PAU RUÉ

TRANSIENT AND STOCHASTIC DYNAMICS IN CELLULAR PROCESSES

DIRECTOR: JORDI GARCIA-OJALVO

MEMÒRIA PRESENTADA PER OPTAR AL GRAU DE DOCTOR

DEPARTAMENT DE FÍSICA I ENGINYERIA NUCLEAR
UNIVERSITAT POLITÈCNICA DE CATALUNYA

BARCELONA, MAIG 2013

UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

"If I have seen further, it is by standing upon the shoulders of giants."
— Sir Isaac Newton to Robert Hooke, 1676

The work presented here is the result of not only my own endeavour, but also the immensely valuable guidance, inspiration, advice and assistance of many.

To all of them I express my most sincere gratitude.

Pau Rué
Barcelona, spring 2013

Agraïments

Aquesta Tesi no hauria estat possible sense el suport d'un bon grapat de gent que, d'una manera o altra, m'han ajudat a arribar fins aqui i han estat a prop meu durant tot aquest temps. M'agradaria donar-los les gràcies a tots ells:

Al Jordi, pel seu suport, dedicació i confiança en aquests anys, i per la seva manera de fer desenfadada i optimista que fa que tot sembli més fàcil.

Especialment al Toni i a la Núria, els meus dos "germans grans" científics, per tots el que m'han ensenyat i els bons moments que m'han donat dins i fora de la ciència.

A la Marta, a la Belén, al Jordi Tiana, a la Lorena, a l'Elena, al Jordi Zamora, i al Dani per fer un grup molt acollidor.

També a la Carme, la Cristina Martinez, l'Andrés i tota la gent de Gaia. I als nous companys al PRBB, al Marcal, a la Nara, a l'Alessandro, al Maciek i a la Letícia.

Molt sincerament a l'Alfonso, per ensenyar-me que hi ha altres maneres de prendre's la ciència, per mostrar-me com de fascinant és la biologia i per acollir-me al seu grup.

A la Silvia, per tots els bons moments que hem passat desxifrant el que les cèl·lules ens volien dir.

A la Tina, per fer de Cambridge un lloc assolellat.

Al Joaquín, al Pedro, al David, al Jamie, a la Penny, a la Stefania, al Jonathan, al Christian i a la Chea, per fer-me sentir part d'un grup excepcional.

Al Gürol, al Munehiro, al Mark i al Tolga, per mostrar-me com de complexos poden ser els bacteris.

Al Marcus i al seu grup a l'IMPPC, per ensenyar-me que el laboratori pot ser un lloc agradable.

A la colla d'amics, per allò de *"all work and no play makes Pau a dull boy"*. Molt especialment al Beto, al Pau, a la Carme, al Ferran, a la Gemma, al Dídac i al Marc.

A l'Anna, per la seva inestimable ajuda i al Francesc, per saber escoltar.

Al pare i a la mare, així com al Gerard, al Joan i a la Xènia, per fer-me sempre costat en tot.

A la tieta i el JP per animar-me a continuar fent recerca.

A l'avi i l'avia, per transmetre'm que l'esforç sempre té recompensa.

Al Fausto, company de passeig infal·lible.

I a la Gemma, per recolzar-me en tot moment, per fer-me tocar de peus a terra i per compartir-ho tot.

Resum

En aquesta Tesi s'estudien diferents processos intracel·lulars i de poblacions cel·lulars regits per dinàmica estocàstica i no lineal. El problemes biològics tractats graviten al voltant el concepte de dinàmica transitòria i de relaxació d'un estat dinàmic pertorbat a l'estat estacionari. En aquest sentit, en tots els processos estudiats, les fluctuacions estocàstiques, presents intrínsecament o aplicades de forma externa, hi tenen un paper constructiu, ja sigui empenyent els sistemes fora de l'equilibri, interferint amb les lleis deterministes subjacents, o establint els nivells d'heterogeneïtat necessaris.

