J Stat Phys (2009) 136: 875–901
DOI 10.1007/s10955-009-9819-9

Density-Profile Processes Describing Biological Signaling Networks: Almost Sure Convergence to Deterministic Trajectories

Roberto Fernández · Luiz R. Fontes · E. Jordão Neves

Received: 26 May 2009 / Accepted: 1 September 2009 / Published online: 18 September 2009

Abstract We introduce jump processes in $\mathbb{R}^k$, called *density-profile processes*, to model biological signaling networks. Our modeling setup describes the macroscopic evolution of a finite-size spin-flip model with k types of spins with arbitrary number of internal states interacting through a non-reversible stochastic dynamics. We are mostly interested on the multi-dimensional empirical-magnetization vector in the thermodynamic limit, and prove that, within arbitrary finite time-intervals, its path converges almost surely to a deterministic trajectory determined by a first-order (non-linear) differential equation with explicit bounds on the distance between the stochastic and deterministic trajectories. As parameters of the spin-flip dynamics change, the associated dynamical system may go through bifurcations, associated to *phase transitions* in the statistical mechanical setting. We present a simple example of spin-flip stochastic model, associated to a *synthetic biology* model known as *repressilator*, which leads to a dynamical system with *Hopf* and *pitchfork* bifurcations. Depending on the parameter values, the magnetization random path can either converge to a unique stable fixed point, converge to one of a pair of stable fixed points, or asymptotically evolve close to a deterministic orbit in $\mathbb{R}^k$. We also discuss a simple signaling pathway related to cancer research, called *p53 module*.

Keywords Density-profile processes · Biological signaling networks · Spin-flip dynamics · Non-reversible stochastic dynamics · Thermodynamic limit · Dynamical system · Mean field

R. Fernández
Laboratoire de Mathématiques Raphaël Salem, UMR 6085 CNRS-Université de Rouen, Avenue de l'Université, B.P.12, 76801 St Etienne du Rouvray, France
e-mail: roberto.fernandez@univ-rouen.fr

L.R. Fontes (✉) · E.J. Neves
University of São Paulo, Rua do Matão, 1010, Cidade Universitária, CEP 05508-090, São Paulo, SP, Brasil
e-mail: lrenato@ime.usp.br

E.J. Neves
e-mail: neves@ime.usp.br

1 Motivation and Introduction

The interest in the analysis of the dynamical processes driving biological systems and the corresponding signal-processing mechanisms, together with recent successes in molecular biology and advances in computer technology, spurred a revival of *systems biology* ideas [1]. This approach [3, 4], which exploits ideas from dynamical systems and control theory [5], can be traced back at least to Erwin Schrödinger's meditation on *What is life?* [6], sixty five years ago.

1.1 Signaling Pathways

A central problem in biology is to understand how cells manage to communicate among themselves and respond properly to noisy signals from their environment. Most biological signals are received and transmitted through sequences of chemical reactions that may exhibit a striking similarity to electrical circuitry. Extra-cellular information is often transmitted through cell-membrane receptors activated by chemical entities known as *ligands*, such as hormones, neurotransmitters or growth factors (see Fig. 1). These signals trigger complex time-dependent cascades of internal cellular biochemical interactions that may lead to several different cellular responses, like embryogenesis, motility, differentiation and *apoptosis* (carefully controlled cell suicide) [7]. Such sequences of chemical reactions—called *signaling pathways* when the main interest is on the associated flow of information—can achieve complex signal processing tasks and participate in the control of larger and even more complex multifaceted cellular behaviours. Signaling pathways usually involve the interaction of a large multi-scale hierarchy of subsystems, from processes occurring at molecular level within small intra-cellular compartments to the coordinated dynamics of cells in tissues and organs across the life span of an organism. To illustrate some of the scale and complexity that may be involved in a real signaling pathways the reader could check a recent comprehensive overview in [8] of the important *epidermal growth factor* (EGF) signaling pathway, associated to the control of growth, survival, proliferation and differentiation in mammalian cells. This work, based on the molecular interactions documented in 242 papers accessible from PubMed (http://www.ncbi.nlm.nih.gov/pubmed), identified 211 associated chemical interactions involving 322 types of biochemical components.

The systems biology approach suggests multi-level analyses where the whole system is considered in terms of a hierarchy of interconnected functional *modules* [9, 10], usually small subset of interacting components capable to perform some *basic useful functionality* which contributes to the behavior of the whole system. An example of basic functionality is the capability of generating oscillations which enables time-periodic processes like circadian rhythms and cell cycle. Two other examples of functionalities that may be required in a given signaling pathway are *filters*, which allow responding to an input signal only when it is within a particular range of amplitude or frequency, and the *gradient sensor* module that allows detection of *variations* in concentrations, as required for *chemotaxis*. Our basic modeling setup, described below, can easily describe modules capable of exhibiting these simple functionalities and, to illustrate this, we discuss a simple class of modules, called *cyclic interaction modules*, capable of generating oscillations.

Signaling pathways may exhibit control mechanisms remarkably similar to those found in engineered systems [11] as they had to survive strong selective pressures for good performance and therefore had to be able to deploy efficient strategies to prevent inappropriate physiological responses. Malfunctions in this biochemical circuitry and its control mechanisms may lead to several pathological conditions, like cancer [12].

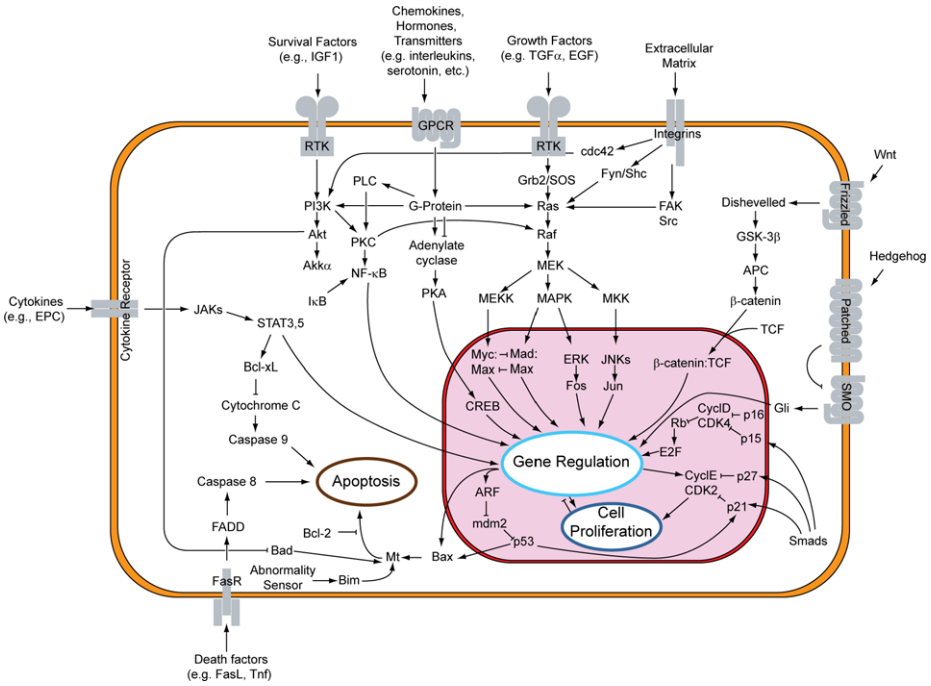

Fig. 1 Overview of Signaling Pathways [2] associated to the decision between proliferation and apoptosis; *pointed arrows* indicate some sort of activation interaction from the biochemical component at the source of the arrow on the component on its tip; similarly, *blunt arrows* indicate inhibition

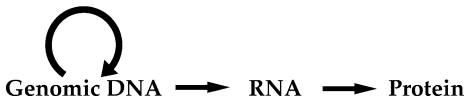

Fig. 2 Basic flow of sequence information processing in the cell

Despite the large amount of information available for a number of pathways [13], the actual biochemical mechanisms are frequently unknown. In fact biochemical components involved in a signaling pathways may frequently be quite large molecules which are, by themselves, information processing machines. Two important biochemical components in signaling pathways are proteins and genes (RNA strings). They are closely related through the *central dogma of molecular biology* [7]: genomic DNA—that contains essentially all genetic information and is copied on each cell division—is read (or *transcribed*) into genes (RNA strings) which is then read (or *translated*) into proteins (Fig. 2). The dynamics of the available quantity of each protein may be regulated by controlling the production of the associated gene in a process known as *gene regulation* [7].

Progress in the understanding of these pathways and their associated control mechanisms demands the development of mathematical models that manage good balance between simplicity and usefulness. Mathematical models should not only provide some understanding for the pathway's dynamical behaviour but also generate experimentally verifiable predictions. In view of the inherent complexity of most signaling pathways, *toy models*, in the old tradition of physics literature, may be particularly useful to illuminate how simple signaling pathways, or modules, manage to perform their basic functionalities and how these functionalities are further integrated into the whole pathway. To some extent, the recent increase

of interest in this systems biology approach came about from recent technological advances that allow exploring cellular processes with very fine experimental detail. Microarrays [14], for instance, which allows—in principle at least—measuring the activity of all the estimated 30,000 human genes in a given cell, is an important new tool in cancer research [15] with great potential in the clinical setting [16]. They also provide a most important tool for experimentally checking model predictions in situations where gene regulation is involved.

1.2 Modeling Signaling Pathways

Systems of chemical reactions where spatial inhomogeneities can be neglected and which involve a large number of molecules can be naturally modelled as nonlinear continuous-time dynamical systems. It is frequently the case that a molecule of a given biochemical component may be able to play different roles according to some internal state, say whether its spatial conformation exposes a particular catalytic region or not. Thus, suppose there are different types of components $\mathcal{T} = \{i_1, \ldots, i_k\}$ and that each one of the large number N of molecules of component i can be in the internal states $\mathcal{S}_i = \{a_1, \ldots, a_{s_i}\}$. Macroscopically, the system is described by giving, for each component i, the s_i densities of molecules in each state (only s_{i-1} of the densities are independent, because the total number of particles, N is fixed). Thus, if $\mathcal{E} = \prod_{i\in\mathcal{T}} \mathcal{S}_i$, the macroscopic description of the system is given by the dynamics of a density-profile $\{x_t\}_{t\geq 0}$ where $x_t \in \mathbb{R}^{\mathcal{E}}$. If the total number of molecules is large, the dynamics of this density-profile should be well described by a dynamical system of the form

$$\dot{x}_t = f(x) - g(x) \tag{1.1}$$

where f and g are bounded smooth functions from $\mathbb{R}^{\mathcal{E}}$ to $\mathbb{R}^{\mathcal{E}}_+$, indicating the velocity at which each one of the $|\mathcal{E}|$ component-states are, respectively, *produced* and *degraded* when the global density-profile is given by x. Behind this *macroscopic* view provided by the density-profile dynamics there is a *microscopic* system describing the interactions at molecular level. A key problem is, precisely, how to *derive* reasonable macroscopic descriptions though equations like (1.1) on the basis of the available information and of reasonable assumptions concerning the microscopic components and their interactions. Once this is done, a general framework based on dynamical systems ideas [18], often relying on numerical methods, can be applied to analyse the corresponding model [19, 20].

A typical method to deduce the dynamical system (1.1) describing a given pathway is to exploit ideas from chemical kinetics, like the law of mass action [21]. In fact, a particular type of equation (Michaelis-Menten) developed originally, under several strong simplifying assumptions, to model a particular type of interaction (enzymatic reaction) is frequently used in the literature as some sort of *default* equation for most reactions. An important drawback associated to this approach is the difficulty in evaluating to what extent the conclusions of a given analysis depend on these somewhat arbitrary choices of equations. Moreover, beyond reasonability issues regarding the models themselves, this approach is hampered by its dependence on too many parameters that can not be measured *in vivo*, that is, inside a living cell.

