this study, we define gene regulation as the control of RNA production from DNA via transcription or chromatin modifications. In cells, genes can be either activated (expressed/transcribed) or deactivated (repressed/silenced) by other gene products (RNA or protein). Stimuli or other signals from inside or outside of the cell can affect the activity of RNA and protein molecules. Thus, gene regulation involves the integration of cellular signals, which affect the expression of genes.

Classical genetic studies have mapped genes into pathways or networks, which define the qualitative regulatory relationships among a set of genes. While these models are useful for classifying and characterizing phenotypic observations in terms of absolute gene requirements, they provide very little information about the dynamics of gene regulation. Moreover, since such models are mainly qualitative in nature, it is often difficult to determine the behavior of networks consisting of many interacting genes.

The common approach to quantitative modeling of the dynamics of gene regulatory networks is to use systems of coupled differential equations of RNA and protein concentrations [2]. However, due to the many elements involved and the high level of details described through kinetic reaction parameters, this method poses difficulties for the description of larger systems.

An alternative approach is to use a more abstract level of description and represent the dynamics of a system in terms of discrete activity states (for a review of common models of gene regulation, see [3]). The first discrete models used for biological networks were the so-called Random Boolean Networks (RBN) introduced by Kauff-

*Electronic address: braunewell@itp.uni-bremen.de
†Electronic address: chapbr@scripps.edu
‡Electronic address: sf287@cam.ac.uk
§Electronic address: cjn24@cam.ac.uk

man and Thomas in [4] and [5].

In these models, the system is described by a set of "nodes", which resemble the different types of chemical reactants and (directed) "edges", which define the interactions between two nodes. Time is discrete and all nodes are updated synchronously according to their respective input function, where each input of a node stems from an incoming edge. Every node can take one of two activity states ("on" or "off") and the inputs (connections) as well as the input functions (drawn at random from all possible Boolean functions) are fixed for a given network. The state of each node is then given by the state of this node's input nodes at the previous time step. As this is a deterministic system with a finite state space, for any initial condition of the system the dynamics eventually runs into an attractor, either a fixed point or a limit cycle. It was hypothesized that the attractors of RBNs represent different cell types.

This discrete approach certainly lacks many of the characteristics of gene regulatory networks. However, despite their limitations, such models have been successfully used to describe the outcome of a dynamical process on a gene regulatory network as in fruit fly development [6] or even the dynamical sequence of activity states as in the case of the budding yeast cell cycle [7].

However, the original assumption of deterministic dynamics in these models is clearly implausible considering that the processes driving these systems are of a molecular nature and thus fluctuations are ubiquitous. To address this fact, modified models have been proposed to introduce a stochastic element into the simulations. In particular, it was shown that attractors in the synchronous updating scheme can vanish when asynchronous updating is used [8]. Indeed, it can be expected that biological systems perform their task reliably despite the inherent stochasticity of the molecular processes involved.

In [9] different asynchronous models are described and implications on the outcome of the fruit fly development are discussed. Commonly used asynchronous models are random-order updating and totally asynchronous updating. In both of these models, the order of updates affects which nodes have already changed their state, thus influencing the state of those nodes to update later. In the random-order algorithm, it is ensured that every node updates exactly once during a system update, whereas in the totally asynchronous version, at every time step a node is drawn at random out of all nodes and is updated. Another class of models, which is more biologically inspired and incorporates some aspects of differential equation models, are the so-called Glass-type networks which have also been used in the description of the yeast cell cycle [10].

Here, we introduce a different modeling framework, which allows the investigation of a system under asynchronous dynamics. We model the system as a Boolean network in which transitions occur as independent Poisson processes in order to investigate the yeast cell cycle with asynchronous updates. In [11] a related approach has been used for the description of the yeast cell cycle. There, however, the underlying dynamics is synchronous. Fluctuations are introduced by allowing (with a small probability) state transitions that are actually forbidden by the update rules. In our model, the updating scheme is fundamentally asynchronous, and the stochastic nature of the system is caused by the randomness of the update sequence.

II. MODEL OF GENE REGULATORY NETWORKS: THE YEAST CELL CYCLE

A. A Network Representation of the Yeast Cell Cycle

The cell cycle is a regulatory process that couples growth, defined by an increase in cell mass, to proliferation, defined by an increase in cell number. The general process consists of a sequence of events through which one cell grows, replicates its components and divides into two daughter cells [12, 13]. In eukaryotes, we discuss the cell cycle in terms of four main phases: G_1 (growth and commitment to replication), S (DNA replication), G_2 ("gap" phase consisting of DNA repair and preparation for division) and M (mitosis or cell division). Progression from one stage to the next is controlled by a network of interacting proteins and RNA that control gene expression. Thus, robust control of the cell cycle is required to maintain normal cellular function. The molecular components of the eukaryotic cell cycle are highly conserved [14]. Thus, detailed studies in organisms such as budding yeast provide a relevant model for cell cycle functions in multicellular organisms.

Approximately 800 genes (15%) in the budding yeast *Saccharomyces cerevisiae* are regulated in response to the cell cycle [15]. Fortunately, there are a small number of regulators that control the overall process. For the purposes of testing our modeling framework, we used a minimal Boolean representation of the cell cycle from budding yeast [7] (Figure 1). This network represents the key regulatory machinery necessary for controlling the transition from G_1 through S, G_2 and M, producing two daughter cells in the G_1 state.

We model the cell cycle in terms of four major biological processes: transcription, phosphorylation/dephosphorylation, protein degradation, and inhibitory binding [16]. Each protein within the network, represented by a single node, is classified according to its predominant biological function (Table I). In our Boolean network representation of the cell cycle, each protein can either activate (turn on) or inhibit (turn off) target proteins through a biological process represented by a single, directed edge (Table II). The processes represented by each edge were chosen to represent the most important regulation step by which a protein activates or inhibits a target protein, based on a review of the literature (see [7, 16] and references therein).

Node Type	Biological classification	Members	Function
transcription	transcription factors	SBF, MBF, Mcm1/SFF, Swi5	Activate proteins by initiating synthesis of mRNA messages.
degradation	ubiquitin ligases	Cdh1, Cdc20,14	Target proteins for degradation by the proteosome machinery.
binding	Cyclin-Kinase Inhibitors (CKI)	Sic1	Bind to and inactivate Cyclin-Cdc28 kinases.
phosphorylation, dephosphorylation	cyclins	Cln1,2, Cln3, Clb1,2, Clb5,6	Bind to Cdc28 and (de)phosphorylate target proteins.