La primera part de la Tesi es dedica a l'estudi de processos cel·lulars transitoris regulats genèticament. En ella analitzem des d'un punt de vista teòric tres circuits genètics de control de polsos excitables i, contràriament al que s'havia especulat anteriorment, establim que tots ells poden treballar en dos tipus de règim excitable. Analitzem també com, en presència de soroll molecular, aquests circuits excitables poden generar polsos periòdics i multimodals degut a la combinació de dos fenòmens induïts per soroll: l'estabilització estocàstica d'estats inestables i la ressonància de coherència.

D'altra banda, estudiem com un mecanisme genètic excitable pot ser el responsable de regular a nivell transcripcional les fluctuacions que s'observen experimentalment en alguns factors de pluripotència en cèl·lules mare embrionàries. En l'embrió, la pluripotència és un estat cel·lular transitori i la sortida de les cèl·lules d'aquest sembla que està associada a fluctuacions transcripcionals. En relació al control de la pluripotència, presentem també un nou mecanisme basat en la regulació post-traduccional d'un petit conjunt de 4 factors de pluripotència. El model teòric proposat, basat en la formació de complexos entre els diferents factors de pluripotència, l'hem validat mitjançant experiments quantitatius en cèl·lules individuals. El model postula que l'estat de pluripotència no depèn dels nivells cel·lulars d'un únic factor, sinó d'un equilibri de correlacions entre diverses proteïnes. A més, prediu el fenotip de cèl·lules mutants i suggereix que la funció reguladora de les interaccions entre les quatre proteïnes és la d'esmorteir l'activitat transcripcional d'Oct4, un dels principals factors de pluripotència.

En el segon apartat de la Tesi estudiem el comportament d'una xarxa computacional de senyalització cel·lular de fibroblast humà en presència de senyals externs fluctuants i cíclics. Els resultats obtinguts mostren que la xarxa respon de forma no trivial a les fluctuacions ambientals, fins i tot en presència d'una senyal externa. Diferents nivells de soroll permeten modular la resposta de la xarxa, mitjançant la selecció de rutes alternatives de transmissió de la informació.

Finalment, estudiem la dinàmica de poblacions cel·lulars durant la formació de biofilms, pel·lícules arrugades d'aglomerats de bacteris que conformen un dels exemples més simples d'estructures multicel·lulars autoorganitzades. En aquesta Tesi presentem un model espai-temporal de creixement i mort cel·lular motivat per l'evidència experimental sobre l'aparició de patrons de mort massiva de bacteris previs a la formació de les ar-

rugues dels biofilms. Aquests patrons localitzats concentren les forces mecàniques durant l'expansió del biofilm i inicien la formació de les arrugues característiques. En aquest sentit, el model proposat explica com es formen els patrons de mort a partir dels canvis de mobilitat dels bacteris deguts a la producció de matriu extracel·lular combinats amb un creixement espacialment heterogeni. Una important predicció del model és que la producció de matriu és un procés clau per a l'aparició dels patrons i, per tant de les arrugues. En aquest aspecte, els nostres resultats experimentals en bacteris mutants que no produeixen components essencials de la matriu, confirmen les prediccions.

Summary

This Thesis presents some new results in the study of different cellular and cell population processes driven by non-linear and stochastic dynamics. Despite the approach taken here is mainly theoretical and computational, the problems addressed are of biological interest and, in most of the cases, the theoretical work is a result of close collaborations with experimentalists. Amongst the systems studied we find the regulation of pulsating genetic responses in transient cellular differentiation, the transcriptional and post-translational control of pluripotency in mouse Embryonic Stem cells, the integration of dynamic external information by cellular signalling networks, and the self-organised spatio-temporal population dynamics of bacterial communities that lead to formation of complex structures in biofilms.

Although the range of non-linear dynamic biological processes tackled in this dissertation is broad, all these systems gravitate around the concepts of transient dynamics and relaxation from a perturbed to a steady state. In this regard, stochastic fluctuations, either intrinsically present in or externally applied to these systems play an important and constructive role, by either driving the systems out of equilibrium, interfering with the underlying deterministic laws, or establishing suitable levels of heterogeneity.

