

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

Roberto Fernández^{1*}, Luiz R. Fontes^{2†}, E. Jordão Neves^{2 ‡}

January 13, 2008

- [1] Laboratoire de Mathématiques Raphaël Salem, UMR 6085 CNRS-Université de Rouen, Avenue de l'Université, B.P.12, F76801 St Etienne du Rouvray, France;
[2] University of São Paulo; Rua do Matão, 1010 - Cidade Universitária
CEP 05508-090 - São Paulo - SP - Brasil

Abstract

We introduce jump processes in \mathbb{R}^k , called *density-profile process*, to model biological signaling networks. They describe the macroscopic evolution of finite-size spin-flip models with k types of spins interacting through a non-reversible Glauber dynamics. We focus on the the k -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. 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 biological 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 .

1 Motivation and introduction

The interest in the analysis of dynamical processes going on within a biological system 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 [2, 3], which exploits ideas from dynamical systems and control theory [28], can be traced back at least sixty five years ago to Erwin Schrodinger's question: *What is life?* [5].

*roberto.fernandez@univ-rouen.fr

†lrenato@ime.usp.br

‡neves@ime.usp.br

Cells are fundamental biological entities and a particularly important problem in biology is to understand how they 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 an striking similarity to electrical circuitry. Extracellular information is often transmitted through cell-membrane receptors activated by chemical entities known as *ligands*, such as hormones, neurotransmitters or growth factors, which may trigger complex time-dependent cascades of internal cellular biochemical interactions and lead to several different cellular responses, like embryogenesis, motility, differentiation and *apoptosis* (carefully controlled cell suicide) [6]. These sequences of chemical reactions, called *signaling pathways* when the main interest is on the associated flow of information, may be capable of performing complex signal processing tasks and be themselves integrated to control even more complex multifaceted cellular behaviours. Signaling pathways usually involve the interaction of a large multi-scale hierarchy of subsystems, from those occurring at molecular level within small intra-cellular compartments to the coordinated dynamics of cells in organs of an organism across its life span. To illustrate the scale and complexity that may be involved in signaling pathways we may check a recent comprehensive overview in [7] of the important epidermal growth factor (EGF) signaling pathway, associated to the control of growth, survival, proliferation and differentiation in mammalian cells, which, based on the molecular interactions documented in 242 papers accessible from PubMed (<http://www.ncbi.nlm.nih.gov/pubmed>), identified 211 reactions involving 322 types of biochemical components.

The systems biology approach suggests multi-level analysis where the whole system is considered in terms of an hierarchy of interconnected functional *modules* [8, 9], each capable to perform some basic useful functionality which contribute to the whole-system behaviour. Signaling pathways may exhibit mechanisms similar to those found in engineered systems [10] as they had to survive strong selective pressures and therefore must be able to deploy sophisticated control strategies to prevent inappropriate physiological responses. Malfunctions in this biochemical circuitry may lead to several pathological conditions, like cancer [11].

As already noted, several different types of components may interact in a given signaling pathway and even when much information is available for a given pathway [12] the actual biochemical mechanisms associated to the associated interactions are frequently unknown. Two main biochemical components in signaling pathways are proteins and genes (RNA). They are closely related through the *central dogma of molecular biology* [6]: genomic DNA, which contains the basic genetic information of an individual, is read (or *transcribed*) into genes (RNA strings) which is then read (or *translated*) into proteins. The process where the available quantity of a given protein is regulated through the control of how much of the associated gene is being produced at each given time is known as *gene regulation* [6].

An important mathematical contribution to this research problem is to propose simple useful models for those pathways, which grasp the most relevant dynamical features of the biological system of interest, allow rigorous analysis and provide experimentally verifiable predictions. Due to the complexity of biological systems, *toy models* in the old tradition of physics literature, may be particularly useful to understand how the modules perform their functionalities

and how their interaction leads to the global behaviour of the system. Model predictions can be verified experimentally exploiting several recent advances in biochemical technologies, like cDNA microarrays [13].

Systems of chemical reactions can be naturally be modelled as a nonlinear continuous time dynamical system. Suppose there are k types of biochemical components in the signaling pathway and, for simplicity sake, assume also that the only relevant time-dependent variable associated to each biochemical component in the pathway is its density. If $x_t = (x_1, \dots, x_k)_t \in \mathbb{R}^k$ represent the global density profile at time t , the typical dynamical system describing a signaling pathway may be written as

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

where f and g are bounded smooth functions from \mathbb{R}^k to \mathbb{R}_+^k , indicating the velocity with which each one of the k components are, respectively, *produced* and *degraded* when the global density-profile is given by x . Once the set of non-linear differential equations are obtained, the general framework based on dynamical systems ideas [14], often relying on numerical methods, can be applied to analyse the corresponding model [15, 16].

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 [17]. 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. It is difficult to evaluate whether the conclusions of an analysis along this line depend on these somewhat arbitrary choices of equations. Moreover, even when reasonable, they may not be very useful as they usually depend on several parameters that can not actually be measured *in vivo*, that is, inside a living cell.

In order to find simpler toy models for this system we seek an alternative approach to derive the associated dynamical systems, which only tries to incorporate the essential *qualitative* information about the biochemical interactions in a given signaling network. This approach borrow ideas from Interacting Particle Systems [20] and represent components of the signaling pathway as different types of spin and the evolution of the whole system as a non-reversible Glauber spin-flip stochastic dynamics as explained in detail in section 3 below. From this stochastic spin systems we derive the *density-profile process*, a jump process in \mathbb{R}^k , that lead to the deterministic dynamical system 1.1 in the limit where the number of spins is very large. The density-profile process itself may be useful to analyse biochemical pathways where stochasticity play an important role [18].

To illustrate our approach in a biological setting we consider a *very* simple set of signaling pathways that we call *cyclic-interaction motif*. As part of a larger pathway it could provide the functionality of generating oscillations, essential for signaling pathways that organise time-modulated biological functions like, for instance, the periodic adjustment of an organism’s physiology to the circadian rhythm.

The precise general definition is given in section 3.2. The simplest cyclic-interaction motif capable to generate oscillations has only three components that we call A, B and C. They interact cyclically through a loop of *feedback inhibition* where each component represses the next one along the loop: component A

represses component B, component B represses component C and component C represses component A. More precisely, the rate of change of the density of component A at a given time depends only on the present density of C in such a way that component C inhibits component A: the density of A tends to decrease if the concentration of C is high. In a biological setting it could be that both A and C are proteins but protein C is an enzyme that accelerates the rate of degradation of protein A. Now, the rate of change of the density of component C depends only on the density of B in such a way that component B inhibits component C and component B depends only on component A in the same manner.