TABLE I. Biological processes of the yeast cell-cycle network: Nodes (proteins/enzymes)

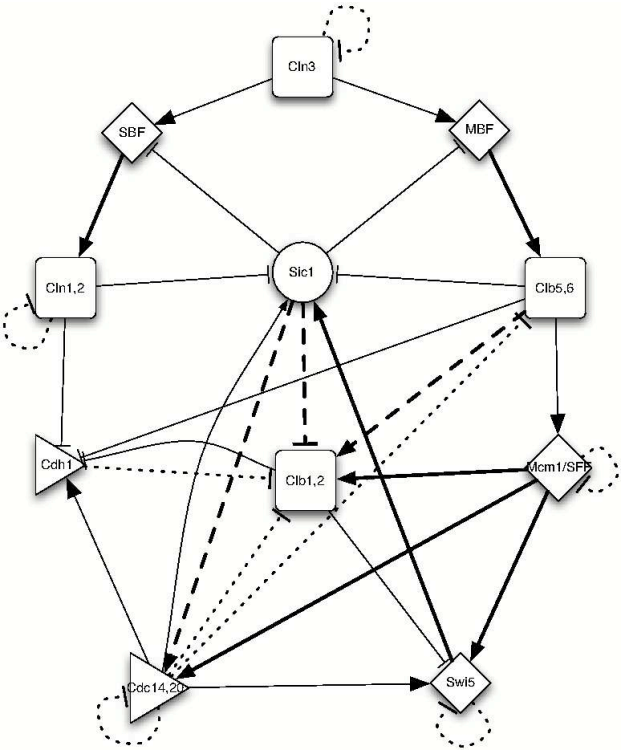

FIG. 1. Schematic diagram of the budding yeast cell cycle representing key regulatory functions. Nodes (shapes) in the network represent proteins classified according to their predominant biological function (Table I): cyclins (□, (de)phosphorylation), cylin-kinase inhibitior (◯, binding), ubiquitin ligases (▷, degradation), transcription factors (◇, transcription). In a Boolean network, each node is either active ("on" or 1) or inactive ("off" or 0). Each node is connected to one or more nodes via edges (lines). The edges have a direction indicating that the source node either activates (→, turns on) or inhibits (⊣, turns off) the target node. The line represention of each edge indicates the predominant biological process through which a given node activates or inhibits its target node (Table II): phosphorylation/dephosphorylation (——), ubiquitin/degradation (· · ·), binding (- - -), transcription (——). In this network, the stationary G_1 phase of the cell cycle is a state in which Sic1 and Cdh1 are "on", and all other nodes are "off".

Edge Type	Biological processes	Relative rate
transcription	transcription, translation, nuclear transport	1
degradation	ubiquitin ligation, proteosome degradation	9
binding	stable binding	25
phosphorylation, dephosphorylation	phosphorylation dephosphorylation	100

TABLE II. Biological processes of the yeast cell-cycle network: Edges (biological processes)

Biological processes represented by the edges can occur on very different timescales. Detailed models of the cell cycle formulated as differential equations require binding and rate constants for each reaction, which were determined through *in vitro* experiments or by fitting parameters to match known experimental outcomes for multi-step processes [19, 20]. In contrast, we model the difference in biological times scales by assigning a single "rate" to each edge, according to the biological process that edge the represents (Table III).

The assigned rates do not correspond to any single measured biological event. Rather, they are derived mainly from the known rates of key proteins or protein complexes that carry out each process. We derived the rate for phosphorylation and dephosphorylation by averaging the steady state phosphorylation rates of the Clb/Cdc28 kinases [21]. Only three proteins regulate activity through inhibitory binding in this network and Sic1 is the most well-characterized. Therefore, we used the estimated equilibrium binding constant for Sic1 [19, 20] as a basis for the binding rate. The rate for protein degradation was the most difficult to estimate due to the lack of detailed experimental information about the degradation mechanism. Since a precise measurement for proteasome degradation is not available, we determined the protein degradation rate from estimates of the total protein turnover in a population of cells [18]. The key transcription factors in the cell cycle network, MBF and SBF, are preloaded onto DNA in the nucleus, but only affect transcription initiation after they are phosphorylated (reviewed in [16]). After the transcription factors become active, mRNA synthesis is limited by the speed of RNA Polymerase II. Thus, we chose the rate for transcription

Process	Rate	Biological Rate[a] (min^{-1})	Literature Reference
transcription	1	1–2	Based on the elongation rate of eukaryotic RNA polymerase II [17]
protein degradation	9	3–5	Proteasome degradation rate based on protein turnover in cells [18]
binding	25	50	Estimated Sic1 association constant [19, 20]
phosphorylation dephosphorylation	100	200	Average Michaelis-Menton k_{cat} for Clb-CDKs [21]

[a]Rates are all approximations used directly or derived from information in the literature references

TABLE III. Relative rates for biological processes

based on the measured elongation rate of eukaryotic RNA Polymerase II (reviewed in [17]).

Although the rates are only crude estimates, they are consistent with biological timescales, given the basic requirements for each process (reviewed in [1]). For example, phosphorylation, dephosphorylation and inhibitory binding events usually only require the activity of a single protein or protein complex. Thus, a kinase or a phosphatase acts by binding to a target protein, covalently adding (kinase) or removing (phosphatase) a phosphate molecule and dissociating from the target protein. Similarly, Sic1 inhibits CDK complexes by noncovalently binding to the CDK complex. The complex remains inhibited as long as Sic1 is bound. For both of these processes, association with the target protein is often rate limiting, thus, we model inhibitory binding as slightly slower than phophorylation, since the inhibitor must remain bound to the target in order to have an effect. In contrast, protein degradation and transcription each encompass more than one sub-process. Cell cycle proteins are degraded by the proteasome complex, which actively unfolds proteins and cleaves them into short peptide fragments in an energy (ATP) dependent reaction [22, 23]. However, in order to be degraded by the proteasome, a target protein must first be tagged by covalently attaching ubiquitin to a specific site on the surface of the target [24]. Thus, protein degradation occurs on a much slower time scale than phosphorylation or binding. Similarly, a protein that is controlled via transcription also requires several steps prior to becoming active: synthesis of mRNA from DNA in the nucleus, transport of the mRNA out of the nucleus, translation into protein, and in some cases, transport back into the nucleus. Thus, compared to regulatory events controlled by phosphorylation or binding, transcriptional control is very slow. By estimating rates for processes that are not completely dependent on the cell cycle, our modelling framework can be readily adapted for analyzing other regulatory networks.

In our simplified cell cycle network, a normal, growing cell in the G_1 state has two active proteins (nodes), Cdh1 and Sic1, which together inactivate the Clb1,2 and Clb5,6 cyclins. These cyclins control entry into S phase (Clb5,6) and M phase (Clb1,2). Thus, in the G_1 state, the cell can grow, but is not committed to dividing, and is said to be in a stationary growth state. Exit from stationary growth requires Cln3 activation, which serves as a "start" signal for irreversible commitment down the path towards division. Once activated, Cln3 activates the transcription factors SBF and MBF, which produce the G_1 cyclins Cln1,2 and Clb5,6. These cyclins inactivate Cdh1 and Sic1, which are stabilizing the cell in G_1, allowing the cell to progress into S phase. Further progression through S, G_2 and eventually M, results in cell division, with each daughter cell in the stationary G_1 growth state.