Our modeling setup In the present work we propose an alternative approach—a *toy model* approach—to derive the associated dynamical systems, which only seeks to incorporate the essential *qualitative* information about the biochemical interactions in a given signaling network. This approach borrows ideas from interacting particle systems [27], in such a way that spins represent the internal states of components of the signaling pathway and

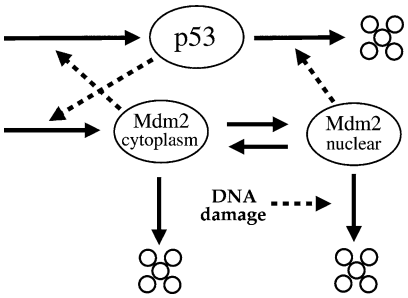

Fig. 3 Graphical representations of p53/Mdm2 interactions as given in Zhang et al. [25]; ovals represent proteins, *solid arrows* represent chemical reactions, *dashed arrows* represent activation effect of a protein on a reaction, and a set of five small circles represent degradation products

the global system evolves through a non-reversible Glauber spin-flip stochastic dynamics as explained in detail in Sect. 5 below. This stochastic spin system is projected onto a jump process in $\mathbb{R}^k$—called the *density-profile process*—which drives the discretized densities of the biochemical components in the signaling pathway. In the limit of a very large number of spins, these densities converge to time dependent functions satisfying the deterministic dynamical system 1.1. Thus, in statistical mechanics notation [34], the stochastic spin system and the density-profile process provide, respectively, the *microscopic* and *macroscopic* views of the same model. Rates with which molecular states change, which correspond to *spin flips*—transitions in the spin internal state—in our model, are chosen to depend on the global density-profile is such a way as to mimic the qualitative description of the associated interactions. In fact, the intermediate density-profile process may be of independent interest to analyse biochemical pathways where stochasticity plays an important role [22].

Note however that a comprehensive discussion of interesting signaling networks, like those described in [13] for instance, and their mathematical modeling can not be given here as it would unavoidably require delving into far too many biological considerations. Thus, as a compromise we will illustrate our modeling approach and its potential in basically two simple—but real—biological problems. Both may be thought as *modules* [10], rather than complete signaling pathways, as both of them, while involving a small number of components, still manage to provide some basic cellular functionality. We start with the discussion of the dynamics of a synthetic biologic system implemented *in vivo* in a bacteria [23]. First this is done informally, but in some detail, as a motivation for our approach. A more detailed explanation is given later in Sect. 6, where we also derive the corresponding differential equations and discuss the associated bifurcation analysis (Theorem 6.1). Next we present the second example, which concerns the puzzling response [24] of p53—an important *tumor suppressor gene* [7] involved in preventing cancer—to DNA damage. A graphic representation of the interactions described in this module is given in Fig. 3. From the corresponding dynamical equations we find that, as for the usual strategy based on chemical kinetics ideas [25], our approach can also explain the dynamical behavior of p53 by the presence of a *sub-critical Hopf* bifurcation with respect to the parameter associated to DNA damage (see Fig. 4).

2 Examples

2.1 Cyclic-interaction Module

Our first example is called the *cyclic-interaction* module and provides the basic functionality of generating oscillations. Such functionality is essential to organise time-modulated

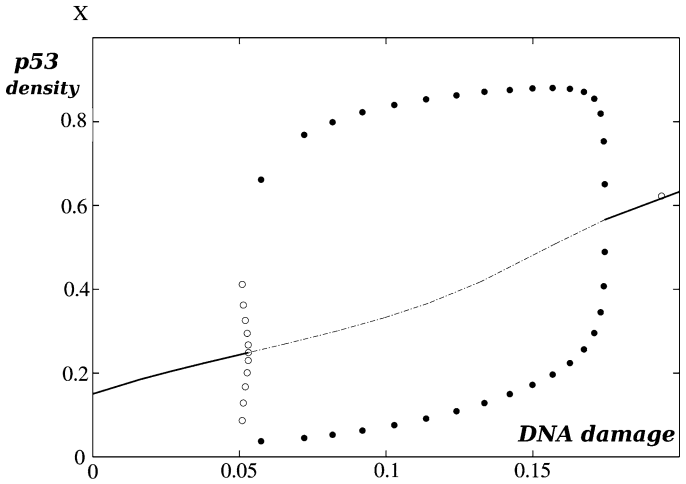

Fig. 4 Bifurcation diagram for our *toy model* for the p53 module. There is a sub-critical Hopf bifurcation with respect to γ, the parameter associated to DNA damage; *solid lines* indicate stable points; *dotted lines* indicate unstable points; *white circles* indicate maximum and minimum values of unstable orbits and *black circles* indicate the same for stable orbits; bifurcation diagram obtained with xpp-aut [26]

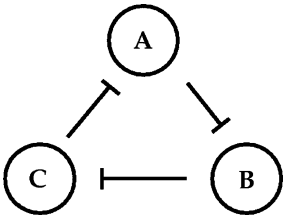

Fig. 5 Basic clock module based on a loop of feedback inhibition among three components indicated by A, B and C. *The blunt arrows* indicate inhibition

biological functions like, for instance, the required periodic adjustment of an organism's physiology to the circadian rhythm.

Basic clock The simplest cyclic-interaction module, which will be called *basic clock* module, has only three components that we call A, B and C. They interact cyclically through a loop of *feedback inhibition* where each component acts upon the next one along the loop so that A represses B, B represses C and C represses component A (see Fig. 5). More precisely, the rate of change of component A density at each time depends only on C density in an inhibitory manner: the density of A tends to decrease if the concentration of C is high. In a biological setting, A and C could represent, for instance, proteins with C being an enzyme that accelerates the rate of degradation of A. A similar dependence holds between C and B and between B and A. For large inhibition strengths, dynamical instabilities would be expected where concentrations of all three components oscillate. A precise general definition is given in Sect. 6.

Repressilator A slightly more complex version of the cyclic-interaction module is known in the literature as *repressilator*. It was implemented (that is, biochemically engineered) in the bacteria Escherichia coli [23] and is a milestone in *synthetic biology*. Three genes, that for simplicity sake we will call a, b and c, instead of their biological identifications λcl, LacI

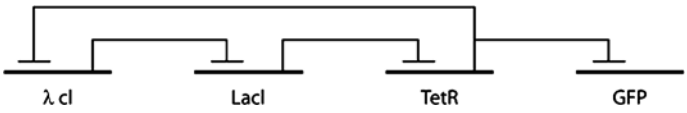

Fig. 6 A synthetic biology oscillatory network of transcriptional regulators [2] engineered in *E. coli*. *Blunt arrows* indicate inhibition through the gene regulation process described in the text

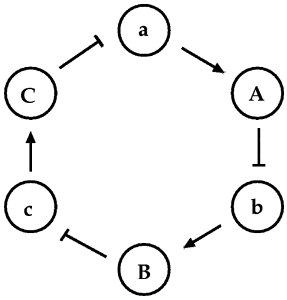

Fig. 7 Cyclic-interaction module with three genes, indicated by a, b, and c, whose product proteins, indicated respectively by A, B and C, repress the transition of the following gene. *Pointed arrows* indicate activation while *blunt arrows* indicate inhibition

and TetR (check Fig. 6), were introduced into the genome of the bacteria so that the corresponding proteins, denoted by A, B and C, interact in an cyclic inhibition loop. The fourth component, indicated by GFP in Fig. 6, plays no role in the dynamics and serves only as a fluorescent marker [17] for the concentration of TetR. The directed inhibition loop between components A, B and C arises from gene regulation as follows. Each gene, since it is translated into the corresponding protein, can be said to *promote* that protein; thus, a large density of a given gene in the cell indicates that the density of the corresponding protein tends to increase. On the other hand, each protein *inhibits* the production of the gene associated to the next protein in the same loop as before. That is, if a, b and c denote genes and A, B and C the corresponding proteins, protein A inhibits the production of gene b while protein B inhibits gene c and protein C inhibits gene a. See Fig. 7 for a graphical representation of these interactions. As before, if interaction strengths are large enough, oscillations are expected to occur. Indeed, these oscillations were experimentally verified in the synthetic biology system in *E. coli* engineered to have these extra genes. Our toy model (Theorem 6.1 below) explains this dynamical behaviour through a (supercritical) Hopf bifurcation with respect to the interaction strength parameter.

Modeling approach for the basic clock model Let us first discuss informally our basic mathematical modeling approach for the three component cyclic-interaction module that leads to what we call the *basic clock* model. The general case, which includes the repressilator model, will be considered later. Suppose the three types of biochemical components of this simple module interact inside some cellular container where N sites, labelled from 1 up to N, are available for each one of the types A, B and C. We write $\eta_t(i,n) = +1$ if there is a component of type $i \in \{A, B, C\}$ in site n, $1 \le n \le N$, at time t and write $\eta_t(i,n) = -1$ otherwise. The numbers $\eta_t(i,n)$ are thought as *spins* (up or down) for type i at site n in the usual nomenclature of statistical mechanics. For each $i \in \{A, B, C\}$ denote by $c(i)$ the type of the component that inhibits it, that is, $c(A) = C$, $c(B) = A$ and $c(C) = B$. (Remark: to simplify our notation and treat arbitrarily large numbers of types, below we label them with numbers rather than letters.) Now assume that each spin $\eta_t(i,n)$ flips with rates that depend only on the density at time t of the component $c(i)$. We write $\eta_t \in \{-1,+1\}^{\{A,B,C\}\times\{1,\ldots,N\}}$ for the global *configuration space* of the system at time t and denote by $c(i,n,\eta)$ the rate

with which the spin of type i at site n flips when the present configuration is η. Several choices of flip rates would be natural to represent this inhibition loop. A simple choice is given by

$$c(i,n,\eta) = \exp\left\{ J\eta(i,n)\frac{1}{N}\sum_{l=1}^{N}\eta\big(c(i),l\big)\right\} \tag{2.1}$$

where J is the parameter measuring inhibition strength. The region $J > 0$ is of interest for mimicking the inhibition loop, as the resulting rates favors spins of type i pointing opposite to the majority of spins of type $c(i)$. In the Interacting Particle Systems setting this stochastic model is a spin-flip model with mean-field Glauber dynamics. Due to the asymmetric nature of the interaction among the three types of spins the stochastic dynamics is no longer reversible and interesting new phenomena may occur.

In this paper we analyse the thermodynamic limit (that is, the limit as N goes to infinity) of these stochastic spin flip models and the main result is given by Theorem 4.1. In this limit, both the simplest cyclic-interaction module model ($k = 3$) and the repressilator model ($k = 6$) give rise to deterministic dynamical systems of type (1.1) with Hopf bifurcation (Theorem 6.1) at $J = 4$ and $J = 4\sqrt{3}/3$, respectively. Therefore these *toy models* are capable of exhibiting the qualitative behaviour observed for the associated real biological signaling pathway.

2.2 p53 Module

Let us now describe our second simple biological example, the *p53 module*. p53 is a tumor suppressor gene coding for a protein that plays an important role in maintaining cellular genomic integrity and preventing cancer. Recent experiments [24] verified that DNA damage elicits—in each single damaged cell—a number of pulses of p53 which have essentially the same amplitude and shape instead of, as expected originally, damped oscillations. This behavior can not be explained by the usual Hopf bifurcation, as is the case for the *repressilator* model. Several publications dealt with this problem and tried to understand the source of this dynamical behaviour via mathematical models. For instance, Zhang et al. [25], described four simple mechanisms by which p53 could be regulated and proposed mathematical models describing them, based on the usual chemical kinetics ideas and additional simplifying assumptions.

We consider here one of these mechanisms which involves positive and negative feedback loops between p53 and another type of molecule, called Mdm2, as described graphically in Fig. 3. As indicated in the figure, Mdm2 is a protein that has a dual role activation/inhibition with respect to p53: Mdm2 in the cytoplasm enhances translation of p53 RNA strings into p53 protein (thus activating p53) while Mdm2 in the nucleus mediates p53 degradation (thus inhibiting p53). Finally, Mdm2 transcription is induced by p53 itself. Thus three states are required to describe the different roles of Mdm2, say state 0 to indicate Mdm2 *not present* (or *degraded*), state 1 for Mdm2 *present in the cytoplasm* and state 2 for Mdm2 *present in the nucleus*. On the other hand, only two states (present/absent) are enough to describe p53.