If the inhibition strength is large one could expect the onset of a dynamical instability on this system where concentrations of all three components oscillate.

A slightly more complex version of the cyclic-interaction motif, where the directed inhibition loop between components A, B and C arises from gene regulation, is known in the literature as the *repressilator* and was implemented (biochemically engineered) in the bacteria *Escherichia coli* [19] as a *synthetic biology system*. In this cyclic pathway, besides the proteins A, B and C, we have to consider their genes, indicated, respectively, by a, b and c. The cyclic interaction now is as follows: each gene, since it is translated into the corresponding protein, *promotes* that protein: 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. Namely, protein A inhibits the production of gene b, protein B inhibits the production of gene c and protein C inhibits gene a. Again, if interaction strengths are large enough, one expects oscillations to occur. Indeed the synthetic biology system in *E. coli* was engineered to be able to sustain oscillations, verified experimentally. For the toy models derived from our approach this dynamical instability is manifest through Hopf bifurcation with respect to the interaction strength parameter (Theorem 3.1).

Let us now discuss informally our modelling approach for the simplest cyclic-interaction motif with only three components. The general case, which includes the repressilator model, is considered below.

Suppose the dynamics of the basic clock motif evolves inside some cellular container where N sites, labelled from 1 up to N , are available for each one of the three types of components A, B and C. If we indicate types A, B and C by the numbers 1, 2 and 3, respectively, we write $\eta_t(i, n) = +1$ if there is a component of type i , $1 \leq i \leq 3$, in site n , $1 \leq n \leq N$, at time t and by $\eta_t(i, n) = -1$ the opposite. Write also $\eta_t \in \{-1, +1\}^{\{1,2,3\} \times \{1, \dots, N\}}$ for the global *configuration* of the system at time t . For each $i \in \{1, 2, 3\}$ denote by $c(i)$ the type of the component that inhibits it, that is, $c(1) = 3$, $c(2) = 1$ and $c(3) = 2$. Finally, we assume that each *spin* $\eta_t(i, n)$ flips with rates that depend only on the density at that time of the component $c(i)$. Denote by $c(i, n, \eta)$ the rate with which the spin of type i at site n flips when the present configuration is $\eta(i, n)$. A simple choice of rates to represent the qualitative inhibition loop in the cyclic-interaction motif is given by

$$c(i, n, \eta) = \exp\{J\eta(i, n) \frac{1}{N} \sum_{l=1}^N \eta(c(i), l)\} \quad (1.2)$$

where J is the inhibition strength. If $J > 0$ then the rates favors the situation

where spins of type i have the opposite sign of the majority of spins of type $c(i)$, an is therefore the situation of interest to mimic the inhibition loop. In the Interacting Particle Systems setting this stochastic model is a spin-flip model with non-reversible (due to the asymmetric nature of the rate dependence) mean-field Glauber dynamics. There are other natural choices for the rates 1.2 with the same qualitative appeal.

In this paper we analyse the thermodynamic limit (that is, the limit as N goes to infinity) of these models, as presented in Theorem 2.1 in a more general setting. In this limit, both the simplest cyclic-interaction motif model and the repressilator model yield deterministic dynamical systems of type 1.1 with Hopf bifurcation (Theorem 3.1) 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.

In the general setting considered in section 3 the dynamics of the k -dimensional empirical density vector of these spin models are described by the so-called density-profile process introduced in Section 2.1 below. These 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 2.1) is the proof that the paths of such a process converge almost surely to the trajectories of the dynamical system having V as the velocity field within arbitrary time intervals with some control of the distance between the stochastic and deterministic trajectories. As we show, the resulting dynamical systems can exhibit a very rich behavior, including bifurcations.

Similar problems were considered in the literature [22, 24]. It should be pointed out that ideas from the so called martingale problem and random time changes developed by T. Kurtz [23] provide an alternative method to prove almost sure convergence for our models (see also [25]). We did not investigate, though, the use of these techniques to get the a.s. convergence *together* with the control in the distance between the stochastic and deterministic trajectories as given in Theorem 2.1. The approach presented here, strongly exploiting coupling ideas [30] is arguably quite natural within the Interacting Particle Systems setting.

The main mathematical steps in the proof of our convergence result are the following:

- (i) A graphical construction (Section 4.2) that allows a coupled simultaneous construction of density-profile processes for different N .
- (ii) An auxiliary process $\{\widehat{m}_t^{x^0, N}\}_{t \geq 0}$ (Section 4.1) defined through a simple spin-flip model (independent flips with time-dependent rates) which shadows the deterministic dynamical system (Lemma 4.1).
- (iii) A coupling between the auxiliary and the density-profile processes that keep both processes as close to each other as possible (Section 4.4). Instants where they move further apart define a process of *discrepancies*. Bounds on the rate of these discrepancies yield our convergence theorem (Theorem 2.1).

2 Convergence of density profile process to dynamical systems

In this section we define our basic processes and present the main mathematical result of the paper. In next section we shall realize them through the macroscopic dynamics of stochastic Ising models.

2.1 Density-profile processes $\{m_t^{x^0 N}\}_{t \geq 0}$

A density-profile process is a continuous time jump process in the hypercube $\mathcal{D}_N = (-\frac{1}{N}, 1 + \frac{1}{N})^k$, for $k, N \geq 1$. At each jump, a point $x \in \mathcal{D}_N$ changes one of its coordinates by $\pm 1/N$, with rates that depend smoothly on x . These rates are defined in the following way. We start with two bounded Lipschitz functions $\lambda, \mu : \mathbb{R}^k \rightarrow \mathbb{R}_+^k$ which, in turns, define functions $f, g : \mathbb{R}^k \rightarrow \mathbb{R}_+^k$ through the relations

$$f_i(x) = \begin{cases} (1 - x_i)\lambda_i(x) & \text{if } 0 \leq x_i \leq 1 \\ \lambda_i(0^+) & \text{if } x_i \leq 0 \\ 0 & \text{if } x_i \geq 1 \end{cases} \quad (2.1)$$

and

$$g_i(x) = \begin{cases} x_i\mu_i(x) & \text{if } 0 \leq x_i \leq 1 \\ 0 & \text{if } x_i \leq 0 \\ \mu_i(1^-) & \text{if } x_i \geq 1 \end{cases} \quad (2.2)$$

for $i = 1, \dots, k$.