In cells, the transitions between the four different phases of the cell cycle are sensitive to physical properties of the cell. The most important physical properties that affect cell cycle regulation are: cell size, firing of replication origins, absence of DNA damage, and spindle attachment to chromosomes. Cells have sensory networks that act as feedback control mechanisms for each of these four properties. The sub-networks that act as feedback control mechanisms by sensing physical properties are referred to as "checkpoints". Since the networks that control checkpoint responses are complex, our network model does not explicitly account for any of these. Instead, proteins that would otherwise be the main regulatory targets of the checkpoint pathways are modelled as self-degrading proteins [7], with a self-degradation rate equal to the rate assigned for proteasome-mediated degradation (Table III).

B. Model Details

As outlined above, we model the gene regulation of the yeast cell cycle with discrete but asynchronous network dynamics. The nodes of the network represent the key regulators of the cell cycle, while the edges correspond to the processes of activation and inhibition. Our model is Boolean in the sense that every node can assume one of two possible values: "on" (1) or "off" (0). A network state is defined as the vector of all node states. In our yeast model, where 11 nodes are considered, a network state is thus one of $2^{11} = 2048$ states.

A transition between two network states occurs with a characteristic rate specific to the individual process causing the transition. As we are considering an asyn-

chronous model, only transitions between network states that differ in exactly one node state are possible. However, the set of allowed transitions is further restricted by the topology of the network. In this section we introduce two different schemes that determine the allowed transitions between network states and the rates with which these transitions occur: the *independent edge model* and the *Boolean threshold model*. When the system is in a particular network state, usually several transitions are allowed. The different transitions then compete according to the rates, which we will explain in more detail in section III.

Independent Edge Model

In the independent edge model, each edge switch the state of the node that it regulates. In practice, that is to say an activating edge can switch on an inactive node while an inhibiting edge can cause the opposite event. These transitions happen at the rate associated with the reaction that corresponds to the edge under consideration. In Boolean terms this means that the edges running into a node are OR-connected, as any one of these edges can independently cause the transition.

The rate of transition between two network states in the independent edge model is simply given by the sum of all rates that cause the respective node state switch. For example, node Clb5,6 is down-regulated by Cdc20 and Sic1. Adding the rates attached to the links Cdc20⊣Clb5,6 and Sic1⊣Clb5,6 gives the total Clb5,6 deactivation rate which will then be assigned to all network transitions in which Cdc20, Sic1 and Clb5,6 are active and Clb5,6 then gets deactivated.

Boolean Threshold Model

In the Boolean threshold model, the set of allowed transitions is more restricted. It is based on the dynamical rule introduced in [7] to describe the yeast cell cycle as a synchronous Boolean system. Assume the network is in a state $\mathbf{S} = (s_1, ..s_N)^T$ where s_i denotes the state of node i. Then a node i is allowed to switch to state $s_i' \neq s_i$ if the following condition is met:

$$s_i' = \begin{cases} 1 & \text{if } \sum_j A_{ij} s_j > 0 \\ 0 & \text{if } \sum_j A_{ij} s_j < 0 \end{cases}$$

where A_{ij} is set by the connections of the network and is defined as

$$A_{ij} = \begin{cases} 1, & \text{if node } j \text{ has an } \textit{activating} \text{ link to node } i \\ -1, & \text{if node } j \text{ has an } \textit{inhibiting} \text{ link to node } i \\ 0, & \text{if node } j \text{ does not affect node } i \end{cases}$$

Effectively this rule corresponds to a "majority vote" scheme which decides whether the activating or the inhibiting edges determine the state of the node. An additional subtlety is the treatment of *self-degrading* nodes, i.e. nodes that inhibit themselves. The convention is that a self-degrading link will count as 1/10 of the others in the sums of the threshold rule. This means that the contribution of the self-degradation term will always be neglected unless the sum of the external influence is zero. Using these rules we allow a node to switch (given a particular network state) if and only if, in the synchronous dynamics of [7], the respective node would switch in the next time step. This has the effect that the stationary states of our stochastic model are exactly the same as in the synchronous case.

Regarding the rates of the transitions in the Boolean threshold model, we cannot simply transfer the rates from the edges to the network state transitions as we have done in the independent edge model. In principle, for every network state one would have to determine the rate of a given state switch considering all the competing activating and deactivating processes. In this project, we have chosen a simple approximation and have determined the typical rate at which nodes turn on (off) given their activating (inhibiting) input links.

Node	Activation		Deactivation	
	process[a]	rate[b]	process[a]	rate[b]
Cln3	N/A	1	self-deg	9
MBF	dephos	100	phos	100
SBF	phos	100	dephos	100
Cln1,2	trans	1	self-deg	9
Cdh1	dephos	100	phos	100
Swi5	dephos trans	51	phos degradation	55
Cdc20,14	trans binding	13	self-deg	9
Clb5,6	trans	1	binding degradation	17
Sic1	trans dephos	51	phos	100
Clb1,2	binding trans	13	binding degradation	17
Mcm1	phos	100	self-deg	9

[a] phos = phosphorylation, dephos = dephosphorylation
trans = transcription, self-deg = self-degradation
[b] Rates that depend on more than one process were calculated by averaging the rates for all processes

TABLE IV. Processes and rates for the Boolean threshold model

If a node is activated by only one process (for example Swi5 is activated only through phosphorylation), the rate of that process is used for all activation transitions of that node. If a node is affected through several processes, we average over all rates of these processes. In table IV we give the processes that affect each regulator and the respective rate that we associate with the corresponding node. We want to emphasize that this is a rather crude approximation and should be replaced by a

more sophisticated mechanism in the future.

III. THEORY & METHODS

As outlined previously, in our model of a gene regulatory network, we wish for the updates of individual sites to be asynchronous. The most basic mechanism such a scheme can follow is under the assumption of statistical independence between update "events" at different sites. Updates of the sites happen at the rates defined and reasoned in section II. These correspond to the transitions between states of the network governed by the rules of the independent edge or the Boolean threshold model respectively. Assuming the system to be *memoryless*, such that the probability of making a transition only depends on the current state, this can be formulated rigorously as a continuous-time Markov process [25].