The set of possible state transitions at each $n \in \{1,\dots,N\}$ is indicated in Fig. 8. Solid arrows indicate transitions that are compatible with the graphical description given by Fig. 3. For instance, the transition of Mdm2 from state 2 to state 0 is allowed as it corresponds to a molecule of Mdm2 at the nucleus being degraded while the reverse transition, from 2 to 0, is not allowed as Fig. 3 indicates that Mdm2 can only be produced in the cytoplasm, by RNA translation. Dotted arrows from a state of a given component terminating in a solid transition arrow indicate, in Fig. 8, that transition is regulated by the associated density. For instance,

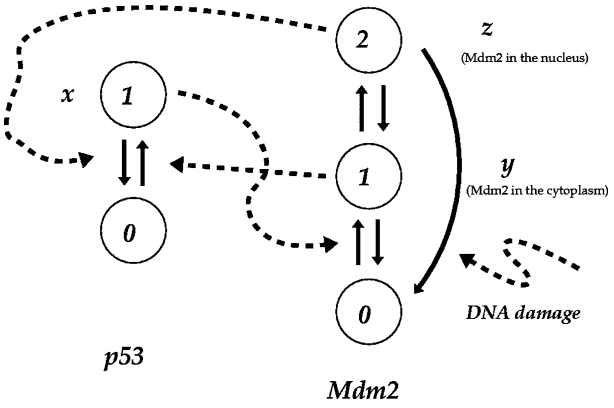

Fig. 8 Possible transitions corresponding to p53/Mdm2 interactions from Zhang et al. [25]; *solid arrows* indicate possible transitions while *dotted arrows* from a state to *a solid arrow* indicates that the rate of the target transition is regulated by the density of the initial state

the transition where p53 is produced, that is, goes from state 0 to state 1, is regulated by Mdm2 in state 1, that is, it depends on the density of Mdm2 in the cytoplasm.

Our heuristic prescription for setting up models The intuitive idea is simple: first we assume that each transition may depend on the density profile in such a way as to represent qualitatively the putative interactions. Next, the differential equations are formally derived from Kolmogorov's equations, as in (7.1).

Let us denote by x, y and z the densities of p53, Mdm2 in the cytoplasm and Mdm2 in the nucleus, respectively. The rates that are not regulated, for instance the transition corresponding to the degradation of Mdm2 (state 1 to state 0) are taken to be constants to be chosen later. The choices of these constants should be guided by the available biochemical information but we assume that their precise values are not crucially important. This *robustness* assumption with respect to parameter changes is justified by the fact that processes that depend on *fine tuning* would not be reliable in the extremely noisy cellular environment.

All three regulated rates indicated in Fig. 8 are qualitatively similar and correspond to *activation*. For instance, the rate of production of p53 depends on the density of Mdm2 in the cytoplasm in such a way that larger (smaller) densities of Mdm2 in the cytoplasm imply larger (smaller) rates of production of p53. To represent this we should assume that the rate of production of p53 is an increasing function of y, the density of Mdm2 in the cytoplasm. A simple choice for this function (related with (5.12)) is given by

$$S(u, V, \alpha, a) = \frac{V}{1 + e^{-\alpha(u-a)}} \tag{2.2}$$

where V, α and a are parameters.

The density-profile process along the lines of Fig. 3 is defined in the following way. There are two types, $\mathcal{T} = \{p, M\}$ (for p53 and Mdm2), with internal spin spaces $\mathcal{S}_p = \{0, 1\}$ and $\mathcal{S}_M = \{0, 1, 2\}$. The respective densities are

$$\begin{aligned} x_{(p,1)} &= x, \\ x_{(p,0)} &= 1 - x, \end{aligned} \qquad \begin{aligned} x_{(M,2)} &= z, \\ x_{(M,1)} &= y, \\ x_{(M,0)} &= 1 - y - z \end{aligned} \tag{2.3}$$

while the transition rates are

$$\begin{aligned} \lambda_p^{0\to 1} &= S(y, V_{yx}, \alpha_{yx}, a_{yx}), & \lambda_M^{1\to 2} &= q_{12}, \\ \lambda_p^{1\to 0} &= S(z, V_{zx}, \alpha_{zx}, a_{zx}), & \lambda_M^{1\to 0} &= q_{10}, \\ \lambda_M^{0\to 1} &= S(x, V_{xy}, \alpha_{xy}, a_{xy}), & \lambda_M^{2\to 1} &= q_{21}, \\ & & \lambda_M^{2\to 0} &= \gamma. \end{aligned} \tag{2.4}$$

Here q_{ij} indicates the (constant) rate of transition of Mdm2 from state i to state j and S is the function defined above. The parameter γ measures the *input* of the dynamical system, namely, the strength of DNA damage.

The resulting differential equations for the p53 are, therefore,

$$\begin{aligned} \dot{x} &= (1-x)S(y, V_{yx}, \alpha_{yx}, a_{yx}) - xS(z, V_{zx}, \alpha_{zx}, a_{zx}), \\ \dot{y} &= (1-y-z)S(x, V_{xy}, \alpha_{xy}, a_{xy}) - y(q_{12}+q_{10}) + zq_{21}, \\ \dot{z} &= yq_{12} - z(q_{21}+\gamma). \end{aligned} \tag{2.5}$$

The main *output* of interest, the response of the system, is the density of p53, given by x. The bifurcation diagram for this variable, with respect to γ is presented in Fig. 4. There is a sub-critical Hopf bifurcation at a particular value of γ, say γ_c, such that the concentration of p53 remains stable for $\gamma < \gamma_c$ but oscillates with large amplitude as soon as γ increases past this threshold. This diagram is in agreement with the findings of Zhang et al. [25] based on the usual approach based on chemical kinetics ideas.

3 Overview of the Paper

The density-profile process and our almost sure convergence result is presented in Sect. 4. Density-profile processes are random walk jump-processes in $\mathbb{R}^k$ with jumps of size $1/N$, whose expected drift velocity $V(x)$ does not depend on N. The main result of this paper (Theorem 4.1) is the proof that, within arbitrary but fixed time intervals, the paths of such a process converge almost surely to the trajectories of the dynamical system having V as the velocity field with explicit bounds for the distance between the stochastic and deterministic trajectories. As we show, the resulting dynamical systems can exhibit a very rich behavior, including oscillations.

The microscopic setup is presented in Sect. 5. We introduce the type-dependent stochastic spin model with non-reversible dynamics in Sect. 5.1. The mean-field version of this model, which is motivated by our biological interest and for which our convergence result applies, is also presented here.

Similar problems where considered in the literature [29]. It should be pointed out that ideas from the so called martingale problem and random time changes developed by T. Kurtz [30, 31] provide an alternative method to prove almost sure convergence for our models (see also [32] for analogous models and results in a discrete time setting). It is, however, not immediately clear to us whether these techniques also lead to an explicit control of the difference between the stochastic and deterministic trajectories similar to that stated in Theorem 4.1. The pathwise approach presented here, which strongly exploits coupling ideas [37], is quite natural within the Interacting Particle Systems setting we adopt. It would, for instance, also be natural to consider metastability issues [33]. One interesting problem, that

we leave for a further investigation, is the analysis of the density-profile fluctuations as bifurcations of the associated dynamical system are crossed.

In Sect. 6 we describe bifurcation results for the repressilator model. Finally, the main section of this paper, Sect. 7, presents the proof of our convergence result. The main mathematical steps in this proof are the following:

(i) A graphical construction (Sect. 7.2) that allows a coupled simultaneous construction of density-profile processes for different N.

(ii) An auxiliary process $\{\widehat{m}_t^{x^0,N}\}_{t\geq 0}$ (Sect. 7.1) defined through a simple spin-flip model (independent flips with time-dependent rates) which shadows the deterministic dynamical system (Lemma 7.1).

(iii) A coupling between the auxiliary and the density-profile processes that keeps both processes close to each other (Sect. 7.4). Instants where they may move further apart define a process of *discrepancies*. Bounds on the rate of these discrepancies yield our convergence theorem (Theorem 4.1).

4 Convergence of Density Profile Process to Dynamical Systems

In this section we define our basic processes and state the main mathematical result of the paper. In next section we shall discuss how this density-profile process can be realized through the dynamics of stochastic Ising models.

4.1 Density-profile Processes $\{m_t^{x^0 N}\}_{t\geq 0}$

The general set-up is as follows. There is a finite family of *types* (components, elements, molecules) $\mathcal{T}$ with cardinal $|\mathcal{T}| = k$. Each type i has an associated space of *internal states* $\mathcal{S}_i = \{a_1, \ldots, a_{s_i}\}$. We denote

$$\mathcal{E} = \left\{\underline{i} = (i,a) : i \in \mathcal{T}, a \in \mathcal{S}_i\right\}. \tag{4.1}$$

A density-profile process on $\mathcal{E}$ is a continuous-time jump-process in the hypercube $\mathcal{D}_N = (-\frac{1}{N}, 1 + \frac{1}{N})^{\mathcal{E}}$, for $N \geq 1$. At each jump, a point $x \in \mathcal{D}_N$ changes two coordinates simultaneously—one by $1/N$ and the other by $-1/N$—, both corresponding to the same type . The rates of these transitions depend smoothly on x and are defined in the following way. We start with a family of matrices of bounded Lipschitz functions $(\lambda_i^{a\to b})_{a,b\in\mathcal{S}_i}$, one for each type $i \in \mathcal{T}$, with each $\lambda_i^{a\to b} : \mathbb{R}^{\mathcal{E}} \longrightarrow \mathbb{R}_+^{\mathcal{E}}$ and $\lambda_i^{a\to a} = 0$. They, in turn, define functions $f_i^{a\to b} : \mathbb{R}^{\mathcal{E}} \longrightarrow \mathbb{R}_+^{\mathcal{E}}$ through the relations

$$f_i^{a\to b}(x) = x^\bullet_{(i,a)} \lambda_i^{a\to b}(x^\bullet) \tag{4.2}$$

for $i \in \mathcal{T}$, $a, b \in \mathcal{S}_i$, where

$$x^\bullet_{\underline{i}} = \begin{cases} x_{\underline{i}} & \text{if } 0 \leq x_{\underline{i}} \leq 1, \\ 0 & \text{if } x_{\underline{i}} < 0, \\ 1 & \text{if } x_{\underline{i}} > 1 \end{cases} \tag{4.3}$$

and $(x^\bullet)_{\underline{j}} = x^\bullet_{\underline{j}}$.

A density-profile process $\{m_t^{x^0 N}\}_{t\geq 0}$ is a random-walk process in $\mathcal{D}_N$ which starts at x^0 and evolves in continuous time through jumps of the form

$$x \longrightarrow x - \frac{e_{(i,a)}}{N} + \frac{e_{(i,b)}}{N} \tag{4.4}$$

where $e_{\underline{i}}$ denotes the unit vector along direction $\underline{i}$. From each position $x \in \mathcal{D}_N$ such a transition occurs with rate $Nf_i^{a\to b}(x)$, that is,

$$Nf_i^{a\to b}(x) = \frac{d}{dt} P\left(m_t^{xN} = x - \frac{e_{(i,a)}}{N} + \frac{e_{(i,b)}}{N} \right)\Big|_{t=0}. \tag{4.5}$$

In our applications, each variable $x_{(i,a)}$ represents the density of objects of type i that have internal state a, and the function $\lambda_i^{a\to b}$ is the rate at which an object of type i changes its state from a to b. Thus, $Nf_i^{a\to b}$ is the total rate of transitions $a \to b$ for all the N objects of type i present at the different sites. As we are fixing the total number N of particles, we shall assume that the initial conditions x^0 belong to

$$\mathcal{H}_{\mathcal{E}} = \left\{ x \in [0,1]^{\mathcal{E}} : \sum_{a\in\mathcal{S}_i} x_{(i,a)} = 1, i \in \mathcal{T} \right\}. \tag{4.6}$$

In the sequel we assume the choice of a common probability space where the density-profile processes (and other auxiliary processes defined below) are simultaneously realized for all N and x^0. The corresponding probability measure will be denoted P. (The graphical constructions introduced in Sect. 7 offer, in fact, a concrete way of defining this common space.)