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 size $1/N$ along the coordinate directions. From each position x , the rates for jumps forward or backwards along the coordinate direction i are, respectively, $Nf_i(x)$ and $Ng_i(x)$. That is, for $1 \leq i \leq k$ and $x \in \mathcal{D}_N$,

$$Nf_i(x) = \left. \frac{d}{dt} P\left(m_t^{x^0 N} = x + \frac{e_i}{N}\right) \right|_{t=0} \quad (2.3)$$

and

$$Ng_i(x) = \left. \frac{d}{dt} P\left(m_t^{x^0 N} = x - \frac{e_i}{N}\right) \right|_{t=0}, \quad (2.4)$$

where e_i denotes the unit vector along direction i .

In our applications, the variable x_1, \dots, x_k represent the densities of k types of objects that can be present at N different sites. The function λ_i is the rate of creation or activation of an object of type i at a site where the object is absent or inhibited. The function μ_i is the rate for the opposite move. Thus, Nf_i (resp. g_i) is the total rate of creation (resp. destruction) of an object of type i throughout the entire collection of sites.

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 Section 4 offer, in fact, a concrete way of defining this common space.)

2.2 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 a given pair of appropriate functions (λ, μ) , and let $V : \mathbb{R}^k \rightarrow \mathbb{R}_+^k$ be its associated *velocity field*:

$$V(x) = \lim_{t \downarrow 0} \frac{\mathbf{E}(m_t^{x^0 N} - x)}{t} = f(x) - g(x), \quad (2.5)$$

that is,

$$V_i(x) = \begin{cases} (1 - x_i)\lambda_i(x) - x_i\mu_i(x) & \text{if } 0 \leq x_i \leq 1 \\ \lambda_i(0^+) & \text{if } x_i \leq 0 \\ \mu_i(1^-) & \text{if } x_i \geq 1 \end{cases} \quad (2.6)$$

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

$$\dot{x}_t = V(x_t) \quad (2.7)$$

starting at $x^0 \in (0, 1)^k$. The global trajectory exists by the smoothness of the field V . Furthermore, the flow does not leave $(0, 1)^k$ because $V_i(0^+) > 0$ and $V_i(1^-) < 0$ for $i = 1, \dots, k$.

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

$$\tau_\epsilon^N = \inf \left\{ t \geq 0 : |m_t^{x^0, N} - x_t^{x^0}| > \frac{1}{N^{\frac{1}{2} - \epsilon}} \right\} \quad (2.8)$$

and, for a given T , $0 \leq T < \infty$, write $\mathcal{A}_{N\epsilon}^T = \{\tau_\epsilon^N < T\}$. The following is our main result.

Theorem 2.1 *Let λ and μ be bounded functions from \mathbb{R}^k to \mathbb{R}_+^k satisfying the Lipschitz condition*

$$|\lambda(x) - \lambda(y)| \leq K|x - y| \quad ; \quad |\mu(x) - \mu(y)| \leq K|x - y| \quad (2.9)$$

for some $K > 0$ and all $x, y \in (0, 1)^k$. Then, for any finite T , initial position x^0 and $\epsilon > 0$,

$$P\left(\overline{\lim}_N \mathcal{A}_{N\epsilon}^T\right) = 0 \quad (2.10)$$

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 (2.7)/(2.6) can exhibit quite complex dynamics—even for simple choices of λ and μ —, including stable orbits and chaotic behavior. In Section 3.2 we present an example of a system leading to a Hopf bifurcation.

3 Associated stochastic spin models

We present here a family of stochastic spin models, motivated by biological cellular systems, whose empirical densities evolve as density-profile processes.

3.1 General definition

We consider a discrete set Λ of *sites* and a finite set $\mathcal{T} = \{1, \dots, k\}$ of *types*. Each type $i \in \mathcal{T}$ can be present at each *site* $\ell \in \Lambda$. Thus, we choose our configuration space as $\Sigma = \{-1, +1\}^{\mathcal{T} \times \Lambda}$, where for each $\eta \in \Sigma$, the value $\eta(i, \ell) = +1$ ($\eta(i, \ell) = -1$) indicates the presence (absence) of a particle of type i at site ℓ .

We consider continuous-time processes in Σ in which only single spin flips are allowed. For $\eta \in \Sigma$ and $(i, \ell) \in \mathcal{T} \times \Lambda$, let $\eta^{(i, \ell)}$ denote the configuration which is equal to η except at site ℓ where the spin of type i is *flipped*.

$$\eta^{(i, \ell)}(j, n) = \begin{cases} \eta(j, n) & \text{if } (j, n) \neq (i, \ell) \\ -\eta(i, \ell) & \text{otherwise} \end{cases} \quad (3.1)$$

We shall consider *type-dependent interaction models*, for which the flip rate for the spin-flip transition $\eta \rightarrow \eta^{(i, \ell)}$ is a monotone non-increasing function of

$$\Delta(i, \ell, \eta) = H_i(\eta^{(i, \ell)}) - H_i(\eta) \quad (3.2)$$

for a *Hamiltonian vector* $H(\eta) = (H_1(\eta), \dots, H_k(\eta))$.

A simple choice is

$$c(i, \ell, \eta) = \exp[-\Delta(i, \ell, \eta)]. \quad (3.3)$$

We focus on the particular case

$$H_i(\eta) = -\frac{1}{2} \sum_{\ell \in \Lambda} \left(\sum_{(j, n) \in \mathcal{T} \times \Lambda} \frac{\alpha_{ji}}{|\Lambda|} \eta(j, n) \eta(i, \ell) + a_i \eta(i, \ell) \right). \quad (3.4)$$

This corresponds to a *mean-field* interaction, where α_{ji} is the strength of the influence of spins of type j on those of type i and each a_i acts as a type-dependent external field. The most interesting phenomena appear when α is not symmetric. The *empirical density profile* of a configuration η , is the k -uple $m(\eta) = (m_1(\eta), \dots, m_k(\eta)) \in \mathbb{R}_+^k$, where

$$m_i(\eta) = \frac{|\{\ell \in \Lambda : \eta(i, \ell) = +1\}|}{|\Lambda|}. \quad (3.5)$$

for $1 \leq i \leq k$.