A. Stochastic Modeling of Gene Regulatory Networks: Master Equation Description

The system as a whole is thought of as being in a certain *microstate*. Such a microstate is a particular configuration of the system, $\mathbf{S} = (s_1, \ldots, s_N)^T$, where the $s_i \in \{0, 1\}$ denote the state of site i which can be on (1) or off (0). The probability of finding the system in a specific microstate $\mathbf{S}$ at time t is denoted by $P(\mathbf{S}, t)$. Transitions between these microstates happen with asynchronous dynamics which means that transitions between $\mathbf{S}$ and $\mathbf{S}'$ happen at a specific rate $w_{\mathbf{S}\to\mathbf{S}'}$ if the states are coupled by the rules of the model. The process is therefore stochastic, however, the time evolution of the probability $P(\mathbf{S}, t)$ is deterministic. This time evolution and therefore the time evolution of the system as such is governed by the master equation

$$\partial_t P(\mathbf{S}, t) = \sum_{\mathbf{S}'} w_{\mathbf{S}'\to\mathbf{S}} P(\mathbf{S}', t) - \sum_{\mathbf{S}'} w_{\mathbf{S}\to\mathbf{S}'} P(\mathbf{S}, t) \quad (1)$$

which describes the flow of probability into and away from the microstate $\mathbf{S}$ [26].

Along with the initial configuration of the system $P(\mathbf{S}, t = 0)$, this master equation fully determines the time-evolution of the system. As at each point in time the further development of the system depends only on the current state, and does not depend on the previous history of the process. This is the memoryless property introduced above and makes this scheme a continuous-time Markov process, or more specifically, a *Poisson process*. For this class of processes, the probability of observing a certain number of events in a set interval follows a Poisson distribution. It can be shown that the waiting times between successive events are exponentially distributed, $P(\tau) \sim \mathrm{e}^{-r\tau}$, where r is the rate of the process. Furthermore, one can show that for a set of independent Poisson processes the waiting time to the next occurring event is exponentially distributed with a rate $R = \sum_i r_i$, which is the sum of the rates of all individual independent processes [25]. These results will be vital when considering computer simulations of our model in section III B.

Summing up, we now have a formalism which describes gene regulation as a stochastic process. Updating events are considered to be happening as independent Poisson processes with specific rates. This enables us to fully describe the state of the system by a probability distribution $P(\mathbf{S}, t)$ of finding the system in its various possible states $\mathbf{S}$ at time t. Its time evolution is given by a linear differential equation in time, the master equation (1). In the following, we will develop a vector representation for this equation in close analogy to the formalism used in Quantum Mechanics.

Given that in our gene regulatory network each site can be in one of two states (on (1) or off (0)), the number of possible microstates of the system is 2^N. Following [26], we can represent each microstate $\mathbf{S}$ by a 2^N-dimensional vector $|\sigma\rangle$ that has all but one non-zero component, such that $\langle\sigma'|\sigma\rangle = \delta_{\sigma',\sigma}$, where $\langle\sigma'|$ denotes the transpose of $|\sigma'\rangle$. Thus, the set of vectors $|\sigma\rangle$ form an orthonormal basis of a 2^N-dimensional vector space. We choose the so-called Dirac *bra(c)ket* notation for the vectors $|\sigma\rangle$ in order to distinguish these 2^N-dimensional vectors from the N-dimensional microstates $\mathbf{S}$.

In this vector space, we can define a time-dependent probability state vector, $|P(t)\rangle = \sum_\sigma P(|\sigma\rangle, t)|\sigma\rangle$, such that $\langle\sigma|P(t)\rangle = P(|\sigma\rangle, t)$. For normalization, the elements of the probability state vector have to sum to unity. This is expressed by defining an identity state vector $\langle I| = \sum_\sigma \langle\sigma|$ such that $\langle I|P\rangle = 1$.

Applying this formalism, the master equation (1) takes the following form

$$\partial_t |P(t)\rangle = -\hat{L}|P(t)\rangle \quad (2)$$

where $\hat{L}$ is called the Liouville operator. This operator governs the time evolution of the probability state vector and therefore the dynamics of the system. The matrix elements of this operator in the $|\sigma\rangle$-basis are given by

$$L_{\sigma'\sigma} = \langle\sigma'|\hat{L}|\sigma\rangle = -w_{\sigma\to\sigma'} \qquad \text{for } \sigma \neq \sigma' \quad (3)$$

$$L_{\sigma\sigma} = \langle\sigma|\hat{L}|\sigma\rangle = \sum_{\sigma'\neq\sigma} w_{\sigma\to\sigma'}. \quad (4)$$

The off-diagonal elements of the matrix $L_{\sigma'\sigma}$ contain the negative transition rates from configuration $|\sigma\rangle$ to $|\sigma'\rangle$ while the diagonal elements govern the decay of state $|\sigma\rangle$ and therefore contain the sum of all transition rates $\sigma \to \sigma'$. This matrix has the property that its columns sum to zero, i.e. $\langle I|\hat{L} = 0$, which makes $\hat{L}$ a *stochastic* matrix [27]. This follows from probability current conservation and can be easily seen by pre-multiplying equation (2) by $\langle I|$. $\hat{L}$ is non-Hermitian and therefore does not necessarily possess a complete set of eigenvectors that span a 2^N-dimensional eigenspace. Furthermore, its eigenvalues are not guaranteed to be real. However, there are a range of properties that are outlined in appendix A which can be used to ensure that

what follows is justified. Chief among those is a condition of diagonalizability of any given matrix that is satisfied by all the stochastic matrices that we consider in this work. In the following we are going to assume that the eigenvectors of $\hat{L}$ are $|\psi_0\rangle, |\psi_1\rangle, \ldots, |\psi_{2^N-1}\rangle$ with eigenvalues $\mu_0, \mu_1, \ldots, \mu_{2^N-1}$ and that these form a basis for the 2^N-dimensional eigenspace of $\hat{L}$, which is shown in appendix A. These can be normalized, such that $\langle\psi_i|\psi_j\rangle = \delta_{ij}$ and completeness can be expressed as

$$\mathbf{1} = \sum_i |\psi_i\rangle\langle\psi_i| \tag{5}$$

It is important to note that the left eigenvectors $\langle\psi_i|$ are in general not the transpose or Hermitian conjugate of the right eigenvectors $|\psi_i\rangle$.

The master equation expressed via the Liouville operator (equation (2)) has the formal solution

$$|P(t)\rangle = \exp\left(-\hat{L}t\right)|P(0)\rangle \tag{6}$$

$$= \sum_i \exp\left(-\mu_i t\right)\langle\psi_i|P(0)\rangle\ |\psi_i\rangle \tag{7}$$

where the resolution of identity (5) has been used in the last step. This equation gives the full time evolution of the state of the system. As mentioned in the appendix, all eigenvalues of $\hat{L}$ are expected to be non-negative, which means that in the long-time limit, only states with zero eigenvalue contribute. These states are the stationary states of the system that admit no further time evolution and are of key interest in this study. They correspond to pure states in the eigenbasis of $\hat{L}$, but can in general be linear combinations of states in the basis $|\sigma\rangle$. The all-zero configuration $(0, \ldots, 0)^T$ is a trivial stationary (or *absorbing*) state of the system. Eigenvectors that are combinations of states in the $|\sigma\rangle$-basis correspond to limit cycles of the dynamics.