In Part I we introduce and motivate, from both biological and physical points of view, all cellular and cell population processes studied in the following Parts. We devote Part II to transient cellular processes, from the mathematical analysis of mechanisms of genetic excitability for the regulation of transitory cellular responses (Chapter 2) and generation of genetic pulses with non-trivial statistics (Chapter 3), to the analysis and modelling of transcriptional and post-translational mechanisms of regulation of pluripotency in Embryonic Stem (ES) cells (Chapter 4). In Chapter 2 we show, from a dynamical systems perspective, how simple interactions between two genes can regulate transient pulsating responses via different bifurcations from limit cycle attractors in the presence of molecular noise. A special case of noise-induced pulses, namely periodic pulses with polymodal distribution of durations, is worked out in Chapter 3. The genetic circuit designs studied in Chapters 2 and 3 allow for stereotyped transient responses to external stimuli and the generation of highly heterogeneous distributions of cellular phenotypes via regulated fluctuations. This regulated heterogeneity, which is known to be controlled by simple circuits, at least, in some stress responses of the soil bacterium *Bacillus subtilis*, has also been observed in the expression of Nanog, a key regulator of pluripotency in mouse ES cells. Indeed, heterogeneity in Nanog expression is thought to mediate cell priming and commitment to differentiation. Using experimental data on transcriptional interactions, we propose and analyse excitable mechanisms of regulation of pluripotency by which Nanog levels drop in a stochastic and regulated manner (Chapter 4).

The second part of Chapter 4 is committed to the analysis of regulation of pluripotency beyond the transcriptional mechanisms. Using mathematical modelling and quantitative experimental evidence at the single cell

level, we uncovered an association between pluripotency and the correlated levels of Nanog protein with other pluripotency factors, namely Oct4, Tcf3 and β-catenin. On this basis, we propose a novel (and complementary) molecular mechanism of regulation based on competitive binary protein interactions (Sections 4.2 and 4.3). The post-translational protein interaction network proposed is capable of anticipating the phenotypes of mutant cell lines. Additionally, the model predicts that pluripotency is inversely related to the levels of transcriptionally active Oct4, and that the function of the protein network that is to buffer these levels in a dynamic manner. We further support this view with a bioinformatic reanalysis of existing data of Nanog and Oct4 function and genetic interaction.

Part III of this essay addresses some aspects of transient dynamics, signal processing, integration and propagation in cellular signalling networks under fluctuating environments, from a purely theoretical viewpoint. In particular, we develop some results on the relaxation dynamics and frequency response of a model of the fibroblast signalling network in the presence of background fluctuations (Sections 5.1 and 5.2), showing that the ensemble dynamics of the network is driven by a characteristic time scale. We show that if the network is stimulated by a cycling external signal with a time scale comparable to that of the network, environmental fluctuations produce a network-wide modulation of signal detection (Sections 5.3 to 5.6), a situation reminiscent of classic stochastic resonance. This emergent resonance property can be understood by the selection of internal information transmission pathways that external noise exerts on the network. We further explore how multiple signals integrate in the presence of background noise and find that the network is capable of displaying simple patterns of signal integration. Our computational analysis shows that these integration patterns depend on the levels of fluctuating background activity carried by other cell inputs. These results indicate that signal integration is sensitive to environmental fluctuations, and that these effectively determine the information integration capabilities of the cell.

The last Part of this Thesis (Part IV) is devoted to the transient spatio-temporal dynamics of bacterial populations during biofilm formation. In this Part, bacterial growth and death processes, extra-cellular matrix (ECM) production and waste generation are related to the generation of transient localised spatio-temporal patterns of cell death that precede formation of the characteristic biofilm wrinkles. We present a simple mesoscopic model that accounts for the observed experimental patterns that arise from the amplification of microscopic heterogeneities. Together with the experimental deletion of genes implicated in biofilm development, computational simulations of the model reveal that ECM production underlies the formation of localised patterns of massive bacterial death, The proposed model thus provides a conceptual explanation for how the extra-cellular matrix can promote a localised pattern.