Let $\{\sigma_t^{\eta_0}\}_{t \geq 0}$ be the spin-flip process starting at the configuration $\eta_0 \in \Sigma$ with rates defined by (3.3) and (3.4). Then, if $|\Lambda| = N$, the density-profile process $m(\sigma_t^{\eta_0})$ approximates, in the sense of Theorem 2.1 the dynamical system (2.7)/(2.6), defined by

$$\lambda_i(x) = \exp\left(\sum_{j \in \mathcal{T}} \alpha_{ji} x_j + a_i\right) \quad (3.6)$$

and

$$\mu_i(x) = \exp\left(-\sum_{j \in \mathcal{T}} \alpha_{ji} x_j - a_i\right). \quad (3.7)$$

Given our biological motivation, models with mean-field interaction like (3.4) are natural. They are also the simplest models with type-dependent rates of the form (3.3). From the mathematical point of view, though, the analysis of models with local interaction [26], possibly with *stirring* [27] are more interesting.

3.2 Example: Cyclic-interaction model

We exhibit now a model —called the *cyclic-interaction model*— defined through a simple choice of the interaction matrix α which nevertheless leads to interesting (deterministic) dynamical behavior. Two particular cases of these models, the basic clock motif and the repressilator, were informally presented in the introduction. Think $\{1, \dots, k\}$ as points on the circle and, for each $i \in \mathcal{T}$ let $c(i)$ denote the nearest-neighbor of i in the counter-clockwise direction. We assume that $\alpha_{ji} = 0$ unless $j = c(i)$ and that, furthermore, all non-zero terms in α have the same absolute value. That is,

$$\alpha_{ji} = \begin{cases} s_i J & \text{if } j = c(i) \\ 0 & \text{otherwise} \end{cases} \quad (3.8)$$

where $s_i \in \{-1, +1\}$ representing the signals, and $J > 0$. We also set $a_i = -J/2$, for $1 \leq i \leq k$. In this way, once the signs $\{s_i\}_{i=1}^k$ 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 $+1$ [defined in (3.6)], is an increasing 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* [21]. 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 [defined in (3.7)], decreases as a function of $x_{c(i)}$, and the interaction of spins of type $c(i)$ with those of type i is *anti-ferromagnetic*. In biochemical terms, the component $c(i)$ *inhibits* the production of component i .

The dynamical system (2.7) associated to the cyclic-interaction model (3.8) is:

$$\dot{x}_i = e^{s_i J(x_{c(i)} - \frac{1}{2})} - x_i \left(e^{s_i J(x_{c(i)} - \frac{1}{2})} + e^{-s_i J(x_{c(i)} - \frac{1}{2})} \right) \quad (3.9)$$

for $1 \leq i \leq k$. 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 (3.9) 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 [28] the associated directed graph is not consistent. This system exhibits a Hopf bifurcation [14] 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 3.1 Consider the dynamical system (3.9) with $k \geq 3$

- (a) If $\prod_{i=1}^k s_i = 1$, there is a bifurcation at $J_c = 2$: the fixed point $(1/2, \dots, 1/2)$ loses 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 (3.9) is close to $\dot{x} = A(x - \mathbf{1}/2)$, where A is a $k \times k$ real matrix with eigenvalues $sJ e^{\frac{2\pi l}{k} i} - 2$, $l = 0, 1, \dots, k-1$. 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 $J_c = 2/\cos(\pi/k)$. \square

Remark 3.2 For instance, if $k = 3$ and all interactions are antiferromagnetic ($s_i = -1$ for $i = 1, 2, 3$), the dynamical system has stable orbits for $J > J_c = 4$. This is the basic clock motif discussed before. The convergence result, Theorem 2.1, implies that, within any finite time interval, the density-profile process evolves as close to this orbit as wished, for N sufficiently large.

4 Proof of the convergence theorem

4.1 The auxiliary process $\{\widehat{m}_t^{x^0, N}\}_{t \geq 0}$

To prove Theorem 2.1 we introduce an auxiliary stochastic spin model with independent spins flips but *time-dependent rates*.

Let $\Lambda = \{1, \dots, N\}$ and let $\{\eta_t(i, n) : (i, n) \in \mathcal{T} \times \Lambda\}_{t \geq 0}$, be kN independent Markov chains with state space $\{-1, +1\}$. Thus, for each $t \geq 0$, $\eta_t \in \Sigma = \{-1, +1\}^{\mathcal{T} \times \Lambda}$, with $\mathcal{T} = \{1, \dots, k\}$, as defined in the previous section. For each Markov chain $\{\eta_t(i, n)\}_{t \geq 0}$ the flips from -1 to $+1$ and from $+1$ to -1 have time-dependent rates given, respectively, by $\lambda_i(x_t^{x^0})$ and $\mu_i(x_t^{x^0})$, where $\{x_t^{x^0}\}_{t \geq 0}$ is the solution of the dynamical system (2.7) from the initial position x^0 . We initialize these chains with the uniform distribution on configurations η_0 with $m(\eta_0) = x^0$, where $m(\eta_t)$ is defined in (3.5). The total number of spins of each type i is, thus, fixed and equal to x_i^0 , but the initial density components $m_1(\eta_0), \dots, m_k(\eta_0)$ are independent. 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 $p_t(i, n) = P(\eta_t(i, n) = +1)$, we have

$$\dot{p}_t(i, n) = [1 - p_t(i, n)] \lambda_i(x_t^{x^0}) - p_t(i, n) \mu_i(x_t^{x^0}). \quad (4.1)$$

Therefore each function $t \rightarrow p_t(i, n)$ is a solution of the differential equation (2.7) with V as in (2.6). Hence

$$p_0(i, n) = (x^0)_i \implies p_t(i, n) = (x_t^{x^0})_i \quad \forall t \geq 0, \quad (4.2)$$

for all $i \in \mathcal{T}$ and $n \in \Lambda$ [$(y)_i$ indicates the i -th component of vector $y \in \mathbb{R}^k$]. While (4.2) is true for the auxiliary process $\{\widehat{m}_t^{x^0, N}\}$, we are interested in following the actual empirical densities. Next lemma proves that also the path followed by these densities remain, at all times, close to the trajectories of the dynamical system.