4.2 Convergence to a Dynamical System $\{x_t^{x^0}\}_{t\geq 0}$

Let $\{m_t^{x^0 N}\}_{t\geq 0}$ be the density-profile process in $\mathcal{D}_N$ defined as above for appropriate functions $\lambda_i^{a\to b}$, and let $V : \mathbb{R}^{\mathcal{E}} \longrightarrow \mathbb{R}_+^{\mathcal{E}}$ be its associated *drift velocity field*:

$$V(x) = \lim_{t\downarrow 0} \frac{\mathbf{E}(m_t^{xN} - x)}{t}, \tag{4.7}$$

that is,

$$V_{(i,b)}(x) = \sum_{a\in\mathcal{S}_i} x_{(i,a)}^{\bullet} \lambda_i^{a\to b}(x^{\bullet}) - x_{(i,b)}^{\bullet} \sum_{a\in\mathcal{S}_i} \lambda_i^{b\to a}(x^{\bullet}). \tag{4.8}$$

Let $\{x_t^{x^0}\}_{t\geq 0}$ be the solution of the dynamical system

$$\dot{x}_t = V(x_t) \tag{4.9}$$

starting at $x^0 \in \mathcal{H}_{\mathcal{E}}$. The global trajectory exists by the smoothness of the field V. Furthermore, the flow does not leave $[0,1]^{\mathcal{E}}$ if the initial condition is in $\mathcal{H}_{\mathcal{E}}$. Indeed, in this case, each $V_{\underline{i}}$ becomes strictly positive if $x_{\underline{i}}$ reaches 0 and strictly negative if $x_{\underline{i}}$ reaches 1.

The main result of this paper is the convergence of the sequence of density profile processes $(m_t^{x^0 N})_N$ to the trajectory $x_t^{x^0}$. For $\epsilon > 0$ let τ_ϵ^N be the stopping time

$$\tau_\epsilon^N = \inf\left\{ t \geq 0 : |m_t^{x^0 N} - x_t^{x^0}| > \frac{1}{N^{\frac{1}{2}-\epsilon}} \right\} \tag{4.10}$$

where $|x|$ indicates the ℓ^1-norm of x in $\mathbb{R}^{\mathcal{E}}$,

$$|x| = \sum_{n\in\Lambda} \sum_{\underline{i}\in\mathcal{E}} \left| x_{(\underline{i},n)} \right|. \tag{4.11}$$

For a given T, $0 \leq T < \infty$, let us write

$$\mathcal{A}^T_{N\epsilon} = \{\tau^N_\epsilon < T\}. \tag{4.12}$$

The following is our main result.

Theorem 4.1 *Let $\lambda_i^{a\to b}$, $a, b \in S_i$, $i \in \mathcal{T}$ be bounded functions from $\mathbb{R}^{\mathcal{E}}$ to $\mathbb{R}^{\mathcal{E}}_+$ satisfying the Lipschitz condition*

$$\left|\lambda_i^{a\to b}(x) - \lambda_i^{a\to b}(y)\right| \leq K\,|x-y| \tag{4.13}$$

for some $K > 0$, all $x, y \in [0,1]^{\mathcal{E}}$ and all $i \in \mathcal{T}$ and $a, b \in S_i$. Then, for any finite T, initial position $x^0 \in \mathcal{H}_{\mathcal{E}}$ and $\epsilon > 0$,

$$P\left(\overline{\lim}_N \mathcal{A}^T_{N\epsilon}\right) = 0. \tag{4.14}$$

That is, for typical realizations, there exists some $N_{\epsilon,T}$ such that for $N > N_{\epsilon,T}$ each process $\{m_t^{x^0 N}\}_{t\geq 0}$ stays within a distance $N^{-1/2+\epsilon}$ of the deterministic path $\{x_t^{x^0}\}_{t\geq 0}$ at least up to time T.

Dynamical systems of the form (4.9)/(4.8) can exhibit quite complex dynamics—even for simple choices of the rates λ—, including stable orbits and chaotic behavior.

5 Type-dependent Stochastic Models and Density-profile Processes

5.1 Type-dependent Stochastic Spin Models

We define here a family of stochastic spin-flip models which extends the usual definition [27] to allow asymmetric dependence of rates on the energy function, the *Hamiltonian*. This family of models—that we call *type-dependent* stochastic models—have mathematical interest in themselves. Given our biological motivation though, our main interest here will be on the *mean-field* type-dependent model which provides our basic theoretical framework for generating useful models for signaling pathways.

We consider a finite set Λ of *(spatial) positions*, at each point of which there is a copy of the set-up of Sect. 4.1, namely a family $\mathcal{T}$, of cardinal k, of *spin types*; each type $i \in \mathcal{T}$ having available a set of *internal states* $\mathcal{S}_i = \{a_1, \ldots, a_{s_i}\}$. We call *sites* the pairs formed by a type and a spatial position. Thus in the usual statistical mechanics notation, our spin system has site-space $V = \mathcal{T} \times \Lambda$ and configuration space $\Sigma = \prod_{i\in\mathcal{T}} \mathcal{S}_i^\Lambda$. For a configuration $\eta \in \Sigma$, the value of a spin at site (i, n)—that is, of the spin of type i at position n—is denoted $\eta(i, n)$.

For simplicity sake we assume here that only single-spin flips (more precisely, single-site internal-state transitions) are allowed in our continuous-time stochastic dynamics. The corresponding rates are determined in terms of a function $H : \Sigma \to \mathbb{R}$, the *Hamiltonian* or energy function of the spin system, that we define presently.

The Hamiltonian is determined by a family of *interaction matrices* $\mathbb{I}_{n,\ell} : \mathcal{E} \times \mathcal{E} \to \mathbb{R}$, one for each pair of spatial sites $n, \ell \in \Lambda$. Due to the applications in sight, we do not assume the matrices $\mathbb{I}_{n,\ell}[\cdot;\cdot]$ to be symmetric. Rather, the quantity $\mathbb{I}_{n,\ell}[(i,a);(j,b)]$ will indicate the strength of the influence that a spin at a site $(i, n) \in V$ in internal state $a \in \mathcal{S}_i$ has *upon* a spin at site $(j, \ell) \in V$ that is in internal state $b \in \mathcal{S}_j$. This influence may be noticeably

different that the one in the opposite direction (e.g. activation vs. deactivation). Asymmetries associated with types are natural in our modeling setup, as illustrated in Sect. 2. For instance, in our *basic clock* example, presented there, type A components act upon type B components while the reciprocal interaction does not occur.

The Hamiltonian is then defined by

$$H(\eta) = \sum_{(i,n)\in V} H_{(i,n)}(\eta) \tag{5.1}$$

with

$$H_{(i,n)}(\eta) = -\sum_{(j,\ell)\in V} \mathbb{I}_{\ell,m}[(j,\eta(j,\ell));(i,\eta(i,n))]. \tag{5.2}$$

Thus, $H_{(i,n)}(\eta)$ describes the collective influence of the global configuration η *upon* the spin at (i,n) which is in an internal state $\eta(i,n)$.

The Hamiltonian (5.1)–(5.2) generalizes usual Potts interactions in statistical mechanics on two counts: First, the coupling parameters $\{\mathbb{I}_{n,\ell}\}_{n,\ell\in\Lambda}$ depend on the internal spin states, and second and more importantly, *we do not assume interactions to be symmetric*. We restrict, however, our model by assuming that the asymmetry of the interaction is only associated to types, that is

$$\mathbb{I}_{n,\ell}[(i,a);(j,b)] = \mathbb{I}_{\ell,n}[(i,a);(j,b)] \tag{5.3}$$

for $n,\ell\in\Lambda$, $i,j\in\mathcal{T}$, $a\in\mathcal{S}_i$ and $b\in\mathcal{S}_j$.

Given a configuration $\eta\in\Sigma$, a site (i,n) and an internal state $a\in\mathcal{S}_i$ let us denote $\eta^a_{(i,n)}$ the configuration with

$$\big[\eta^a_{(i,n)}\big](j,\ell) = \begin{cases} a & \text{if } (j,\ell)=(i,n),\\ \eta(j,\ell) & \text{otherwise.}\end{cases} \tag{5.4}$$

Now consider the energy cost $\Delta^{a\to b}_{(i,n)}(\eta)$ of the transition $\eta^a_{(i,n)}\to\eta^b_{(i,n)}$, given by

$$\Delta^{a\to b}_{(i,n)}(\eta) = H(\eta^b_{(i,n)}) - H(\eta^a_{(i,n)}). \tag{5.5}$$

The asymmetry of the interaction leads naturally to the decomposition

$$\Delta^{a\to b}_{(i,n)}(\eta) = \Delta[\,{}^{\text{upon}}_{\text{self}}\,]^{a\to b}_{(i,n)}(\eta) + \Delta[\,{}^{\text{unto}}_{\text{others}}\,]^{a\to b}_{(i,n)}(\eta). \tag{5.6}$$

The term

$$\Delta[\,{}^{\text{upon}}_{\text{self}}\,]^{a\to b}_{(i,n)}(\eta) = \sum_{(j,\ell)\in V}\big[\mathbb{I}_{\ell,n}[(j,\eta(j,\ell));(i,b)] - \mathbb{I}_{\ell,n}[(j,\eta(j,\ell));(i,a)]\big] \tag{5.7}$$

collects the change in the influence of the configuration η upon the site (i,n) when the internal state there changes from a to b. On the other hand,

$$\Delta[\,{}^{\text{unto}}_{\text{others}}\,]^{a\to b}_{(i,n)}(\eta) = \sum_{(j,\ell)\in V}\big[\mathbb{I}_{\ell,n}[(i,b);(j,\eta(j,\ell))] - \mathbb{I}_{\ell,n}[(i,a);(j,\eta(j,\ell))]\big] \tag{5.8}$$

collects the change of the influence that the site (i,n) has on all other sites when its internal state jumps from a to b.

For the dynamics we adopt what can be labeled a *heat-bath* prescription: each transition rate depends only on the energy changes brought upon the site; the rest of the configuration acts as a "heat reservoir" that remains unperturbed. Formally, the rate $c^{a\to b}_{(i,n)}(\eta)$ of a transition flipping $\eta^a_{(i,n)}$ to $\eta^b_{(i,n)}$ depends only on $\Delta[\,{}^{\text{upon}}_{\text{self}}\,]^{a\to b}_{(i,n)}(\eta)$,

$$c^{a\to b}_{(i,n)}(\eta) = \Phi\big(\Delta[\,{}^{\text{upon}}_{\text{self}}\,]^{a\to b}_{(i,n)}(\eta)\big) \tag{5.9}$$

where Φ is a non-increasing $\mathbb{R}_+$-valued function satisfying the "detailed-balance" condition

$$\Phi(E)e^{E} = \Phi(-E)e^{-E}. \tag{5.10}$$

The natural choices are the usual heat-bath rates

$$\Phi(E) = e^{-E}, \tag{5.11}$$

or the *Metropolis form*

$$\Phi(E) = e^{-2[E]_+} \tag{5.12}$$

where $[x]_+ = \max\{0, x\}$ is the positive part of $x \in \mathbb{R}$.

Let us formalize our definitions.

Definition 5.1

(i) The prescription contained in (5.1), (5.2) and (5.3) is a *type-dependent interaction*.
(ii) A *type-dependent stochastic spin model* is the continuous-time process defined by a spin model with a type-dependent interaction and a dynamics with rates (5.9).

Let us comment that in usual stochastic spin models the transition rates depend rather on the total change of energy associated to the flip, in the form

$$c^{a\to b}_{(i,n)}(\eta) = \Phi\big(\Delta^{a\to b}_{(i,n)}(\eta)\big). \tag{5.13}$$

This total energy cost includes, thus, also the change of energy brought to other sites by the transition at (i,n). If the matrices $\mathbb{I}_{n,\ell}$ are symmetric, $\Delta^{a\to b}_{(i,n)}(\eta) = 2\Delta[\,{}^{\text{upon}}_{\text{self}}\,]^{a\to b}_{(i,n)}(\eta)$ and both definitions become equivalent [modulo a factor 2 in the argument of Φ in (5.9)]. They yield the so called *stochastic spin models* [27] which are reversible with respect to the Gibbs measure

$$\mu(\eta) = \frac{e^{-H(\eta)}}{\sum_{\xi\in\Sigma} e^{-H(\xi)}}. \tag{5.14}$$

In the asymmetric case, however, there is an important difference. With rates given by (5.13) the symmetry or asymmetry of the couplings plays no role and the process is always reversible with respect to the Gibbs measure (5.14). With our rates (5.9), the type-dependent interaction process with asymmetric couplings is no longer reversible for the Gibbs measure (5.14). This dynamics, however, is the one leading to the right Kolmogorov equations (7.1) and hence to the right dynamical system (4.9)–(4.8). The lack of reversibility may lead to a *frustrated dynamics* as illustrated in Sect. 6 below for a mean-field model.