Finally, we discuss and summarise the main results presented in this Thesis and its implications in Part V, and add some concluding remarks.

Contents

List of Figures

List of Tables

I

Introduction

1

Introduction

Biological processes arise from complex and dynamic interactions between simple elements ruled by the laws of physics and chemistry. These interactions often span multiple scales in space and time: from nanometric molecular machines interacting at the nanosecond scale, to whole macroscopic animal systems developing over several years. Such broad range of scales, which are more than a reflection of the many levels of interactions at increasing complexity, make it almost impossible for these processes to be stated and described with the postulates of the above-mentioned fields of chemistry and physics.

Yet, following the spirit behind the aphorism put forward by the celebrated biochemist Jacques Monod, "What is true of *E. coli* is true of the elephant", one can try to uncover some basic principles common to all, or most, biological processes. At the cellular level, for instance, mechanisms of regulation of gene expression are one such example. Since the pioneering work by Ptashne on the mechanisms of gene regulation, it has become more and more clear that there exist some basic principles underlying genetic processes. Today, these principles define a common (although yet to be polished) language in which these processes can be defined.

Biological mechanisms, however, do not exclusively rely on gene expression regulation. In many cellular processes, for instance, gene expression can be seen as the final output of a cellular decision that involves the processing of external signals and integration with the internal cellular state. Indeed, signalling interactions are one of the key mechanisms of living cell. Only with an efficient flow of information across its biochemical architecture, cells are capable of reacting to the environment and to fulfil the basic processes needed to preserve life. Biochemical protein-

protein interactions are basic mechanism of signal transduction and a modification of the cellular state. These interactions are also ruled by basic principles that link physical mechanisms to their ultimate function.

At a larger scale, we can think of cells as individual processing units that interact to perform macroscopic, tissue- or colony- level biological functions. This extra level of complexity comes with added benefits: as individual life units, cells are able to reproduce and decease. As a matter of fact, regulation of growth and death has largely proven to be a fundamental mechanism in the development of multicellular organisms and communities.

A fundamental property of mechanisms implementing biological functions, across all their scales, is their dynamic nature. Most biological and cellular processes involve factors that evolve in time. In this sense, physics and mathematics have brought up a large repertoire of tools that have proven useful in elucidating the rules and mechanisms behind these dynamic cellular and biological processes.

The work presented in dissertation mainly relies on these tools, namely dynamical systems theory, to address some very specific problems in gene regulation, intra-cellular signalling, and population-level cellular interactions that share some basic features. The following Sections succinctly describe each of the systems and processes studied.

1.1 *Dynamics of transient stochastic cellular processes*

Transient processes and excitability

Understanding the processes that regulate living cells is a challenging task. Most of these processes involve complex networks of interacting genes and proteins and imply time-varying changes in the concentrations of the molecular species. The most basic example of dynamic behaviour is provided by periodic oscillations. Rhythmic intracellular activity is the essence of the well known circadian and cell-cycle oscillations. Other regulation mechanisms require, instead, non-rhythmic dynamic processes in which intracellular proteins and molecules suffer transient changes in their levels. In certain situations, for instance, cells need to respond to external conditions by activating a new cellular state that is not beneficial if sustained. In other circumstances, populations of genetically identical cells benefit from highly heterogeneous and reproducible phenotype profiles, which sometimes arise from tightly regulated fluctuations at the single cell level. The adequate strategy in those circumstances is to produce transient pulses or drops of gene expression.