As mentioned above, in the long-time limit $t \to \infty$, the probability of finding the system on a non-stationary state decays exponentially. Therefore, whatever state one starts in, in this limit the system will always end up with a stationary distribution $|P^*\rangle$.

$$\lim_{t\to\infty} e^{-\hat{L}t}|P_0\rangle = |P^*\rangle \tag{8}$$

From equation (7) it follows that we can define a projection operator

$$T^* = \lim_{t\to\infty} e^{-\hat{L}t} = \sum_{j=1}^{k} |\psi_0^j\rangle\langle\psi_0^j| \tag{9}$$

where the sum runs over the stationary states $|\psi_0^j\rangle$ of the system. The late-time behavior for a process with an initial probability distribution $|P_0\rangle$ is then given by

$$|P^*\rangle = T^*|P_0\rangle = \sum_{j=1}^{k} \langle\psi_0^j|P_0\rangle|\psi_0^j\rangle \tag{10}$$

This means that starting from an initial probability distribution $|P_0\rangle$ in the long-time limit the system's state will be described by the probability distribution $|P^*\rangle$. In connection with this project, this result will enable us to investigate the probability weight of certain gene expression patterns.

B. Simulating Stochastic Gene Regulation

An alternative method of investigating the dynamic process defined on the directed network which models gene regulation in budding yeast is Monte Carlo simulation. In this type of computer simulation [28], a random number generator is employed to evolve the system over time starting from a certain initial condition, according to the stochastic rules of the process.

As mentioned in section III, when considering the update events to be independent Poisson processes, the waiting time to the next possible event is given by an exponential distribution $P(\tau) \sim \mathrm{e}^{-R\tau}$ where $R = \sum_i r_i$ is the sum of the rates of all processes. Once a waiting time has been sampled from this distribution, the corresponding event is selected randomly among all possible events weighted by their associated rate r_i. This yields a simple algorithm for a time-dependent simulation of our model of stochastic gene regulation.

While in the formalism presented above, the infinite-time limit can be explicitly taken, a simulation will inevitably have to be terminated after a maximum runtime. Furthermore, a single run of a stochastic simulation will reach an absorbing state (one that can be entered but cannot be left). Due to the stochastic nature of the simulation method, this may happen after a very small number of steps thus not necessarily giving a representative value for the observables to be monitored. Also, ultimately we are interested in the fraction of times the system gets trapped in different stationary states. Thus, in order to obtain accurate average values for these quantities, a large number of simulations have to be performed and averaged, in which the runtime is chosen such that the observable quantities have attained stationary values. This Monte Carlo simulation technique will be used to support the findings obtained from the formalism presented above.

IV. RESULTS & DISCUSSION

The above formalism to represent our stochastic model of gene regulation was applied to the budding yeast and the stationary gene expression patterns (stationary states) in the two updating schemes were determined.

	Cln3	MBF	SBF	Cln1,2	Cdh1	Swi5	Cdc20,14	Clb5,6	Sic1	Clb1,2	Mcm1		Probability Weight homog. rates	Probability Weight heterog. rates
1.	0	0	0	0	1	0	0	0	1	0	0	G_1	0.47	0.80
2.	0	0	0	0	0	0	0	0	1	0	0		0.22	0.05
3.	0	0	0	0	0	0	0	0	0	0	0		0.06	0.06
4.	0	0	0	0	1	0	0	0	0	0	0		0.07	0.01
5.	0	0	1	1	0	0	0	0	0	0	0	} limit cycle	0.09	0.06
6.	0	0	1	0	0	0	0	0	0	0	0		0.09	0.02

TABLE V. Probability weight of observed stationary states in the independent edge model

	Cln3	MBF	SBF	Cln1,2	Cdh1	Swi5	Cdc20,14	Clb5,6	Sic1	Clb1,2	Mcm1		Probability Weight homog. rates	Probability Weight heterog. rates
1.	0	0	0	0	1	0	0	0	1	0	0	G_1	0.56	0.84
2.	0	0	0	0	0	0	0	0	1	0	0		0.18	0.02
3.	0	0	0	0	0	0	0	0	0	0	0		0.02	0.03
4.	0	0	0	0	1	0	0	0	0	0	0		0.01	0.00
5.	0	0	1	1	0	0	0	0	0	0	0		0.13	0.06
6.	0	1	0	0	0	0	0	0	1	0	0		0.02	0.01
7.	0	1	0	0	1	0	0	0	1	0	0		0.08	0.04

TABLE VI. Probability weight of observed stationary states in the Boolean threshold model

In particular, the probability of finding the system in a particular state given an arbitrary initial condition was considered. In practice, starting from a uniform initial probability distribution $|P_0\rangle$ this amounts to analyzing the elements of the resulting long-time distribution $|P^*\rangle$ as defined in equation (10). In order to further support our findings, we investigated the fraction of times the system gets trapped in particular stationary gene expression states in the Monte Carlo simulation averaged over all initial conditions.

As a starting point, each of the models, that is the independent edge and the Boolean threshold model, were considered with a uniform set of rates in order to gain a first insight into the effect stochasticity has on our system. In table V we show the six stationary states of the independent edge model which we will refer to as IE1 through IE6. In table VI, the stationary states of the Boolean models (BT1,...,BT7) are listed. As mentioned above, these seven stationary states are the same as in the synchronous deterministic model in [7].

Comparing the stationary states of the two models, we find that the first five expression patterns are present in both models. The first state is the biological G_1 stationary state. In the investigation of the synchronous model [7] this was found to be the by far dominating fixed point, to which about 86% of all initial conditions lead. The two other states with the largest basin sizes (number of initial conditions leading to the respective states) in the synchronous model were (in that order) IE5/BT5 and BT7. The other four states of the synchronous model were found to have very small basin sizes of less than 0.6%.

In the independent edge model, two fixed points of the synchronous dynamics are not present (BT6, BT7) and we want to discuss this point as it is a typical example of the difference between our two models. Both of these effects can be easily understood when looking at the update rules. In the Boolean model some transitions are prohibited because, for example, the influence of an activating edge is suppressed by two inhibiting edges. This is the case with Clb5,6 in the stationary states BT6 and BT7. As Sic1 and Cdc20,14 both inhibit Clb5,6, MBF cannot activate it in the Boolean model ("majority vote"). In the independent edge model, the activation of Clb5,6 can happen although two inhibitors are present.