Lemma 4.1 *For $\delta > 0$ there exists $c > 0$ such that*

$$P\left(|\widehat{m}_t^{x^0, N} - x_t^{x^0}| > N^{\delta-1/2}\right) < \exp(-cN^\delta) \quad (4.3)$$

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 with $P(\eta_0(i, n) = +1) = (x^0)_i$. Hence, the density components $m_1(\eta), \dots, m_k(\eta)$ are independent and each $m_i(\eta)$ has a binomial distribution with parameters N and $p_i = (x^0)_i$, $i = 1, \dots, k$. (This means that, for large N , $\widehat{m}_t^{b(x^0),N}$ starts at a random position in \mathcal{D} close to x^0 , while $\widehat{m}_t^{x^0,N}$ starts precisely at x^0 .)

This new auxiliary process also satisfies (4.2) but has the advantage that the spin chains remain independent at all times, and, by (4.2), the proportions of spins of each type coincide with the components of $x_t^{x^0}$. Therefore $N\widehat{m}_t^{b(x^0),N}$ is a vector of independent binomial random variables

$$N\left(\widehat{m}_t^{b(x^0),N}\right)_i \sim \text{Bin}(N, (x_t^{x^0})_i) \quad (4.4)$$

for $t \geq 0$ and $i \in \mathcal{T}$. In particular the variances of the components of $\widehat{m}_t^{b(x^0),N}$ are proportional to $1/N$. Thus, the large-deviation properties of binomial distributions [29] imply that for any $\delta > 0$ there exists a constant c such that

$$P\left(\left|\widehat{m}_t^{b(x^0),N} - x_t^{x^0}\right| > \frac{1}{2}N^{\delta-1/2}\right) < \exp(-cN^\delta) \quad (4.5)$$

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 [30, 20]. 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 (4.1). At sites (i, n) with $\eta_0^{b(x^0),N}(i, n) = \eta_0^{x^0,N}(i, n)$, the spins evolve together. Otherwise, the spins for both processes evolve independently until one of them makes a transition, thereby bringing them to a common value. They evolve together ever after. As the distance between the corresponding coupled density profiles decreases with time,

$$\begin{aligned} \left|m(\eta_t^{b(x^0),N}) - m(\eta_t^{x^0,N})\right| &\leq \left|m(\eta_0^{b(x^0),N}) - m(\eta_0^{x^0,N})\right| \\ &= \left|m(\eta_0^{b(x^0),N}) - x^0\right| \end{aligned} \quad (4.6)$$

and, therefore,

$$\begin{aligned} P\left(\left|\widehat{m}_t^{x^0,N} - x_t^{x^0}\right| > N^{\delta-1/2}\right) &\leq \\ P\left(\left|\widehat{m}_t^{x^0,N} - \widehat{m}_t^{b(x^0),N}\right| > \frac{1}{2}N^{\delta-1/2}\right) &+ P\left(\left|\widehat{m}_t^{b(x^0),N} - x_t^{x^0}\right| > \frac{1}{2}N^{\delta-1/2}\right). \end{aligned} \quad (4.7)$$

To prove (4.5) we bound the right-hand side using (4.6) and (4.5) (for $t = 0$) for the first term and again (4.5) for the second one. \square

To prove Theorem 2.1 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.

4.2 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 $y \in \mathcal{D}_N$ we associate $2k$ independent Poisson processes: $N_t^{1+}(y)$, $N_t^{1-}(y)$, \dots , $N_t^{k+}(y)$, $N_t^{k-}(y)$, where each $N_t^{i+}(y)$ has rate $f_i(y)$ and each $N_t^{i-}(y)$ rate $g_i(y)$. 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 $N_t^{i+}(y)$ ($N_t^{i-}(y)$) carries the instruction to jump along the positive (negative) i coordinate direction.

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 a coordinate direction 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 (2.3) and (2.4)] are N times those of the former. Thus, one process can be constructed from the other by a simple change in the time scale:

$$m_t^{x^0, N} = g_{Nt}^{x^0, N}. \quad (4.8)$$

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

4.3 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 minimize 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, $g_0^0 = \widehat{g}_0^0 = x^0$ 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 $i = 1, \dots, k$ we define six Poissonian mark processes:

- (i) *Marks associated to jumps from y to $y + e_i/N$* : They are defined by independent Poisson processes $\widehat{N}_t^{i,+}(y)$, $\widehat{E}_t^{i,+m}(y)$ and $\widehat{E}_t^{i,+,\widehat{m}}(y)$ with respective rates

$$\begin{aligned}\widehat{u}_t^{i,+}(y) &= \min\{(1-y_i)\lambda_i(y), (1-y_i)\lambda(x_t^{x_0^0}/N)\}, \\ \widehat{e}_t^{i,+m}(y) &= |(1-y_i)\lambda_i(y) - \widehat{u}_t^{i,+}|_+ \quad \text{and} \\ \widehat{e}_t^{i,+,\widehat{m}}(y) &= |(1-y_i)\lambda_i(x_t^{x_0^0}/N) - \widehat{u}_t^{i,+}|_+.\end{aligned}\tag{4.9}$$

$[|z|_+] = \max\{z, 0\}$. Note the rescaling in time for the deterministic path $\{x_t^{x_0^0}\}$ needed to represent it on the graphical construction time scale.

- (ii) *Marks associated to jumps from y to $y - e_i/N$* : Defined by three independent Poissonian processes $\widehat{N}_t^{i,-}(y)$, $\widehat{E}_t^{i,-m}(y)$ and $\widehat{E}_t^{i,-,\widehat{m}}(y)$ which are independent from the precedent ones and have respective rates

$$\begin{aligned}\widehat{u}_t^{i,-}(y) &= \min\{y_i\mu_i(y), y_i\mu(x_t^{x_0^0}/N)\}, \\ \widehat{e}_t^{i,-m}(y) &= |y_i\mu_i(y) - \widehat{u}_t^{i,-}|_+ \quad \text{and} \\ \widehat{e}_t^{i,-,\widehat{m}}(y) &= |y_i\mu_i(x_t^{x_0^0}/N) - \widehat{u}_t^{i,-}|_+.\end{aligned}\tag{4.10}$$