The simplest type-dependent stochastic spin model, which is nevertheless important both for mathematical and biological reasons, is when all internal spin space are binary. Two standard choices are the *Ising spins* $\mathcal{S}_i = \{-1, +1\}$, representing deactivation/activation, and the *lattice-gas variables* $\mathcal{S}_i = \{0, 1\}$, representing absence/presence. The Hamiltonian for Ising spins may be written as

$$H_{(i,n)}(\eta) = \eta(i,n)\Big[\sum_{(j,\ell)\in V} J_{(n,\ell)}(i,j)\eta(j,\ell) + h(i,n)\Big] \tag{5.15}$$

which corresponds to the choice of interaction matrix given by

$$\mathbb{I}_{n,\ell}[(i,a);(j,b)] = aJ_{n,\ell}(i,j)b \tag{5.16}$$

for $(i,n) \neq (j,\ell)$ and

$$\mathbb{I}_{n,n}[(i,a);(i,a)] = ah(i,n) \tag{5.17}$$

for $a \in \{-1,+1\}$ and $i, j \in \mathcal{T}$.

Definition 5.2 A *type-dependent stochastic Ising model* is a continuous-time process defined by:

(i) A configuration space $\Sigma = \{-1,+1\}^{\mathcal{T}\times\Lambda}$.
(i) A *type-dependent Ising interaction* defined through the prescription (5.1), (5.15) and (5.3).
(ii) A dynamics with rates (5.11) or (5.12).

Given our motivation in biology we will focus on the following family of models.

Definition 5.3 A type-dependent stochastic spin model is *mean-field* if the Hamiltonian parameters in (5.2) are of the form

$$\mathbb{I}_{n,\ell}[(i,a);(j,b)] = \frac{\alpha_{(i,a),(j,b)}}{|\Lambda|} \tag{5.18}$$

where $\{\alpha_{\underline{i},\underline{j}}\}_{\underline{i},\underline{j}\in\mathcal{E}}$, is a real matrix .

5.2 Density Profile Processes Defined by Stochastic Spin Models

For each stochastic spin model, an *empirical density profile* of a configuration η, $m(\eta) = (m_{\underline{i}}(\eta)) \in \mathbb{R}_+^{\mathcal{E}}$, where

$$m_{(i,a)}(\eta) = \frac{|\{\ell \in \Lambda : \eta(i,\ell) = a\}|}{|\Lambda|} \tag{5.19}$$

for $i \in \mathcal{T}$, $a \in \mathcal{S}_i$.

The following proposition is immediate from the fact that the density profile process in this case satisfies the conditions of the previous section and thus Theorem 4.1 applies.

Proposition 5.4 *Let* $\{\sigma_t^{\eta_0}\}_{t\geq 0}$ *be the mean-field spin-flip process starting at the configuration* $\eta_0 \in \Sigma$ *with rates defined by* (5.9), (5.1) *and* Φ *given by* (5.11). *Then, if* $|\Lambda| = N$, *the*

density-profile process $m(\sigma_t^{\eta^0})$ *approximates, in the sense of Theorem* 4.1, *the dynamical system* (4.9)/(4.8), *defined by rates* λ *as in* (4.2) *given by*

$$\lambda_j^{a\to b}(x) = \exp\Bigg(\sum_{\substack{i\in\mathcal{T}\\ d\in\mathcal{S}_i}} x_{(i,d)}[\alpha_{(i,d),(j,a)} - \alpha_{(i,d),(j,b)}] + h_{(j,a)} - h_{(j,b)}\Bigg) \tag{5.20}$$

for $j \in \mathcal{T}$ *and* $a, b \in \mathcal{S}_j$.

In the Ising case, the only independent variables are the densities of activated types that we denote x_i. Replacing $x_{(i,1)} = x_i$, $x_{(i,-1)} = 1 - x_i$ in (5.20) we obtain that the Ising density profile process is defined by rates

$$\lambda_j^{-1\to 1}(x) = \exp\Bigg(\sum_{i\in\mathcal{T}} \alpha_{i,j}x_i + h_j\Bigg), \tag{5.21}$$

$$\lambda_j^{1\to -1}(x) = \exp\Bigg(-\sum_{i\in\mathcal{T}} \alpha_{i,j}x_i - h_j\Bigg). \tag{5.22}$$

where $\alpha_{i,j}$ and h_j are real numbers for $i, j \in \mathcal{T}$.

6 Bifurcation Results for the Cyclic-interaction Module

The *cyclic-interaction model* is an Ising stochastic model defined through a simple choice of the parameters (5.18), which nevertheless leads to interesting (deterministic) dynamical behavior in the thermodynamic limit. Two particular cases of these models, the *basic clock* module and the *repressilator*, were informally presented in the introduction. We think the spin types as points $\{1, \ldots, k\}$ on the circle and, for each $i \in \mathcal{T}$, we denote $c(i)$ the nearest-neighbor of i in the counter-clockwise direction. We assume that $\alpha_{ji} = 0$ unless $j = c(i)$ and, furthermore, that all non-zero terms in α have the same absolute value. That is,

$$\alpha_{ji} = \begin{cases} s_i J & \text{if } j = c(i),\\ 0 & \text{otherwise} \end{cases} \tag{6.1}$$

where $s_i \in \{-1, +1\}$ and $J > 0$. We also set $h_i = -J/2$, for $1 \le i \le k$. In this way, once the interaction signals $\{s_i\}_{i=1}^k$ are chosen, J is the only free parameter of the model.

If $s_i = 1$, the rate with which spins of type i flip from -1 to $+1$ ($+1$ to -1) [defined in (5.21) and (5.22), respectively], is an increasing (decreasing) function of $x_{c(i)}$, the density of spins $+1$ of type $c(i)$. Borrowing statistical mechanical nomenclature, we say that the interaction of spins of type $c(i)$ with those of type i is *ferromagnetic* [28]. In the biochemical context, where x_i measures the density of some chemical component i, this means that the component $c(i)$ *activates* the production of component i. On the other hand, if $s_i = -1$ the rate for a spin of type i to flip from -1 to $+1$ ($+1$ to -1), decreases (increases) as a function of $x_{c(i)}$, and the interaction of spins of type $c(i)$ with those of type i is said to be *anti-ferromagnetic*. In biochemical terms, the component $c(i)$ *inhibits* the production of component i. Figure 5, representing the basic clock module, graphically represents the situation with $k = 3$ and $s_i = -1$, $i \in \{1, 2, 3\}$ such that $\boldsymbol{\alpha}$ is given by

$$\boldsymbol{\alpha} = \begin{pmatrix} 0 & -J & 0\\ 0 & 0 & -J\\ -J & 0 & 0 \end{pmatrix}. \tag{6.2}$$

For the flipping rates given by (5.21) and (5.22) one easily checks that the dynamical system (4.9) associated to the cyclic-interaction model (6.1) is:

$$\dot{x}_i = e^{s_i J(x_{c(i)}-\frac{1}{2})} - x_i\left(e^{s_i J(x_{c(i)}-\frac{1}{2})} + e^{-s_i J(x_{c(i)}-\frac{1}{2})}\right) \tag{6.3}$$

for $1 \le i \le k$. If J is small, this system has a single stable equilibrium point at $(1/2, \dots, 1/2) \in \mathbb{R}^k$, whichever the choice of signs s_i. For larger J, the behavior of the dynamical system (6.3) crucially depends on whether the product of signals is positive or negative. If $\prod_{i=1}^k s_i = -1$—a *frustrated* model in statistical mechanical terms—there is no (global) density-profile where all pairs of types of spins minimize their mutual interaction. In the notation introduced by E. Sontag [35] the associated directed graph is not consistent. This system exhibits a Hopf bifurcation [18] as J exceeds a critical value, which depends on k. In the non-frustrated case, the model behaves as the *Curie-Weiss* model. Formally:

Theorem 6.1 *Consider the dynamical system* (6.3) *with* $k \ge 3$

(a) *If* $\prod_{i=1}^k s_i = 1$, *there is a bifurcation at* $J_c = 2$: *the fixed point* $(1/2, \dots, 1/2)$ *looses stability and two stable points appear for* $J > J_c$.
(b) *If* $\prod_{i=1}^k s_i = -1$, *there is a Hopf bifurcation at* $J_c = 2/\cos(\pi/k)$.

Proof Write $s = \prod_{i=1}^k s_i$. A simple computation shows that near $\mathbf{1/2} = (1/2, \dots, 1/2) \in \mathbb{R}^k$ the dynamical system (6.3) is close to $\dot{x} = A(x - \mathbf{1/2})$, where A is a $k \times k$ real matrix with eigenvalues $sJe^{\frac{2\pi l}{k}i} - 2$, $l = 0, 1, \dots, k-1$. The fixed point is stable if the real parts of all eigenvalues are strictly negative, and stability is lost when one of the real parts becomes positive. Thus, if $s = 1$ the fixed point $\mathbf{1/2}$ loses stability at $J_c = 2$ when the eigenvalue corresponding to $l = 0$ crosses the imaginary axis through the origin. On the other hand, if $s = -1$, the stability is lost when two eigenvalues, symmetric around the real axis, cross the imaginary axis. This occurs at $J_c = 2/\cos(\pi/k)$. □

Remark 6.2 For instance, for the basic clock model where $k = 3$ and all interactions are antiferromagnetic ($s_i = -1$ for $i = 1, 2, 3$), as indicated by Fig. 5, the dynamical system has stable orbits for $J > J_c = 4$. The convergence result, Theorem 4.1, implies that, within any finite time interval, the stochastic density-profile process evolves as close to this deterministic orbit as wished, for N sufficiently large.

7 Proof of the Convergence Theorem

7.1 The Auxiliary Process $\{\widehat{m}_t^{x^0,N}\}_{t\ge 0}$

To prove Theorem 4.1 we introduce an auxiliary stochastic spin model with independent spins flips but *time-dependent rates*.

Let $\Lambda = \{1, \dots, N\}$ and let $\{\eta_t(i,n) : (i,n) \in \mathcal{T} \times \Lambda\}_{t\ge 0}$, be Nk independent Markov chains. Each chain $\eta_t(i,n)$ has state space $\mathcal{S}_i$. Thus, for each $t \ge 0$, $\eta_t \in \Sigma = \prod_{i\in\mathcal{T}} S_i^\Lambda$ tells the internal state of each type $i \in \mathcal{T}$ at each site $n \in \Lambda$ at time t. Each Markov chain $\{\eta_t(i,n)\}_{t\ge 0}$ undergoes transitions from state a to state b with time-dependent rate $\lambda_i^{a\to b}(x_t^{x^0})$, where $\{x_t^{x^0}\}_{t\ge 0}$ is the solution of the dynamical system (4.9) from the initial position x^0. We initialize these chains with the uniform distribution on configurations η_0 with empirical density profile $m(\eta_0) = x^0$ [$m(\eta_t)$ is defined in (5.19)]. The total number of

spins of each type and internal state $\underline{i}$ is, thus, fixed and equal to $Nx^0_{\underline{i}}$, but the initial state at each site is chosen independently for each type. We denote $\{\widehat{m}_t^{x^0,N}\}_{t\geq 0}$ the corresponding density-profile process.