Another important difference between the results of our two models is the state BT5/IE5. In the independent edge model, this state is not a fixed point, but actually forms a limit cycle of the dynamics together with IE6. This is also immediately apparent when looking at Cln1,2 which is a self-degrader, i.e. deactivates itself; but is also activated by SBF. These two processes compete which leads to a fixed point in the Boolean model but a continuous "blinking" of Cln1,2 in the independent edge model.

All of these differences, however, concern stationary states that do not have a clear biological significance. Next, we want to investigate the probabilities that a specific stationary state is taken by the system. In both models, the states 1 and 2 are most likely. In the independent edge model the stationary state G_1 (IE1) occurs with a probability of 47% and IE2 with 22%. In the Boolean threshold model, BT1 collects 56% of initial conditions, BT2 18%.

It is interesting to note that IE2/BT2 has a very small basin size in the synchronous dynamics. It seems that it is over-represented in our stochastic model as compared to the deterministic dynamics because it only differs from

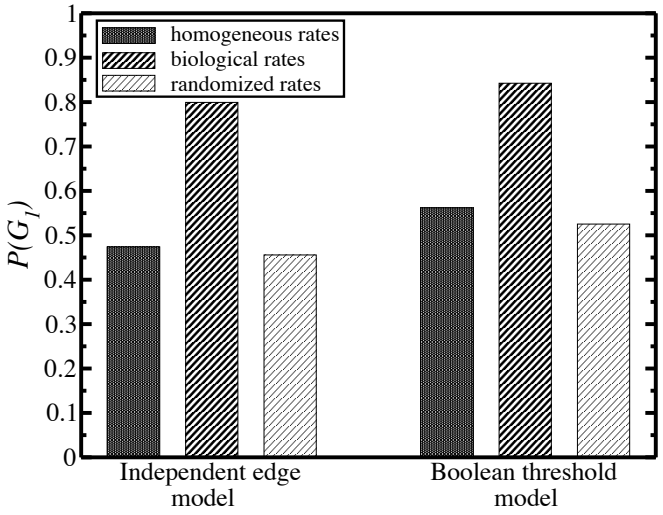

FIG. 2. Probability $P(G_1)$ that the system enters that G_1 state for the independent edge model and the Boolean threshold model, each with homogeneous, biological and randomized heterogeneous rates.

G_1 by one node state: Cdh1 is inactive. During the biological cell cycle regulation Sic1 is activated before Cdh1 becomes active. If its activator, Cdc20,14 gets deactivated via self-degradation, Cdh1 will not turn on and the system can end up in the IE2/BT2 state. We don't want to elaborate on the state sequence during the biological pathway (see [7] for details), but it is important to note that there is a competition between self-degradation and phosphorylation (Cdc20,14→Cdh1).

The next most likely stationary state is BT5 (13%), IE5 (18% for the limit cycle IE5,6). This is not surprising as this state also has a large basin size in the synchronous model. The other stationary states are rather improbable and sum to a probability weight of about 13% in both models. To summarize, in our stochastic asynchronous model of the yeast cell cycle, we reproduce the biological G_1 stationary state as one of several possible stationary states. With the unrealistic homogeneous rates, the G_1 stationary state is reached by the system with a probability of about 50% in both models.

In order to investigate the effect of introducing our biologically inspired rates, the above analysis was repeated for both models with the rates associated to the processes as given in section II A. As can be seen from the last column of tables V and VI, the probability weights are shifted considerably by the change in rates. Interestingly, the probability weight of the biological G_1 stationary state strongly increases when the heterogeneous rates are used. In the independent edge model, it increases by 70%, in the Boolean threshold model by 56%. All other states (with the exception of the trivial state BT3) lose in probability weight. In both models, the G_1 state is thus reached with a probability of more than 80%.

We then investigated whether this increase in probability weight of the stationary G_1 state is simply due to heterogeneity of the rates (as opposed to the specific numbers in the biological rates). In figure 2 we show the probability weights of the G_1 state for homogeneous rates, biologically inspired heterogeneous rates and randomized rates with values between 1 and 100. As can be clearly seen, the randomized rates results do not show the increase in G_1 probability weight.

From the above results we conclude that while our model allows for the occurrence of non-biological stationary states, the G_1 stationary state features prominently in all cases considered. Furthermore, the introduction of biologically inspired rates raises the probability weight of the biological G_1 stationary state considerably.

Having found that introducing heterogeneous rates in both the independent edge and the Boolean threshold model dramatically increases the expression of the biological G_1 state, we consider in the following the effects of the specific ordering of the rates. By ordering we mean a specific assignment of the four relative rates, 1, 9, 25 and 100, to the four biological processes described in section II, transcription, phosphorylation, degradation and binding. Of course, the ordering in section II is the biological sensible one, but it is still of interest to study how changing the ordering affects the expression of G_1. To that end, we computed the probability of arriving in state G_1 for all possible permutations of the ordering, i.e. all possible assignments of the four rates to the four processes. We find that for two processes, binding and transcription, their rank in the ordering (ranking the processes from the fastest to the slowest, e.g. transcription being rank one if it is assigned the relative rate of 100) does not affect the probability of ending up in G_1 very much. However, the ranks of degradation and phosphorylation have a strong influence so that the ordering of phosphorylation being the fastest and degradation the slowest is most favorable one. In order to quantify these qualitative results, we analyzed the probabilities of arriving in G_1, $P(G_1)$, for a certain permutation of rates in the following way: given a particular process l, for each rank r, we averaged $P(G_1)$ over all permutations in which the process has this particular rank. This averaging yields four quantities, $\langle P\rangle_r^l$, per process and, carried out for all four processes, a total of 16 values. These values of $\langle P\rangle_r^l$ are shown in Figures 3 (a) and (b) for the independent edge mode and the Boolean threshold model, respectively. The figures clearly show the trends that we described above. Of course, the standard deviations (not shown in the figure), especially for binding and degradation, are very large (up to three-quarters of the means) but this is to be expected: given that the ranking of phosphorylation affects $P(G_1)$ very much for example, all orderings with only degradation or binding at a fixed rank should produce large variations. The figures suggest that the ordering that produces the largest $P(G_1)$ is phosphorylation rate > transcription rate > binding rate > degradation rate. This is indeed the case, as this particular ordering yields the largest probabilities,

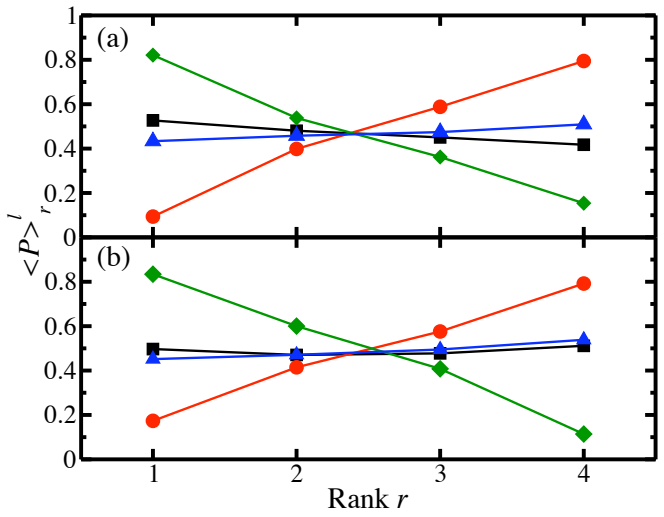

FIG. 3. Averaged probability of entering the G_1 state, $\langle P\rangle_r^l$, for the four different processes, transcription (□), degradation (○), phosphorylation (◇) and binding (△) as a function of the rank r of the process in the (a) *independent edge model* and the (b) *Boolean threshold model*.