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 $\{g_t^{x_0^0,N}\}$ occur at the marks of $\{\widehat{E}_t^{i,+m}(y)\}$ and $\{\widehat{E}_t^{i,-m}(y)\}$; while those of the process $\{\widehat{g}_t^{x_0^0,N}\}$ are at $\{\widehat{E}_t^{i,+,\widehat{m}}(y)\}$ and $\{\widehat{E}_t^{i,-,\widehat{m}}(y)\}$. The marks of the four processes $\{\widehat{E}_t^{i,+m}(y)\}$, $\{\widehat{E}_t^{i,+,\widehat{m}}(y)\}$, $\{\widehat{E}_t^{i,-m}(y)\}$ and $\{\widehat{E}_t^{i,-,\widehat{m}}(y)\}$ are thus seen by only one of $\{g_t^{x_0^0,N}\}$ or $\{\widehat{g}_t^{x_0^0,N}\}$ and will be called *discrepancies*. The *basic Poisson processes* $\{\widehat{N}_t^{i,\pm}(y)\}$, on the other hand, are introduced to ensure that $\{g_t^{x_0^0,N}\}$ and $\{\widehat{g}_t^{x_0^0,N}\}$ 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*

$$D_t^0 = \sum_{i=1}^k \left[\widehat{E}_t^{i,+m}(y^0) + \widehat{E}_t^{i,-m}(y^0) + \widehat{E}_t^{i,+,\widehat{m}}(y^0) + \widehat{E}_t^{i,-,\widehat{m}}(y^0) \right]\tag{4.11}$$

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 2.1. 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

$$g_{\tau_l}^{x_0^0,N} = x^l \quad , \quad \widehat{g}_{\tau_l}^{x_0^0,N} = x^l + \Delta^l.\tag{4.12}$$

[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} . We define the following Poisson mark processes for each $y \in \mathcal{D}_N$ and coordinate direction $i = 1, \dots, k$:

- (i) *Marks associated to jumps from y to $y + e_i/N$* : Let $\widehat{N}_t^{i,+m}(y)$, $\widehat{N}_t^{i,+,\widehat{m}}(y)$, $\widehat{E}_t^{i,+m}(y)$ and $\widehat{E}_t^{i,+,\widehat{m}}(y)$ be Poisson processes with respective rates

$$\begin{aligned}\widehat{u}_t^{i,+m}(y, \Delta^l) &= \min\{(1 - y_i) \lambda_i(y), (1 - y_i - \Delta_i^l) \lambda(x_{t/N}^{x^0})\}, \\ \widehat{u}_t^{i,+,\widehat{m}}(y, \Delta^l) &= \min\{(1 - y_i) \lambda_i(x_{t/N}^{x^0}), (1 - y_i + \Delta_i^l) \lambda_i(y - \Delta^l)\}, \\ \widehat{e}_t^{i,+m}(y, \Delta^l) &= |(1 - y_i) \lambda_i(y) - \widehat{u}_t^{i,+m}(y, \Delta^l)|_+ \quad \text{and} \\ \widehat{e}_t^{i,+,\widehat{m}}(y, \Delta^l) &= |(1 - y_i) \lambda_i(x_{t/N}^{x^0}) - \widehat{u}_t^{i,+,\widehat{m}}(y, \Delta^l)|_+.\end{aligned}\tag{4.13}$$

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

$$\widehat{N}_t^{i,+m}(y) = \widehat{N}_t^{i,+,\widehat{m}}(y + \Delta^l).\tag{4.14}$$

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

- (ii) *Marks associated to jumps from y to $y - e_i/N$* : They are determined by Poisson processes $\widehat{N}_t^{i,-m}(y)$, $\widehat{N}_t^{i,-,\widehat{m}}(y)$, $\widehat{E}_t^{i,-m}(y)$ and $\widehat{E}_t^{i,-,\widehat{m}}(y)$, respectively with rates

$$\begin{aligned}\widehat{u}_t^{i,-m}(y, \Delta^l) &= \min\{y_i \mu_i(y), (y_i + \Delta_i^l) \mu(x_{t/N}^{x^0})\}, \\ \widehat{u}_t^{i,-,\widehat{m}}(y, \Delta^l) &= \min\{y_i \mu_i(x_{t/N}^{x^0}), (y_i - \Delta_i^l) \mu(y - \Delta^l)\}, \\ \widehat{e}_t^{i,-m}(y, \Delta^l) &= |y_i \mu_i(y) - \widehat{u}_t^{i,-m}(y, \Delta^l)|_+ \quad \text{and} \\ \widehat{e}_t^{i,-,\widehat{m}}(y, \Delta^l) &= |y_i \mu_i(x_{t/N}^{x^0}) - \widehat{u}_t^{i,-,\widehat{m}}(y, \Delta^l)|_+;\end{aligned}\tag{4.15}$$

with the identification

$$\widehat{N}_t^{i,-m}(y) = \widehat{N}_t^{i,-,\widehat{m}}(y + \Delta^l).\tag{4.16}$$

All these processes are independent among themselves, except for the preceding identification, and independent of other mark processes.

The process $\{g_t^{x^0,N}\}$ jumps only at the marks placed by the processes $\{\widehat{E}_t^{i,\pm,m}(y)\}$, while process $\{\widehat{g}_t^{x^0,N}\}$ does so at the marks of $\{\widehat{E}_t^{i,\pm,\widehat{m}}(y)\}$. Due to identifications (4.14)/(4.16), the basic Poisson marks $\{\widehat{N}_t^{i,\pm,m}(y)\}$ 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 $\{\widehat{E}_t^{i,+m}(y) + \widehat{E}_t^{i,-m}(y)\}$ or because $\{\widehat{g}_t^{x^0,N}\}$, at the corresponding position $y + \Delta^l$, meets a mark of

$\{\widehat{E}_t^{i,+,\widehat{m}}(y + \Delta^l) + \widehat{E}_t^{i,-,\widehat{m}}(y + \Delta^l)\}$. Therefore, this discrepancy, corresponding to the stopping time τ_{l+1} , is the first event of the l -th-discrepancy process $\{D_t^l\}_{t \in [\tau_l, \infty)}$, given by

$$D_t^l = \sum_{i=1}^k \left[\widehat{E}_t^{i,+,\widehat{m}}(y_t^l) + \widehat{E}_t^{i,-,\widehat{m}}(y_t^l) + \widehat{E}_t^{i,+,\widehat{m}}(y_t^l + \Delta^l) + \widehat{E}_t^{i,-,\widehat{m}}(y_t^l + \Delta^l) \right] \quad (4.17)$$

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. To improve it, 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 $|\Delta_{l+1} - \Delta_l| = 1/N$.