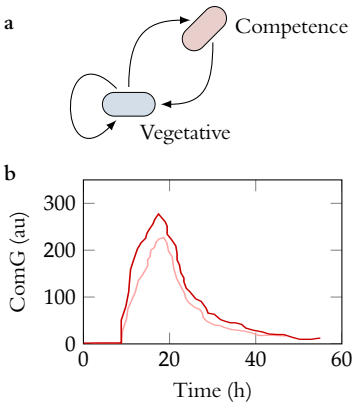

Figure 1.1: **Genetic competence in *Bacillus subtilis* soil bacterium.** Bacterial competence is a transient bacterial differentiation process temporarily adopted by cells when under nutritional stress. (a) A small fraction of normal (vegetative) *B. subtilis* cells randomly switch to the competent state and eventually recover its original state. (b) Competence involves the transient change in the levels and activities of proteins like ComK or ComG, as illustrated here by the time traces of ComG activity of two different cells. Experimental data adapted from Suel et al. (2007).

This is the case, for instance, in the response of *Bacillus subtilis* to nutritional stress. This soil bacterium has evolved an elaborate suite of stress responses that include developing the ability to take up exogenous DNA. This response, known as genetic competence, requires the activation of a large number of genes in order to prepare the cell to absorb and process the high-molecular-weight extracellular DNA molecules. Given the substantial physiological changes incurred by the cells in that process, competence is understood as a cellular state different from the vegetative state of normal cell growth and division. In the competence state cells can grow, although at a lower rate (Süel et al., 2006), and replicate their DNA, but they cannot divide. Therefore, a sustained state of competence is disadvantageous to the cells, since due to nutrient shortage long non-dividing cells run the risk of dying without progeny, while non-competent cells keep dividing at high rates. Time-lapse fluorescence microscopy confirms that competence is a transient cellular state that occurs in a small fraction of cells under stress (Süel et al., 2006; Suel et al., 2007, and see Figure 1.1).

Transient behaviours are also a common response mechanism in eukaryotic cells. Upon DNA damage, for instance, proliferating mouse and human cells react by generating transient pulses of expression of the

tumour suppressor protein p53. Indeed, p53 protein is believed to play a role in preventing cancer development by triggering cell-cycle arrest and apoptosis in response to DNA damage and therefore its prolonged expression might be detrimental in sane cells. Recent experimental results (Loewer et al., 2010) on the dynamics of p53 in basal conditions show that cells display, even in the absence of DNA damage stress, spontaneous pulses of p53 that are similar to those in stress conditions, hinting at a p53-activation mechanism analogous to the competence initiation programme in *B. subtilis*.

Other examples of transient behaviour in mammalian cells exist. As an example, Embryonic Stem (ES) cells in their pluripotent state have been seen to exhibit a highly heterogeneous profile in the expression of key pluripotency genes such as Nanog (Martinez Arias and Brickman, 2011). In particular, a small sub-population of cells in the culture has low levels of Nanog, which correlate with a higher probability of priming to differentiation (Kalmar et al., 2009). Flow cytometry studies using fluorescence-activated cell sorting showed that those cells expressing low Nanog were able to de-differentiate back into fully pluripotent cells expressing high Nanog levels (Singh et al., 2007). Such dynamic de-differentiation process, which has also been observed by means of time-lapse fluorescence microscopy (Kalmar et al., 2009), hints also at a transient mechanism of cellular priming to differentiation.

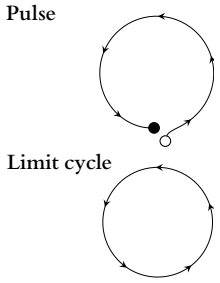

Figure 1.2: **Excitable dynamics.** Schematic representation of excitable dynamics (top) and its associated limit-cycle behaviour (bottom). The solid circle in the top plot represents the stable fixed point of the system. Representation adapted from Izhikevich (2006).

Thus, transient pulsed activity is an ubiquitous mechanism of cellular regulation, and some cellular processes that take the form of activity pulses can be interpreted in terms excitable dynamics driven by noise. Mathematically, excitability regards a dynamical system with a stable fixed point which is forced to undergo, when subject to a large enough perturbation, a large excursion in phase space before relaxing back to the fixed point (Lindner et al., 2004, and see top panel in Figure 1.2). This feature arises when the system is close to a bifurcation point beyond which the dynamics has the form of a limit cycle (bottom panel in Figure 1.2).