We observe that the chain at each $(i,n) \in \mathcal{T} \times \Lambda$ satisfies Kolmogorov's equation. Hence, for $p_t^n(i,a) = P(\eta_t(i,n) = a)$, we have

$$\dot{p}_t^n(i,b) = \sum_{a\in\mathcal{S}_i} p_t^n(i,a)\lambda_i^{a\to b}(x_t^{x^0}) - p_t^n(i,b)\sum_{a\in\mathcal{S}_i}\lambda_i^{b\to a}(x_t^{x^0}). \tag{7.1}$$

Therefore each function $t \to p_t^n(\underline{i})$ is a solution of the differential equation (4.9) with V as in (4.8). Hence

$$p_0^n(\underline{i}) = (x^0)_{\underline{i}} \quad \Longrightarrow \quad p_t^n(\underline{i}) = (x_t^{x^0})_{\underline{i}} \quad \forall t \geq 0, \tag{7.2}$$

for all $\underline{i} \in \mathcal{E}$ and $n \in \Lambda$. While (7.2) is true for the auxiliary process $\{\widehat{m}_t^{x^0,N}\}$, we are interested in following the actual empirical densities. The next lemma proves that also the path followed by these densities remains, at all times, close to the trajectories of the dynamical system.

Lemma 7.1 *For $\delta > 0$ there exists $c > 0$ such that*

$$P\left(\left|\widehat{m}_t^{x^0,N} - x_t^{x^0}\right| > N^{\delta-1/2}\right) < \exp(-cN^\delta) \tag{7.3}$$

for $t \geq 0$.

Proof Let us introduce yet another auxiliary process, denoted $\{\widehat{m}_t^{B(x^0),N}\}_{t\geq 0}$, defined exactly as $\{\widehat{m}_t^{x^0,N}\}_{t\geq 0}$ but with initial spins chosen independently for each pair $(i,n) \in \mathcal{T} \times \Lambda$ with $P(\eta_0(i,n) = a) = (x^0)_{(i,a)}$. Hence, the density vectors $m_i(\eta)$ are independent for different types $i \in \mathcal{T}$, and for each $a \in \mathcal{S}_i$ each $[m_i(\eta)]_a$ has a binomial distribution with parameters N and $p_{(i,a)} = (x^0)_{(i,a)}$. (This means that, for large N, $\widehat{m}_t^{B(x^0),N}$ starts at a random position in $\mathcal{H}_\mathcal{E}$ close to x^0, while $\widehat{m}_t^{x^0,N}$ starts precisely at x^0.)

This new auxiliary process also satisfies (7.2) but has the advantage that the spin chains for different types and sites remain independent at all times, and, by (7.2), the expected proportion of spins of each type having each internal state coincide with the components of $x_t^{x^0}$. We may therefore describe $N\widehat{m}_t^{B(x^0),N}$ as an array $N(\widehat{m}_t^{B(x^0),N})_{i\in\mathcal{T}}$ of independent multinomial vectors

$$N\left(\widehat{m}_t^{B(x^0),N}\right)_i \sim \text{Mult}(N, (x_t^{x^0})_i) \tag{7.4}$$

for fixed $t \geq 0$ and $i \in \mathcal{T}$, where $(x_t^{x^0})_i = ((x_t^{x^0})_{i,a})_{a\in\mathcal{S}_i}$. Thus each coordinate of each vector is a binomial random variable, i.e.,

$$N\left(\widehat{m}_t^{B(x^0),N}\right)_{\underline{i}} \sim \text{Bin}(N, (x_t^{x^0})_{\underline{i}}) \tag{7.5}$$

for each $t \geq 0$ and $\underline{i} \in \mathcal{E}$; in particular its variance is bounded above uniformly by $1/N$. Now the large-deviation properties of binomial distributions [36] imply that for any $\delta > 0$ there exists a constant c such that

$$P\left(\left|\widehat{m}_t^{B(x^0),N} - x_t^{x^0}\right| > \frac{1}{2}N^{\delta-1/2}\right) < \exp(-cN^\delta) \tag{7.6}$$

for any $t \geq 0$.

To conclude the proof of the lemma we must show that both auxiliary processes $\widehat{m}_t^{B(x^0),N}$ and $\widehat{m}_t^{x^0,N}$ remain close to each other. This is more easily done through a coupling argument [27, 37]. We construct a coupled realization $(\eta_t^{B(x^0),N}, \eta_t^{x^0,N})$ of the spin systems defining both processes as follows. Spins in both systems flip with the same time-dependent rates given in (7.1). At coordinates $(\underline{i}, n)$ with $\eta_0^{B(x^0),N}(\underline{i}, n) = \eta_0^{x^0,N}(\underline{i}, n)$, the spins evolve together. Otherwise, the spins for both processes evolve independently until transitions bring them to a common value. They evolve together ever after. The distance between the corresponding coupled density profiles equals $2/N$ times the number of pairs $(\underline{i}, n)$ where $\eta_0^{B(x^0),N}(\underline{i}, n) \neq \eta_0^{x^0,N}(\underline{i}, n)$. The coupled construction causes this number to decrease with time, thus

$$\begin{aligned} \left| m\big(\eta_t^{B(x^0),N}\big) - m\big(\eta_t^{x^0,N}\big) \right| &\leq \left| m\big(\eta_0^{B(x^0),N}\big) - m\big(\eta_0^{x^0,N}\big) \right| \\ &= \left| m\big(\eta_0^{B(x^0),N}\big) - x^0 \right| \end{aligned} \tag{7.7}$$

and, therefore,

$$\begin{aligned} P\Big(\big| \widehat{m}_t^{x^0,N} - x_t^{x^0} \big| > N^{\delta-1/2} \Big) &\leq P\Big(\big| \widehat{m}_t^{x^0,N} - \widehat{m}_t^{B(x^0),N} \big| > \frac{1}{2} N^{\delta-1/2} \Big) \\ &\quad + P\Big(\big| \widehat{m}_t^{B(x^0),N} - x_t^{x^0} \big| > \frac{1}{2} N^{\delta-1/2} \Big). \end{aligned} \tag{7.8}$$

To prove (7.3) we bound the right-hand side using (7.7) and (7.6) (for $t = 0$) for the first term and again (7.6) for the second one. □

To prove Theorem 4.1 we will show that, for N large, $\widehat{m}_t^{x^0,N}$ and $m_t^{x^0,N}$ remain close within arbitrary finite time intervals. To achieve this we will couple both stochastic evolutions with the help of a graphical construction.

7.2 Graphical Construction: The Process $\{g_t^{x^0,N}\}_{t\geq 0}$

We resort to a graphical construction of density-profile processes with different N through time-rescaling of auxiliary processes $\{g_t^{x^0,N}\}_{t\geq 0}$. The latter is defined through paths determined by Poissonian "marks". This construction will be adapted in next section to couple the processes $\widehat{m}_t^{x^0,N}$ and $m_t^{x^0,N}$.

To each $y \in \mathcal{D}_N$ and $i \in \mathcal{T}$, we associate $|\mathcal{S}_i|(|\mathcal{S}_i| - 1)$ independent Poisson processes: $N_t^{i,a\to b}(y)$, $i \in \mathcal{T}$, $a, b \in \mathcal{S}_i$, $a \neq b$, with respective rates $f_i^{a\to b}(y)$. We associate a particular type of mark for the events of each type of process and place these marks along the time axis of y. A Poisson mark associated to the process $N_t^{i,a\to b}(y)$ carries the instruction to add $1/N$ at coordinate (i, b) and subtract the same amount at (i, a).

The process $\{g_t^{x^0,N}\}_{t\geq 0}$ is defined by *open paths* in $\mathcal{D}_N \times \mathbb{R}_+$ determined by the marks. These are piecewise linear curves that move along the positive time axis until a Poisson mark is met. At these times the trajectory moves by $\pm 1/N$ along two coordinate directions according to the type of mark. The process $\{g_t^{x^0,N}\}_{t\geq 0}$ is at position x at time t if there exists an open path from $(x^0, 0)$ to (x, t).

We see that the evolution of $\{g_t^{x^0,N}\}$ differs from that of $\{m_t^{x^0,N}\}$ only in that the rates of the latter [given by (4.5)] are N times those of the former. Thus, one process can be constructed from the other by a simple change in the time scale:

$$m_t^{x^0,N} = g_{Nt}^{x^0,N}. \tag{7.9}$$

In words: a *density-profile time* t corresponds to a *graphical-construction time* Nt.

7.3 Main Coupling and Discrepancy Process

We now use the graphical-construction strategy to produce coupled realizations of the density-profile processes $m_t^{x^0,N}$ and $\widehat{m}_t^{x^0,N}$ with an appropriate control of their distance. Our coupling forces both processes to keep their relative distance as much as possible, evolving as a rigid system. Of course, since their rates are not equal, they will make occasional asynchronous moves that may take them increasingly apart with the passing of time. The coupling is designed so to control this asynchrony.

The coupling involves a number of Poissonian mark processes at different sites which are updated every time there is an asynchronous move. The successive times of these moves correspond to a sequence of stopping times $\{\tau_n\}_{n\geq 1}$; the coupling is defined in a recursive fashion within successive time intervals $[\tau_{n-1}, \tau_n)$, $n \geq 1$. The auxiliary processes, which arise directly from such graphical coupled construction will be denoted, respectively, by $g_t^{x^0,N}$ and $\widehat{g}_t^{x^0,N}$. They differ from the density profiles $m_t^{x^0,N}$ and $\widehat{m}_t^{x^0,N}$ only in the time scale, which in the graphical construction is slower by a factor N.

Initial stage of the coupling Initially, $g_0^{x^0} = \widehat{g}_0^{x^0} = x^0$ and up to the first stopping time τ_1 (to be defined) we couple them through what is known as *basic coupling* in particle systems. For each $y \in \mathcal{D}_N$ and coordinate direction (i,a) we define $3(|\mathcal{S}_i| - 1)$ Poissonian mark processes $\widehat{N}_t^{i,a\to b}(y)$, $\widehat{E}_t^{i,a\to b,m}(y)$ and $\widehat{E}_t^{i,a\to b,\widehat{m}}(y)$, $b \neq a$, with respective rates

$$\begin{aligned}
\widehat{u}_t^{\,i,a\to b}(y) &= \min\big\{y_{(i,a)}\lambda_i^{a\to b}(y),\, y_{(i,a)}\lambda_i^{a\to b}(x_{t/N}^{x^0})\big\},\\
\widehat{e}_t^{\,i,a\to b,m}(y) &= y_{(i,a)}\lambda_i^{a\to b}(y) \;-\widehat{u}_t^{\,i,a\to b} \quad \text{and}\\
\widehat{e}_t^{\,i,a\to b,\widehat{m}}(y) &= y_{(i,a)}\lambda_i^{a\to b}(x_{t/N}^{x^0}) - \widehat{u}_t^{\,i,a\to b}.
\end{aligned} \tag{7.10}$$

Note the rescaling in time for the deterministic path $\{x_t^{x^0}\}$ needed to represent it on the graphical construction time scale.

As before, we think that occurrence of each of these processes are associated to particular marks indicating where to jump. The jumps of the process $\{g_t^{x^0,N}\}$ occur at the marks of $\{\widehat{E}_t^{i,a\to b,m}(y)\}$ and $\widehat{N}_t^{i,a\to b}(y)$, while those of the process $\{\widehat{g}_t^{x^0,N}\}$ are at $\{\widehat{E}_t^{i,a\to b,\widehat{m}}(y)\}$ and $\widehat{N}_t^{i,a\to b}(y)$. The marks of the processes $\{\widehat{E}_t^{i,a\to b,m}(y)\}$ and $\{\widehat{E}_t^{i,a\to b,\widehat{m}}(y)\}$ are thus seen by only one of $\{g_t^{x^0,N}\}$ or $\{\widehat{g}_t^{x^0,N}\}$ and will be called (potential) *discrepancies*. The *basic Poisson processes* $\{\widehat{N}_t^{i,a\to b}(y)\}$, on the other hand, are introduced to ensure that $\{g_t^{x^0,N}\}$ and $\{\widehat{g}_t^{x^0,N}\}$ remain equal until they find the first discrepancy. This defines a stopping time τ_1 at which the processes get separated by a distance of $1/N$. At this time we can not continue using the basic coupling.