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. The process

$$\overline{D}_t = D_t^l \quad \text{if } t \in (\tau_l, \tau_{l+1}] \quad (4.18)$$

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

4.4 Discrepancy rates

The proof of Theorem 2.1 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 $1/N$,

$$|m_t^{x^0, N} - \widehat{m}_t^{x^0, N}| \leq \frac{\overline{D}_t}{N} \quad (4.19)$$

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

Lemma 4.2 *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*

$$\mathcal{R}_\delta^{NT} = \left\{ R_l \leq N^{\delta-1/2} + \frac{Al}{N} \quad \forall l \text{ s.t. } \tau_l \leq NT \right\} \quad (4.20)$$

satisfy

$$P(\underline{\lim}_N \mathcal{R}_\delta^{NT}) = 1. \quad (4.21)$$

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

$$\Delta_t = \sum_{l \geq 0} \Delta^l \mathbf{1}_{t \in [\tau_l, \tau_{l+1})}. \quad (4.22)$$

The discrepancy process can be written as

$$\begin{aligned} \bar{D}_t = \sum_{i=1}^k & \left[\widehat{E}_t^{i,+m}(g_t^{x^0,N}) + \widehat{E}_t^{i,-m}(g_t^{x^0,N}) \right. \\ & \left. + \widehat{E}_t^{i,+\widehat{m}}(g_t^{x^0,N} + \Delta_t) + \widehat{E}_t^{i,-\widehat{m}}(g_t^{x^0,N} + \Delta_t) \right]. \end{aligned} \quad (4.23)$$

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 (4.13) and (4.15) that to bound this rate we must compare values of $x \lambda(y)$ and $x \mu(y)$ for different densities x and y . Due to the Lipschitz hypothesis (2.9), these difference increase at most linearly, and there exists a constant A such that the rate of \bar{D}_t is bounded above by

$$A |\widehat{g}_t^{x^0,N} - x_{t/N}^{x^0}| + A |g_t^{x^0,N} - \widehat{g}_t^{x^0,N}|. \quad (4.24)$$

For a given realization of the graphical construction, the second term in (4.24) is bounded above by $\frac{\bar{D}_t}{N}$, as remarked in (4.19). Therefore

$$|g_t^{x^0,N} - \widehat{g}_t^{x^0,N}| \leq \frac{l}{N} \quad \text{if } t \in [\tau_l, \tau_{l+1}]. \quad (4.25)$$

For the first term in (4.24) we apply first the probabilistic bound

$$\begin{aligned} & P\left(|\widehat{g}_t^{x^0,N} - x_{t/N}^{x^0}| > \frac{1}{2A} N^{\delta-1/2}\right) \\ &= P\left(|\widehat{m}_{t/N}^{x^0,N} - x_{t/N}^{x^0}| > \frac{1}{2A} N^{\delta-1/2}\right) \\ &< \exp(-cN^\delta). \end{aligned} \quad (4.26)$$

valid for *each* $t > 0$. The last inequality follows from (4.3). We need, however, a bound valid for *all* $t \in [0, NT]$. To this end, we apply (4.26) 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 $0 \leq j \leq M$ let

$$C_j = \left\{ |\widehat{g}_{jNT/M}^{x^0,N} - x_{jT/(NM)}^{x^0}| \leq N^{\delta-1/2} \right\} \quad (4.27)$$

and

$$\Theta = \inf \left\{ t : |\widehat{g}_t^{b(x^0),N} - x_{t/N}^{x^0}| > N^{\delta-1/2} \right\}. \quad (4.28)$$

Then,

$$\begin{aligned} P(\Theta \leq NT) &\leq P(\Theta \leq NT, \cap_{l=0}^M C_l) + \sum_{l=0}^M (1 - P(C_l)) \\ &\leq M \left[1 - \left(1 - \frac{dNT}{M} \right) \exp(-dNT/M) \right] + M \exp(-cN^\delta) \\ &\leq cN^{2-\gamma} \end{aligned} \quad (4.29)$$

where d and c are positive constants. In the second line we used (4.26) to bound the last term. For the other term, we just observed that the conditions $\Theta \leq NT$

and $\cap_{l=0}^M C_l$ 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 $\{\tilde{g}_t^{b(x^0), N}\}_{t \geq 0}$ (we can take $d = \sum_{i=1}^k \|\lambda_i\|_\infty + \sum_{i=1}^k \|\mu_i\|_\infty$). The choice $\gamma > 3$ yields a summable bound in (4.29), which implies

$$P\left(\overline{\lim}_N \{\Theta \leq NT\}\right) = 0 \quad (4.30)$$

This result together with the bound (4.25) proves (4.21). \square

4.5 Conclusion of the proof

Due to Lemma 4.1 and relation (4.19), the following lemma concludes the proof of Theorem 2.1.

Lemma 4.3 *For any $\varepsilon > 0$ and $0 \leq t \leq T$,*

$$P\left(\overline{\lim}_N \{\overline{D}_{NT} \geq N^{\varepsilon+1/2}\}\right) = 0. \quad (4.31)$$

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

$$\mathbf{T}_N = \min\{t : \tilde{N}_t \geq N^{\varepsilon+1/2}\}. \quad (4.32)$$

It can be written in the form

$$T_N = \sum_{i=1}^{N_+^{\varepsilon+1/2}} T_i \quad (4.33)$$

where T_1, T_2, \dots 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 4.2, the events

$$\mathcal{D}_r = \{\text{condition (4.20) is valid for } N \geq r\} \quad (4.34)$$

satisfy

$$P\left(\bigcup_{r \in \mathbb{N}} \mathcal{D}_r\right) = 1. \quad (4.35)$$

In the sequel we shall show that

$$\sum_N P(\mathbf{T}_N < T; \mathcal{D}_r) < \infty \quad (4.36)$$

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

$$P\left(\overline{\lim}_N \{\mathbf{T}_N < T\}\right) \leq \sum_r P\left(\overline{\lim}_N \{\mathbf{T}_N < T\}; \mathcal{D}_r\right) = 0. \quad (4.37)$$

To prove (4.36) we partially re-sum the decomposition (4.33) in blocks of size

$$Q = \frac{N_+^{\varepsilon+1/2}}{N_+^{\delta+1/2}} \sim N^{\varepsilon-\delta} \xrightarrow{N \rightarrow \infty} \infty \quad (4.38)$$