Functionally, excitability provides cells with a mechanism to amplify molecular noise and transform it into a macroscopic cellular response. Biochemical activity pulses have been reported to exist in cAMP signalling in amoebae (Hofer, Sherratt, and Maini, 1995), as well as in the response to DNA damage in human cells (Lahav et al., 2004), competence in bacteria (Süel et al., 2006; Suel et al., 2007), and differentiation

priming in embryonic stem cells (Kalmar et al., 2009) described above. In a different (but still biological) context, electrical activity pulses in a specific cell type, namely neurons, have long been associated with excitable dynamics (Izhikevich, 2006).

Within that context, early work by Hodgkin (1948) identified two main different types of excitability on the basis of the frequency response of the neuron across a bifurcation leading from excitable to oscillatory behaviour (Figure 1.3). Type I excitability corresponds to a situation in which the limit cycle is born at the bifurcation with a frequency equal to zero, i.e., the oscillatory activity pulses become infinitely sparse as the bifurcation is approached (first two panels in Figure 1.3). In type II excitability, on the other hand, the limit cycle is born with a non-zero frequency (and thus with a finite period), as shown in the bottom panel of Figure 1.3. The distinction is not purely academic: given that neurons encode information mainly in the timing between pulses, the two types of excitability correspond to two fundamentally different modes of information transmission (St-Hilaire and Longtin, 2004). Additionally, they exhibit distinct statistical features in their response to noise (ibid.).

The two types of excitability shown in Figure 1.3 are associated with different classes of bifurcation to oscillatory behaviour. Two kinds of bifurcation lead to type I excitability. The first is a saddle-node on an invariant circle (SNIC) bifurcation, in which a stable node and a saddle point appear together on a limit cycle. In that regime, the trajectory along the remnant of the limit cycle delimits the excitable pulse. The second bifurcation leading to type I excitability is a saddle-homoclinic bifurcation, which occurs when a stable limit cycle collides with a saddle point, being transformed into a homoclinic orbit at the bifurcation point. Again this orbit delimits the excitable pulse. In both cases the existence of a saddle provides the system with a well defined excitability threshold, given by the stable manifold of the saddle. The two bifurcation scenarios are characterised by qualitatively distinct scaling relations of the frequency with respect to the control parameter (Strogatz, 1994), as specified inside the plots at the left column of Figure 1.3.

Type II excitability, on the other hand, is associated with a Hopf bifurcation, either supercritical or subcritical, or to a saddle-node bifurcation occurring near (not on top of) a limit cycle. In the latter case, the saddle separates the stable fixed point from the periodic orbit, and again its stable manifold constitutes a separatrix that determines the excitability threshold. In the case of the Hopf bifurcation, on the other hand, it is not

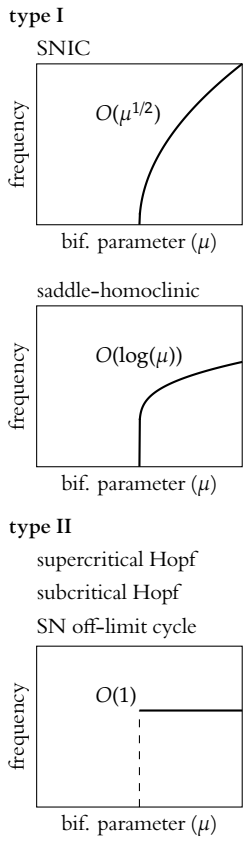

Figure 1.3: **Frequency response to the increase of a given control parameter,** for the two types of excitable behaviour discussed in the text, as a bifurcation from excitability to oscillations is crossed. The different bifurcation types underlying the different scenarios are listed in top of each plot. In each figure, the scaling law of the period of the cycle with the bifurcation parameter is shown.

possible to define in a quantitative way an excitation threshold in phase space.