$P(G_1) = 0.92$ and $P(G_1) = 0.86$, for the independent edge model and the Boolean threshold model, respectively. These are higher than the values of $P(G_1) = 0.80$ and $P(G_1) = 0.84$ for the biological ordering, phosphorylation > binding > degradation > transcription rate.

The importance of the rates of phosphorylation and degradation can be understood when looking at the network topology. We have mentioned before that these processes compete at Cdc20,14 which (if degradation happens faster) can prevent activation of Cdh1, thus not reaching G_1. The same argument can be used at a number of other nodes, such as Swi5, Mcm1, Cln1,2. Most important however, seems to be the activation of SBF and MBF by Cln3. In [10] it was shown in a different model that if Cln3 activates SBF and MBF before degrading itself, the probability of reaching the G_1 state is very high. As this affects 50% of all initial conditions, this seems to be an important factor in our model as well.

V. SUMMARY & OUTLOOK

In this project, we have introduced a discrete Boolean model that makes use of a stochastic asynchronous updating mechanism to describe gene regulatory networks and have applied it to the yeast cell-cycle dynamics. To investigate this model, we have used both a master equation approach and explicit Monte Carlo simulations.

We find that the biological G_1 fixed point is the most probable stationary state of our model. With the introduction of biologically inspired rates, as opposed to all processes happening at the same rate, the probability weight of this state further increases considerably. We show that this is in fact due to the specific choice of the rates and not merely a result of the heterogeneity of rates.

We further investigate how the choice of rates affects the probability of the system to end up in the G_1 state and find that the spread between phosphorylation rate and degradation rate is most important. We explain these results with respect to the network topology.

In the future we plan to further investigate the significance of our specific choice of rates. More elaborate robustness analyses of the dynamics with respect to changes in these would be in order. Furthermore, we plan to compare a set of optimal rates found via genetic algorithms to the ones used in this work. Lastly, we envisage the application of the framework developed above to other networks that control apoptosis, segment polarity in fruit fly embryos and differentiation in the development of flowers.

Acknowledgments

We thank S. Leibler, P. Stadler and A. Wagner for helpful discussions and S. G. Zeitlin for comments on the manuscript. This work was carried out as part of the Santa Fe Institute Complex Systems Summer School 2006 and we are grateful for the institute's kind hospitality. The authors acknowledge financial support from the PIP program (SB), the Skaggs Institute for Chemical Biology (BC), and the UK EPSRC and the Cambridge European Trust (SF and CN).

APPENDIX A: SPECTRAL PROPERTIES OF THE LIOUVILLE OPERATOR

In this appendix, a general condition for the diagonalizability will be described that justifies all the manipulations described in the text. *Diagonalizability* means that there exists a basis for a given vector space in which a linear transformation is represented by a diagonal matrix. The basis is necessarily given by the set of eigenvectors of the linear transformation. However, even though any $n \times n$-matrix has n eigenvectors, they do not necessarily form a basis, i.e. a set of linearly independent vectors that span the vector space.

It can be shown [29] that if the polynomial

$$m(x) = (x-\mu_1)\dots(x-\mu_r) = \prod_{i=1}^{r}(x-\mu_i), \qquad \text{(A1)}$$

where μ_i are the r distinct eigenvalues of a matrix L, is the *minimum polynomial* then its eigenvectors form a linearly independent set of vectors that span the vector space. The minimum polynomial of a matrix L, $m_L(x)$, is the monic [31] polynomial of least degree that satisfies the equation

$$m_L(L) = L^n + \lambda_{n-1}L^{n-1}\cdots + \lambda_1 L + \lambda_0 \mathbf{1} = \mathbf{0} \qquad \text{(A2)}$$

for some scalars λ_i. So, if a given matrix satisfies the following equation,

$$m_L(L) = \prod_{i=1}^{r}(L - \mu_i) = \mathbf{0}, \tag{A3}$$

as all the matrices in this work do, then L satisfies the condition for diagonalizability.

Given that the eigenvectors of the Liouville operator, $\hat{L}$, form a basis of the vector space, we can now show that the resolution of identity can be written as

$$\mathbf{1} = \sum_i |\psi_i\rangle\langle\psi_i|. \tag{A4}$$

First of all, we note that right and left eigenvectors corresponding to different eigenvalues are orthogonal. (In this moment, it might be appropriate to emphasize again that the left eigenvectors are in general not the transpose or Hermitian conjugate of the right eigenvectors and vice versa.) This can be easily seen when we consider the defining eigenvalue equations:

$$\hat{L}|\psi_j\rangle = \mu_j|\psi_j\rangle \tag{A5}$$

$$\langle\psi_i|\hat{L} = \mu_i\langle\psi_i| \tag{A6}$$

Premultiplying equation (A5) by $\langle\psi_i|$ and postmultiplying equation (A6) by $|\psi_j\rangle$ and then subtracting the two yields the following equation

$$0 = (\mu_j - \mu_i)\langle\psi_i|\psi_j\rangle \tag{A7}$$

from which follows that right and left eigenvectors whose eigenvalues are different, $\mu_i \neq \mu_j$, are orthogonal, i.e. $\langle\psi_i|\psi_j\rangle = 0$.

Let us now consider eigenvectors corresponding to the same eigenvalue. Given a set of k right and left eigenvectors, $\{\langle\psi_i^j|, |\psi_i^j\rangle : 1 \geq j \geq k\}$, with the same eigenvalue, we can always construct by linear transformation [32] a new set of left vectors $\{\langle\psi'^l_i| = \sum_{m=1}^{k}\alpha_{ml}\langle\psi_i^m|\}$ such that its elements are orthogonal to all but one right eigenvector, i.e. $\langle\psi'^l_i|\psi_i^j\rangle = \delta_{lj}$ for $1 \geq l \geq k$. The crucial point to note is that since the eigenvectors are degenerate, i.e. they all share the same eigenvalue, the linear combinations $\{\langle\psi'^l_i|\}$ are eigenvectors of $\hat{L}$ with eigenvalue μ_i as well, so that this linear transformation of the basis vectors leaves the matrix L diagonal.