Formally, we define a *first-discrepancy process*

$$D_t^0 = \sum_{i\in\mathcal{T}} \sum_{\substack{a,b\in\mathcal{S}_i\\ a\neq b}} \Big[\widehat{E}_t^{i,a\to b,m}(y^0) + \widehat{E}_t^{i,a\to b,\widehat{m}}(y^0)\Big] \tag{7.11}$$

where y^0 is the density-profile path defined by the preceding (level-0) construction. The *first discrepancy time* τ_1 is the time of the first event of this process. A new coupling definition must be introduced at this time, which will be applied until the second discrepancy time τ_2. This iterative procedure continues up to the time T chosen in Theorem 4.1. We now present the recursive step in the definition of such a coupling.

l-th stage of the coupling Suppose that the graphical construction has been defined up to time τ_l, $l \geq 1$, determining x^l, $\Delta^l \in \mathcal{D}_N$ such that

$$g_{\tau_l}^{x^0,N} = x^l, \qquad \widehat{g}_{\tau_l}^{x^0,N} = x^l + \Delta^l. \tag{7.12}$$

[Thus, $m_{\tau_l/N}^{x^0,N} = x^l$ and $\widehat{m}_{\tau_l/N}^{x^0,N} = x^l + \Delta^l$.] From time τ_l we start a new graphical construction, which defines the evolution of both processes until the next discrepancy appears at time τ_{l+1}. For each $y \in \mathcal{D}_N$ and coordinate direction (i,a) we define $4(|\mathcal{S}_i| - 1)$ Poissonian mark processes $\widehat{N}_t^{i,a\to b,m}(y)$, $\widehat{N}_t^{i,a\to b,\widehat{m}}(y)$, $\widehat{E}_t^{i,a\to b,m}(y)$ and $\widehat{E}_t^{i,a\to b,\widehat{m}}(y)$ with respective rates

$$\begin{aligned}
\widehat{u}_t^{\,i,a\to b,m}(y,\Delta^l) &= \min\Big\{ y_{(i,a)}\lambda_i^{a\to b}(y),\, \big(y_{(i,a)} + \Delta^l_{(i,a)}\big)\lambda_i^{a\to b}(x^0_{t/N}) \Big\},\\
\widehat{u}_t^{\,i,a\to b,\widehat{m}}(y,\Delta^l) &= \min\Big\{ y_{(i,a)}\lambda_i^{a\to b}(x^0_{t/N}),\, \big(y_{(i,a)} - \Delta^l_{(i,a)}\big)\lambda_i^{a\to b}(y-\Delta^l) \Big\},\\
\widehat{e}_t^{\,i,a\to b,m}(y,\Delta^l) &= y_{(i,a)}\lambda_i(y) - \widehat{u}_t^{\,i,a\to b,m}(y,\Delta^l) \quad \text{and}\\
\widehat{e}_t^{\,i,a\to b,\widehat{m}}(y,\Delta^l) &= y_{(i,a)}\lambda_i(x^0_{t/N}) - \widehat{u}_t^{\,i,a\to b,\widehat{m}}(y,\Delta^l).
\end{aligned} \tag{7.13}$$

We observe that $\widehat{u}_t^{\,i,a\to b,m}(y,\Delta^l) = \widehat{u}_t^{\,i,a\to b,\widehat{m}}(y+\Delta^l,\Delta^l)$ for any $y \in \mathcal{D}_N$, so we identify

$$\widehat{N}_t^{i,a\to b,m}(y) = \widehat{N}_t^{i,a\to b,\widehat{m}}(y+\Delta^l). \tag{7.14}$$

Except for this identification, the processes are mutually independent for different (i,n) and independent of all previous Poisson mark processes.

The process $\{g_t^{x^0,N}\}$ jumps only at the marks placed by the processes $\{\widehat{E}_t^{i,a\to b,m}(y)\}$ and $\widehat{N}_t^{i,a\to b,m}(y)$, while process $\{\widehat{g}_t^{x^0,N}\}$ does so at the marks of $\{\widehat{E}_t^{i,a\to b,\widehat{m}}(y)\}$ and $\widehat{N}_t^{i,a\to b,\widehat{m}}(y)$. Due to identifications (7.14), the basic Poisson marks $\{\widehat{N}_t^{i,a\to b,m}(y)\}$ seen by $\{g_t^{x^0,N}\}$ at a given position y coincide with the basic marks seen by $\{\widehat{g}_t^{x^0,N}\}$ at its corresponding position $y+\Delta^l$. Hence, the two graphic processes evolve rigidly, keeping a separation Δ_l, until a discrepancy is met, that is, until one of the processes responds to a Poisson mark that the other ignores. This happens either because $\{g_t^{x^0,N}\}$ at a certain position y meets a mark of $\{\widehat{E}_t^{i,a\to b,m}(y)\}$ or because $\{\widehat{g}_t^{x^0,N}\}$, at the corresponding position $y+\Delta^l$, meets a mark of $\{\widehat{E}_t^{i,a\to b,\widehat{m}}(y+\Delta^l)\}$. This discrepancy defines the stopping time τ_{l+1} and is the first event of the *l-th-discrepancy process* $\{D_t^l\}_{t\in[\tau_l,\infty)}$, given by

$$D_t^l = \sum_{i\in\mathcal{T}} \sum_{\substack{a,b\in\mathcal{S}_i\\ a\neq b}} \Big[\widehat{E}_t^{i,a\to b,m}(y_t^l) + \widehat{E}_t^{i,a\to b,\widehat{m}}(y_t^l+\Delta^l) \Big] \tag{7.15}$$

where y_t^l is the density profile path defined by a realization of the (level l) construction done at this stage.

The construction done at the l-th stage makes sense, and has the correct rates, for $t \geq \tau_l$. Thus, together with the assumed graphical construction for $t \in [0,\tau_l)$, it yields a

well defined coupling for $g_t^{x^0,N}$ and $\widehat{g}_t^{x^0,N}$ at all times. Such a (level-l) coupling, however, loses precision after the next discrepancy is encountered, so we ignore it for $t \geq \tau_{l+1}$, and replace it by the level-$(l+1)$ construction corresponding to the $l+1$ stage. This stage begins with $g_{\tau_{l+1}}^{x^0,N} - \widehat{g}_{\tau_{l+1}}^{x^0,N} = \Delta_{l+1}$ with $|\Delta_{l+1} - \Delta_l| = 2/N$.

This recursive construction is continued, for each trajectory, until the time $t = NT$ is achieved. The procedure involves, almost surely, a finite number of stages, since rates are bounded above. The process

$$\overline{D}_t = D_t^l \quad \text{if } t \in (\tau_l, \tau_{l+1}] \tag{7.16}$$

$l = 0, 1, \ldots$ ($\tau_0 = 0$), counts the number of discrepancies. It satisfies the relation $\{\overline{D}_t \geq l\} = \{\tau_l \leq t\}$.

7.4 Discrepancy Rates

The proof of Theorem 4.1 requires the control of the distance between $m_t^{x^0,N}$ and $\widehat{m}_t^{x^0,N}$. As each discrepancy brings an additional separation of at most $2/N$,

$$\left|m_t^{x^0,N} - \widehat{m}_t^{x^0,N}\right| \leq \frac{2\overline{D}_{Nt}}{N}. \tag{7.17}$$

To estimate the right-hand side we first determine upper bounds on the time-dependent rate of the process $\{\overline{D}_t\}$.

Lemma 7.2 *Consider $N \in \mathbb{N}$, $T \geq 0$ and $\delta > 0$. For each $l \in \mathbb{N}$, let R_t^l be the instantaneous rate of the level-l discrepancy process D_t^l, $t \in [\tau_l, \tau_{l+1}]$ defined above and let $R^l = \sup\{R_t^l : t \in [\tau_l, \tau_{l+1}] \cap [0, NT]\}$. Then there exists a constant $A > 0$ such that the events*

$$\mathcal{R}_\delta^{NT} = \left\{ R_l \leq N^{\delta - 1/2} + \frac{Al}{N} \quad \forall l \text{ s.t. } \tau_l \leq NT \right\} \tag{7.18}$$

satisfy

$$P\left(\underline{\lim}_N \mathcal{R}_\delta^{NT}\right) = 1. \tag{7.19}$$

Proof Let Δ_t be the distance between the coupled geometrical realizations $g_t^{x^0,N}$ and $\widehat{g}_t^{x^0,N}$:

$$\Delta_t = \sum_{l \geq 0} \Delta^l \mathbf{1}_{t \in [\tau_l, \tau_{l+1})}. \tag{7.20}$$

The discrepancy process can be written as

$$\overline{D}_t = \sum_{i \in \mathcal{T}} \sum_{\substack{a,b \in \mathcal{S}_i \\ a \neq b}} \left[\widehat{E}_t^{i,a \to b,m}(g_t^{x^0,N}) + \widehat{E}_t^{i,a \to b,\widehat{m}}(g_t^{x^0,N} + \Delta_t) \right]. \tag{7.21}$$

The rate of this process is zero at $t = 0$, but it increases as the processes $\{g^{x^0,N_t}\}$, $\{\widehat{g}_t^{x^0,N}\}$ and $\{x_t^{x^0}\}$ move away from each other during the stochastic evolution.

We see from (7.13) that to bound this rate we must compare values of $x\lambda_i^{a\to b}(y)$ for different densities x and y. Due to the Lipschitz hypothesis (4.13), these differences increase at most linearly, and there exists a constant A such that the rate of $\overline{D}_t$ is bounded above by

$$A\big|\widehat{g}_t^{x^0,N} - x_{t/N}^{x^0}\big| + A\big|g_t^{x^0,N} - \widehat{g}_t^{x^0,N}\big|. \tag{7.22}$$

For a given realization of the graphical construction, the second term in (7.22) is bounded above by $2\overline{D}_t/N$, as remarked in (7.17). Therefore

$$\big|g_t^{x^0,N} - \widehat{g}_t^{x^0,N}\big| \le \frac{2l}{N} \quad \text{if } t \in [\tau_l, \tau_{l+1}]. \tag{7.23}$$

For the first term in (7.22) we apply first the probabilistic bound

$$\begin{aligned} &P\left(\big|\widehat{g}_t^{x^0,N} - x_{t/N}^{x^0}\big| > \frac{1}{2A}N^{\delta-1/2}\right) \\ &\quad = P\left(\big|\widehat{m}_{t/N}^{x^0,N} - x_{t/N}^{x^0}\big| > \frac{1}{2A}N^{\delta-1/2}\right) \\ &\quad < \exp(-cN^{\delta}) \end{aligned} \tag{7.24}$$

valid for *each* $t > 0$. The last inequality follows from (7.3). We need, however, a bound valid for *all* $t \in [0, NT]$. To this end, we apply (7.24) to a sufficiently thick collection of times. We pick a positive real γ (soon to be chosen larger than 3) and denote M the integer part of N^{γ}. For each $0 \le j \le M$ let

$$C_j = \left\{\big|\widehat{g}_{jNT/M}^{x^0,N} - x_{jT/M}^{x^0}\big| \le N^{\delta-1/2}\right\} \tag{7.25}$$

and

$$\Theta = \inf\left\{t : \big|\widehat{g}_t^{b(x^0),N} - x_{t/N}^{x^0}\big| > N^{\delta-1/2}\right\}. \tag{7.26}$$

Then,

$$\begin{aligned} P(\Theta \le NT) &\le P\left(\Theta \le NT, \bigcap_{l=0}^{M} C_l\right) + \sum_{l=0}^{M}\big(1 - P(C_l)\big) \\ &\le M\left[1 - \left(1 - \frac{dNT}{M}\right)\exp(-dNT/M)\right] + M\exp(-cN^{\delta}) \\ &\le cN^{2-\gamma} \end{aligned} \tag{7.27}$$

where d and c are positive constants. In the second line we used (7.24) to bound the last term; as for the other term, we just observe that the conditions $\Theta \le NT$ and $\bigcap_{l=0}^{M} C_l$ together imply that the process must have at least two transitions during the time interval of length NT/M containing Θ. The constant d bounds the rate of flips of the process $\{\widehat{g}_t^{b(x^0),N}\}_{t\ge 0}$ (we can take $d = \sum_{i\in\mathcal{T}}\sum_{a,b\in\mathcal{S}_i}\|\lambda_i^{a\to b}\|_\infty$). The choice $\gamma > 3$ yields a summable bound in (7.27), which implies

$$P\left(\overline{\lim}_N\big\{\Theta \le NT\big\}\right) = 0 \tag{7.28}$$

This result together with the bound (7.23) proves (7.19). □

7.5 Conclusion of the Proof

Due to Lemma 7.1 and relation (7.17), the following lemma concludes the proof of Theorem 4.1.