In Chapter 2 we present an overview of different gene circuit architectures that exhibit excitable pulses of protein expression when subject to intrinsic molecular noise. Different types of excitable dynamics can occur depending on the bifurcation structure leading to the specific excitable phase-space topology. The bifurcation structure, as we show, is not linked to a particular circuit architecture. Thus a given gene circuit design can sustain different classes of excitable dynamics depending on the system parameters. In Chapter 3 and Section 4.1 we further develop simple models of excitable genetic circuits in particular contexts, such as complex rhythmicity of cell-cycles and transcriptional regulation of pluripotency in ES cells.

Rhythmicity, the temporal organisation of biological processes into self-repeating cascades of events, is the basis of most cellular functions. Biological oscillations, for instance, underlie many physiological functions in cells, from basic processes such as cell growth and division (Tyson, K. Chen, and Novak, 2001) to evolutionary environmental adaptations such as circadian rhythmicity (Goldbeter, 2002). Some circuit architectures have been proposed that explain the observed periodic behaviour in terms of limit-cycle attractors of non-linear dynamical models (Novák and Tyson, 2008). These limit cycles exhibit a perfectly periodic behaviour, which is only slightly perturbed by realistic levels of biochemical random fluctuations, or noise, that are unavoidable in cells (Gonze, Halloy, and Goldbeter, 2002). In some situations, however, cellular oscillations display a degree of variability much larger than what can be obtained from a limit-cycle model with added noise. This is the case, for instance, of the cell cycle oscillations exhibited by Chinese hamster cells (Klevecz, 1976), fission yeast cells (Sveiczer, Novak, and Mitchison, 1996), and *Xenopus laevis* blastomeres (Y. Masui and P. Wang, 1998). In these organisms, cells do not always divide when they are supposed to, giving rise to a distribution of cell-cycle periods that is not unimodal, but that exhibits secondary peaks at multiples of the cell-cycle period.

This quantised behaviour cannot be explained by the usual factors responsible for the heterogeneity of the cell cycle, such as parameter inhomogeneities and the age distribution of cells within a population. Those factors, which are undoubtedly present in any dividing cell population, would only lead to broadening of the period distribution but not to poly-

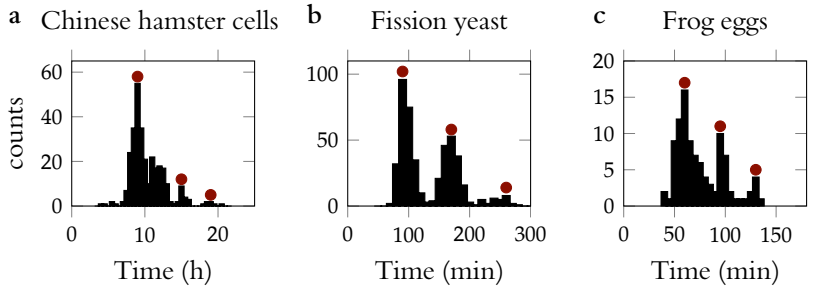

Figure 1.4: **Experimental histograms of polymodal cell-cycle duration.** Polymodal cell-cycles have been observed in Chinese hamster cells (a), fission yeast cells (b), and *Xenopus laevis* frog eggs (c). Figures adapted from Refs. Klevecz (1976), Sveiczer, Novak, and Mitchison (1996), and Y. Masui and P. Wang (1998).

modality. Therefore, detailed mathematical models with a large number of biochemical components (on the order of 10) have been proposed to explain this behaviour (Sveiczer, Csikasz-Nagy, et al., 2000). In those models, period skipping arises already at the deterministic level (i.e., in the absence of sources of heterogeneity and inhomogeneity, (Novak et al., 2001)), while noise is sometimes considered to reproduce the level of irregularity observed in the experiments (Steuer, 2004). Other striking examples of polymodal cycles embedded in an otherwise oscillatory dynamics were reported long ago in sensory neurons (Rose et al., 1967) and bacterial motility (Schimz and Hildebrand, 1992).