Taking the facts that the eigenvectors span the vector space and that we can orthogonalize the left- and right eigenvectors, it can be seen that $\sum_i |\psi_i\rangle\langle\psi_i|$ indeed represents the identity matrix as shown in equation (A4), because it leaves any vector in the space, $|v\rangle$, invariant:

$$\begin{aligned}\sum_i |\psi_i\rangle\langle\psi_i|v\rangle &= \sum_i |\psi_i\rangle\langle\psi_i| \sum_j \gamma_j|\psi_j\rangle \\ &= \sum_{i,j}\gamma_j\delta_{ij}|\psi_j\rangle = |v\rangle.\end{aligned} \tag{A8}$$

According to a theorem by Gershgorin ([30]) there exist limits on the spectrum of matrices having the form of L. It ensures that there is at least one zero eigenvalue and that the real part of all non-zero eigenvalues is > 0.

[1] B. Alberts, A. Johnson, J. Lewis, M. Raff, and K. Roberts. *Molecular Biology of the Cell.* Garland Publishing, Inc., New York, 1994.

[2] J. J. Tyson, K. Chen, and B. Novak. Network dynamics and cell physiology. *Nat. Rev. Mol. Cell Biol.*, 2:908–916, 2001.

[3] P. Smolen, D. A. Baxter, and J. H. Byrne. Mathematical modeling of gene networks. *Neuron*, 26:567–580, 2000.

[4] S. A. Kauffman. Metabolic stability and epigenesis in randomly constructed genetic nets. *J. Theor. Biol.*, 22:437–467, 1969.

[5] R. Thomas. Boolean formalization of genetic control circuits. *J. Theor. Biol.*, 42:437–467, 1973.

[6] R. Albert and H. G. Othmer. The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in drosophila melanogaster. *J. Theor. Biol.*, 223:1–18, 2003.

[7] F. Li, T. Long, Y. Lu, Q. Ouyang, and C. Tang. The yeast cell-cycle network is robustly designed. *Proc. Natl. Acad. Sci. USA*, 101:4781–4786, 2004.

[8] K. Klemm and S. Bornholdt. Stable and unstable attractors in boolean networks. *Phys. Rev. E*, 72:055101, 2005.

[9] M. Chaves, E. D. Sontag, and R. Albert. Methods of robustness analysis for boolean models of gene control networks. arXiv:q-bio/0605004, 2006.

[10] S. Braunewell and S. Bornholdt. Superstability of the yeast cell cycle dynamics: Ensuring causality in the presence of biochemical stochasticity. arXiv:q-bio/0605009, 2006.

[11] Yuping Zhang, Minping Qian, Qi Ouyang, Minghua Deng, Fangting Li, and Chao Tang. Stochastic model of yeast cell cycle network. arXiv:q-bio/0605011, 2006.

[12] A. Murray and T. Hunt. *The cell cycle: An introduction.* Oxford University Press, New York, 1993.

[13] K. Nasmyth. A prize for proliferation. *Cell*, 107(6):689–701, 2001.

[14] P. Nurse. Universal control mechanism regulating onset of M-phase. *Nature*, 344:503–508, 1990.

[15] P. T. Spellman, G. Sherlock, M.Q. Zhang, V.R. Iyer, K. Anders, M.B. Eisen, P.O. Brown, D. Botstein, and B. Futcher. Comprehensive Identification of Cell Cycle-regulated Genes of the Yeast Saccharomyces cerevisiae by Microarray Hybridization. *Mol. Biol. Cell*, 9:3273–3297, 1998.

[16] M. D. Mendenhall and A. E. Hodge. Regulation of Cdc28 Cyclin-Dependent Protein Kinase Activity during the Cell Cycle of the Yeast Saccharomyces cerevisiae. *Microbiol. Mol. Biol. R.*, 62:1191–1243, 1998.

[17] S. J. Greive and P. H. von Hippel. Thinking quantitatively about transcriptional regulation. *Nat. Rev. Mol. Cell Biol.*, 6(3):221–32, 2005.

[18] J. W. Yewdell. Not such a dismal science: the economics of protein synthesis, folding, degradation and antigen processing. *Trends Cell Biol.*, 11(7):294–7, 2001.

[19] K. C. Chen, A. Csikasz-Nagy, B. Gyorffy, J. Val, B. Novak, and J.J. Tyson. Kinetic Analysis of a Molecular Model of the Budding Yeast Cell Cycle. *Mol. Biol. Cell*, 11:369–391, 2000.

[20] K. C. Chen, L. Calzone, A. Csikasz-Nagy, F. R. Cross, B. Novak, and J. J. Tyson. Integrative Analysis of Cell Cycle Control in Budding Yeast. *Mol. Biol. Cell*, 15:3841–3862, 2004.

[21] M. Loog and D. O. Morgan. Cyclin specificity in the phosphorylation of cyclin-dependent kinase substrates. *Nat. Cell Biol.*, 4:329–336, 2002.

[22] J. Pines and C. Lindon. Anaphase-promoting complex-dependent proteolysis of cell cycle regulators and genomic instability of cancer cells. *Nat. Cell Biol.*, 7:731–735, 2005.

[23] Keiichi I. Nakayama and Keiki Nakayama. Ubiquitin ligases: cell-cycle control and cancer. *Nature Reviews Cancer*, 6:369–381, 2006.

[24] D. Nandi, P. Tahiliani, A. Kumar, and D. Chandu. The ubiquitin-proteasome system. *J. Biosci.*, 31:101–119, 2006.

[25] N. G. Van Kampen. *Stochastic Processes in Physics and Chemistry*. North-Holland Publishing Company, New York, 1981.

[26] H. Hinrichsen. Non-equilibrium critical phenomena and phase transitions into absorbing states. *Advances in Physics*, 49:815, 2000.

[27] G. M. Schütz. *Exactly Solvable Models for Many-Body Systems Far From Equilibrium*, volume 19 of *Phase Transitions and Critical Phenomena*. Academic Press, London, 2001.

[28] M. E. J. Newman and G. T. Barkema. *Monte Carlo Methods*. Oxford University Press, New York, 1999.

[29] R. Kaye and R. Wilson. *Linear Algebra*. Oxford Univeristy Press, New York, 1998.

[30] S. Gershgorin. Über die abgrenzung der eigenwerte einer matrix. *Izv. Aked. Nauk. S.S.S.R.*, 7:749, 1931.

[31] coefficient of the highest power is unity

[32] e.g. by Gram-Schmidt orthogonalization [29]