Lemma 7.3 *For any* $\varepsilon > 0$ *and* $0 \le t \le T$,

$$P\Big(\overline{\lim}_N \big\{\overline{D}_{NT} \ge N^{\varepsilon+1/2}\big\}\Big) = 0. \tag{7.29}$$

Proof Let us denote $\widetilde{N}_t = \overline{D}_{Nt}$. This process—which has rates N times higher than those of $\{\overline{D}_t\}_{t\ge 0}$—counts discrepancies in the time scale of $\{m_t^{x^0,N}\}_{t\ge 0}$. Let $\mathbf{T}_N$ be the time needed for the latter to collect $N^{\varepsilon+1/2}$ discrepancies:

$$\mathbf{T}_N = \min\big\{t : \widetilde{N}_t \ge N^{\varepsilon+1/2}\big\}. \tag{7.30}$$

It can be written in the form

$$T_N = \sum_{i=1}^{N_+^{\varepsilon+1/2}} T_i \tag{7.31}$$

where $T_1, T_2, \ldots$ are the independent successive times spent in between jumps and $N_+^{\varepsilon+1/2}$ is the smallest integer following $N^{\varepsilon+1/2}$.

We now choose some δ with $0 < \delta < \varepsilon$. By Lemma 7.2, the events

$$\mathcal{D}_r = \big\{\text{condition (7.18) is valid for } N \ge r\big\} \tag{7.32}$$

satisfy

$$P\Big(\bigcup_{r\in\mathbb{N}} \mathcal{D}_r\Big) = 1. \tag{7.33}$$

In the sequel we shall show that

$$\sum_N P\big(\mathbf{T}_N < T; \mathcal{D}_r\big) < \infty \tag{7.34}$$

for each natural number r. This concludes the proof because it implies that

$$P\Big(\overline{\lim}_N \big\{\mathbf{T}_N < T\big\}\Big) \le \sum_r P\Big(\overline{\lim}_N \big\{\mathbf{T}_N < T\big\}; \mathcal{D}_r\Big) = 0. \tag{7.35}$$

To prove (7.34) we partially re-sum the decomposition (7.31) in blocks of size

$$Q = \frac{N_+^{\varepsilon+1/2}}{N_+^{\delta+1/2}} \sim N^{\varepsilon-\delta} \xrightarrow[N\to\infty]{} \infty. \tag{7.36}$$

We consider intervals $I_l = [(l-1)N_+^{\delta+1/2} + 1, lN_+^{\delta+1/2}]$ and write

$$\mathbf{T}_N = \sum_{l=1}^{Q} G_l, \quad G_l = \sum_{j\in I_l} T_j. \tag{7.37}$$

Within $\mathcal{D}_r$ the process $\{\tilde{N}_t\}_{t\geq 0}$ jumps from i to $i+1$ with rates bounded above by $N^{\delta+1/2}+Ai$. Thus, for each $i\in I_l$ the rate of T_i is bounded above by $N^{\delta+1/2}+Al$, which is smaller than $(l+1)N^{\delta+1/2}$ if N is large enough. This shows that, for such N's, the output of each variable G_l is smaller than that of a sum of $N_+^{\delta+1/2}$ i.i.d. exponential random variables with rate $(l+1)N_+^{\delta+1/2}$. Hence,

$$P\big(\mathbf{T}_N<T;\mathcal{D}_r\big)\leq P\left(\sum_{l=1}^{Q}\frac{G_l(N_+^{\delta+1/2})}{(l+1)N_+^{\delta+1/2}}<T\right) \tag{7.38}$$

where $\{G(N^{\delta+1/2})\}_{l\geq 1}$ denotes an i.i.d. sequence of Gamma random variables with parameters $n=N_+^{\delta+1/2}$ and $\lambda=1$. The large-deviation properties of these distributions imply that each variable $G_{l,N}=G_l(N^{\delta+1/2})/N^{\delta+1/2}$ satisfies

$$P(G_{l,N}<1/2)\leq \exp(-cN^{\delta+1/2}) \tag{7.39}$$

for some positive constant c and all $1\leq l\leq Q$ and N large enough.

Denoting $A_{l,N}=\{G_{l,N}\geq 1/2)\}\cap\mathcal{D}_r$ and $B_{Q,N}=\bigcap_{l=1}^{Q}A_{l,N}\cap\mathcal{D}_r$ we have

$$P\big(\mathbf{T}_N<T;\mathcal{D}_r\big)\leq\big(1-P(B_{Q,N})\big)+P\big(\mathbf{T}_N<T,B_{Q,N}\big). \tag{7.40}$$

On the event $B_{Q,N}$, $\mathbf{T}_N$ is bounded below by

$$\frac{1}{2}\sum_{l=1}^{Q}\frac{1}{l+1}\sim\log Q\xrightarrow[N\to\infty]{}\infty. \tag{7.41}$$

Therefore the second term in the right-hand side of (7.40) is zero for N large enough. Bounding the first term by the large-deviation estimate (7.39) we conclude that

$$P\big(\mathbf{T}_N<T;\mathcal{D}_r\big)\leq Q\exp(-cN^{\delta+1/2}) \tag{7.42}$$

for N large enough. This proves (7.34). □

Acknowledgements It is a pleasure to thank Dr. Luiz Fernando Lima Reis for discussions on the biological motivation of this work. The authors wish to thank the NUMEC (R.F.), the University of São Paulo (R.F.) and the University of Rouen (E.J.N.) for hospitality during the completion of this work. They gratefully acknowledge the following sources of support: USP-COFECUB agreement (R.F. and E.J.N.), Antonio Prudente Cancer Research (FAPESP-CEPID) (E.J.N.), CNRS-FAPESP agreement (R.F.), CNPq grant 307978/2004-4 (L.R.F.), FAPESP grants 2004/07276-2 (E.J.N. and L.R.F.) and 06/03227-2 (E.J.N), and CNPq grant 484351/2006-0 (all three authors).

References

1. Chong, L., Ray, L.B.: Whole-istic biology. Science **295**(5560), 1661 (2002). Special Issue on Systems Biology
2. Harvey, L.: Signal transduction image originally. In: Molecular Cell Biology, 5th edn. Freeman, New York (2003). 973 s. b ill. ISBN: 0-7167-4366-3. Libris: 8926100, repressilator image based on Elowitz and Leibler (2000). http://en.wikipedia.org/wiki/
3. Alon, U.: An Introduction to Systems Biology: Design Principles of Biological Circuits. CRC Press, Boca Raton (2006). ISBN:1584886420
4. Kitano, H.: Systems biology: a brief overview. Science **295**, 1662–1664 (2002)

5. Sontag, E.D.: Some new directions in control theory inspired by systems biology. Syst. Biol. **1**, 9–18 (2004)
6. Wellstead, P.: Schroedinger's legacy systems and life. E.T.S. Walton Lecture, Royal Irish Academy, April (2005)
7. Alberts, B., et al.: Molecular Biology of the Cell, 4th edn. Garland, New York (2002). ISBN: 0-8153-3218-1
8. Oda, K., Matsuoka, Y., Funahashi, A., Kitano, H.: A comprehensive pathway map of epidermal growth factor receptor signaling. Molecular Systems Biology (2005). EMBO and Nature Publishing group; msb4100014
9. Tyson, J.J., Chen, K.C., Novak, B.: Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. Curr. Opin. Cell Biol. **15**(2), 221–231 (2003)
10. Wolf, D.M., Arkin, A.P.: Motifs, modules and games in bacteria. Curr. Opin. Microbiol. **6**(2), 125–134 (2003)
11. Alon, U.: Biological networks: the tinkerer as an engineer. Science **301**, 1866 (2003)
12. Hahn, W.C., Weinberg, R.A.: Modeling the molecular circuitry of cancer. Nat. Rev. Cancer **2**(5), 331–341 (2002)
13. http://stke.sciencemag.org/cm/
14. The chipping forecast III. Nat. Genet. vol. 37, June 2005
15. Carvalho, A.F., Reis, L.F., Brentani, R.R., Carraro, D.M., Verjovski-Almeida, S., Reis, E.M., Neves, E.J., de Souza, S.J., Brentani, H.: Gene expression arrays in cancer research: methods and applications. Crit. Rev. Oncol./Hematol. **54**, 95–105 (2005)
16. Gomes, L.I., Esteves, G.H., Carvalho, A.F., Cristo, E.B., Hirata, J.R.R., Martins, W.K., Brentani, H., Pelosof, A., Zitron, C., Sallum, R.A., Montagnini, A.L., Soares, F.A., Neves, E.J., Reis, L.F.L.: Expression profile of malignant and nonmalignant lesions of esophagus and stomach: differential activity of functional modules related to inflammation and lipid metabolism. Cancer Res. **65**(16), 7127–7136 (2005)
17. Elowitz, M.B., Leibler, S.: A synthetic oscillatory network of transcriptional regulators. Nature **403**(6767), 335–338 (2000)
18. Guckenheimer, J., Holmes, P.: Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields. Springer, Berlin (1990)
19. Tyson, J.J., Csikasz-Nagy, A., Novak, B.: The dynamics of cell cycle regulation. Bioessays **24**(12), 1095–1109 (2002)
20. Lauffenburger, D.A.: Cell signaling pathways as control modules: complexity for simplicity? Proc. Natl. Acad. Sci. **97**(10), 5031–5033 (2000)
21. Murray, J.D.: Mathematical Biology I. Springer, New York (2005)
22. McAdams, H.H., Arkin, A.: Stochastic mechanisms in gene expression. Proc. Natl. Acad. Sci. USA **94**, 814–819 (1997)
23. Elowitz, M.B., Leibler, S.: A synthetic oscillatory network of transcriptional regulators. Nature **403**(6767), 335–338 (2000)
24. Lahav, G., Rosenfeld, N., Sigal, A., Geva-Zatorsky, N., Levine, A.J., Elowitz, M.B., Alon, U.: Dynamics of the p53-Mdm2 feedback loop in individual cells. Nat Genet. **36**(2), 147–150 (2004)
25. Zhang, T., Brazhnik, P., Tyson, J.J.: Exploring mechanisms of the DNA-damage response: p53 pulses and their possible relevance to apoptosis. Cell Cycle **6**(1), 85–94 (2007)
26. http://www.math.pitt.edu/~bard/xpp/xpp.html
27. Liggett, T.M.: Interacting Particle Systems. Springer, Berlin (1985)
28. Thompson, C.J.: Classical Equilibrium Statistical Mechanics. Clarendon, Oxford (1988)
29. Durrett, R.: Stochastic spacial models. SIAM Rev. **41**(4), 677–718 (1999)
30. Ethier, S.N., Kurtz, T.G.: Markov Processes, Characterization and Convergence. Wiley, New York (1986)
31. Kurtz, T.G.: Approximation of Population Processes. Regional Conference Series in Applied Mathematics. SIAM, Philadelphia (1981)
32. Wormald, N.C.: Differential equations for random processes and random graphs. Ann. Appl. Probab. **5**, 1217–1235 (1995)
33. Schonmann, R.H.: An approach to characterize metastability and critical droplets in stochastic Ising models. Ann. Inst. Henri Poincaré, A Phys. Théor. **55**(2), 591–600 (1991)
34. Ruelle, D.: Statistical Mechanics. Benjamin, Elmsford (1969)
35. Sontag, E.D.: Monotone and near-monotone biochemical networks. Syst. Synth. Biol. **1**, 59–87 (2007)
36. Ellis, R.S.: Entropy, Large Deviations, and Statistical Mechanics. Springer, New York (2005). ISBN-10:3540290591
37. Thorisson, H.: Coupling, Stationarity, and Regeneration (Probability and its Applications), 1st edn. Springer, New York (2001). 536 p.