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

$$\mathbf{T} = \sum_{l=1}^Q G_l \quad , \quad G_l = \sum_{j \in I_l} T_j . \quad (4.39)$$

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 no 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,

$$P(\mathbf{T} < T; \mathcal{D}_r) \leq P\left(\sum_{l=1}^Q \frac{G_l(N_+^{\delta+1/2})}{(l+1)N_+^{\delta+1/2}} < T\right) \quad (4.40)$$

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

$$P(G_{l,N} < 1/2) \leq \exp(-cN^{\delta+1/2}) \quad (4.41)$$

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 $B_{Q,N} = \cap_{l=1}^Q A_{l,N} \cap \mathcal{D}_r$ we have

$$P(\mathbf{T}_N < T; \mathcal{D}_r) \leq (1 - P(B_{Q,N})) + P(\mathbf{T}_N < T, B_{Q,N}) \quad (4.42)$$

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

$$\frac{1}{2} \sum_{l=1}^Q \frac{1}{l+1} \sim \log Q \xrightarrow{N \rightarrow \infty} \infty . \quad (4.43)$$

Therefore the second term in the right-hand side of (4.42) is zero for N large enough. Bounding the first term by the large-deviation estimate (4.41) we conclude that

$$P(\mathbf{T}_N < T; \mathcal{D}_r) \leq Q \exp(-cN^{\delta+1/2}) \quad (4.44)$$

for N large enough. This proves (4.36). \square

Acknowledgments

The authors wish to thank the NUMEC and the University of São Paulo (R.F.) and the University of Rouen (E.J.N.) for hospitality during the completion of this work and to the USP-COFECUB agreement. E.J.N. also thanks Antonio Prudente Cancer Research (FAPESP-CEPID). L.R.F. is partially supported by the CNPq grant 307978/2004-4. E.J.N. and L.R.F. are partially supported by the by FAPESP grant 2004/07276-2. All three authors are partially supported by CNPq grant 484351/2006-0. R.F. benefited from the CNRS-FAPESP agreement.

References

- [1] L. Chong and L. B. Ray; **Whole-istic Biology** ; Science 1 March 2002: Vol. 295. no. 5560, p. 1661; Special Issue on Systems Biology.
- [2] U. Alon; **An Introduction to Systems Biology: Design Principles of Biological Circuits** ; 2006; CRC Press; ISBN 1584886420.
- [3] H. Kitano, Systems Biology: a brief overview, Science, 295:1662-1664, 2002.
- [4] E. D. Sontag; Systems Biology (2004) 9-18; **Some new directions in control theory inspired by systems biology.**
- [5] Peter Wellstead *Schroedinger's Legacy: Systems and Life*, E.T.S. Walton Lecture, Royal Irish Academy, April, 2005.
- [6] *Molecular biology of the cell*; Bruce Alberts, et al. 4th ed. ISBN 0-8153-3218-1, Published by Garland Science.
- [7] K. Oda, Y. Matsuoka, A. Funahashi and H. Kitano; **A comprehensive pathway map of epidermal growth factor receptor signaling**; Molecular Systems Biology, 2005, EMBO and Nature Publishing group; msb4100014.
- [8] J. J Tyson, K. C Chen and B. Novak; **Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell**; Current Opinion in Cell Biology, Volume 15, Issue 2, April 2003, Pages 221-231;
- [9] Wolf D.M.; Arkin A.P.; **Motifs, modules and games in bacteria**; Current Opinion in Microbiology, April 2003, vol. 6, no. 2, pp. 125-134(10).
- [10] U. Alon; Science 301, 1866 (2003); **Biological Networks: The Tinkerer as an Engineer.**
- [11] Hahn WC, Weinberg RA; **Modelling the molecular circuitry of cancer**; Nat. Rev. Cancer; 2002 May;2(5):331-41.
- [12] <http://stke.sciencemag.org/cm/>
- [13] **The chipping forecast III**; Nat. Genetics; June 2005, Volume 37.
- [14] J. Guckenheimer and P. Holmes. **Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields**. Springer-Verlag, 1990.
- [15] Tyson JJ, Csikasz-Nagy A, Novak B.; **The dynamics of cell cycle regulation**. Bioessays. 2002 Dec;24(12):1095-109.
- [16] Lauffenburger DA.; *Cell signaling pathways as control modules: complexity for simplicity?*; Proc. Natl. Acad. Sci. ; 2000;97(10):5031-3.
- [17] *Mathematical Biology I*; J.D. Murray, Publisher: Springer;2005.
- [18] H. H. McAdams and A. Arkin; **Stochastic mechanisms in gene expression**; Proc. Natl. Acad. Sci. USA; Vol. 94, pp. 814-819, February 1997.

- [19] M.B. Elowitz and S. Leibler; Nature. 2000 Jan 20;403(6767):335-8; **A Synthetic Oscillatory Network of Transcriptional Regulators.**
- [20] T. M. Liggett (1985): **Interacting particle systems.** Springer-Verlag.
- [21] Colin J. Thompson, **Classical Equilibrium Statistical Mechanics** (Clarendon Press, 1988).
- [22] R. Durrett; *Stochastic Spacial Models*; SIAM review; Vol. 41; No. 4; pp. 677-718, 1999.
- [23] S. N. Ethier and T.G. Kurtz, **Markov processes, characterization and convergence**, Wiley (1986).
- [24] T.G. Kurtz, **Approximation of population processes**, Regional Conference Series in Applied Mathematics, SIAM, 1981.
- [25] N. C. Wormald, **Differential equations for random processes and random graphs**; 1995, Ann. Appl. Probab.,5, 1217-1235.
- [26] D. Ruelle; *Statistical Mechanics*; W. A. Benjamin, Reading, MA.
- [27] A. De Masi, P.A. Ferrari, J.L. Lebowitz Reaction-diffusion equations for interacting particle systems J. Stat. Phys. 44, 3/4:589-644 (1986).
- [28] E.D. Sontag; **Monotone and near-monotone biochemical networks**; Systems and Synthetic Biology, 1, 59-87, 2007.
- [29] Richard S. Ellis; *Entropy, Large Deviations, and Statistical Mechanics*; Publisher: Springer; (2005); ISBN-10: 3540290591.
- [30] Hermann Thorisson; *Coupling, Stationarity, and Regeneration (Probability and its Applications)*; Hardcover: 536 pages; Publisher: Springer; 1 edition (2001).