Using the phenomenologies described above as motivation, we have addressed the general question of how a limit cycle behaviour with polymodal period distribution can arise in minimal oscillator models (Chapter 3). To that end we consider one of the most basic oscillator architectures, namely an activator-inhibitor circuit operating in an excitable regime and subject to molecular noise. Turcotte, Garcia-Ojalvo, and Süel (2008) reported a noise-induced stabilisation effect on a similar system that generated excitable excursions with a polymodal distribution of relaxation times. We show that the behaviour described by Turcotte, Garcia-Ojalvo, and Süel (ibid.) is general to simple activator-repressor genetic circuits and that it can robustly coexists with the another noise-induced effect, namely stochastic coherence (Gang et al., 1993; Lindner et al., 2004; A. S. Pikovsky and Kurths, 1997), by which excitable excursions occur in a highly periodic manner.

Dynamics underlying embryonic stem-cell pluripotency

Embryonic Stem (ES) cells are cells derived from the epiblast tissue in inner cell mass (ICM) of the mammalian blastocyst (see Figure 1.5, for a review see A. G. Smith, 2001). ES cells are characterised by two distinctive aspects: their potential to differentiate to any adult cell type in response to proper external signals —which is termed pluripotency— and their ability to self-renew indefinitely in culture. Both properties are what make ES cells appealing from both the biomedicine and developmental biology points of view. Pluripotency is, in fact, a short transient cellular state in developing embryos in vivo. Nonetheless, cultured clonal populations of these cells can be kept in the pluripotent state in vitro thanks to the activity of a number of factors (Chambers and Tomlinson, 2009; Young, 2011) whose interactions with the cell still need to be fully understood. Indeed, hundreds of studies over the last years have grounded the idea that pluripotency in ES cells is maintained through the activity of a small set of transcription factors centred around Oct4 and Nanog, which control the expression of self-renewal and differentiation genes.

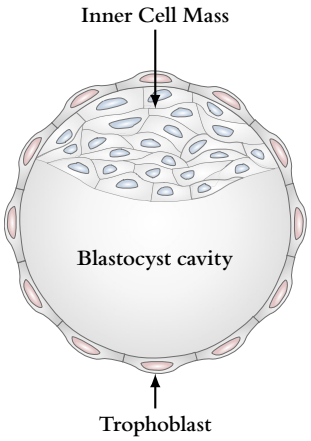

Figure 1.5: **ES cells are derived from the mammalian blastocyst.** The blastocyst is composed of the trophoblast —the outer-most layer of cells—; the blastocyst cavity; and the Inner Cell Mass, where ES cells are produced.

Nanog, a transcription factor which is essential for the establishment of the pluripotent state both in ES cells and artificially induced pluripotent stem (iPS) cells, is thought to be also crucial in pluripotency maintenance. Even though this central role attributed to Nanog, its expression is highly heterogeneous amongst populations of pluripotent cells (Chambers, Silva, et al., 2007; Kalmar et al., 2009; Martinez Arias and Brickman, 2011) cultured in standard conditions. In this sense, Chambers, Silva, et al. (2007) used an ES cell line containing a fluorescent reporter of the *nanog* promoter activity to show that its expression was both dynamic and highly heterogeneous. When mouse Embryonic Stem (mES) cells are cultured in standard self-renewal conditions (in Serum and LIF (S+L)), most of them exhibit high levels of Nanog expression (a state referred to as HN) but a relatively small subpopulation displays low levels (termed LN, for low Nanog) which correlate with a higher probability of differentiation (Figure 1.6a and see Kalmar et al., 2009). In these conditions, a third group of cells, transiting from high to low levels of Nanog (and vice versa) can also be identified (named here MN after middle Nanog, in Figure 1.6). Furthermore, flow cytometry studies using fluorescence-activated cell sorting (FACS) showed that those cells expressing low Nanog were able to de-differentiate back into fully pluripotent cells expressing high Nanog levels (Singh et al., 2007). Such

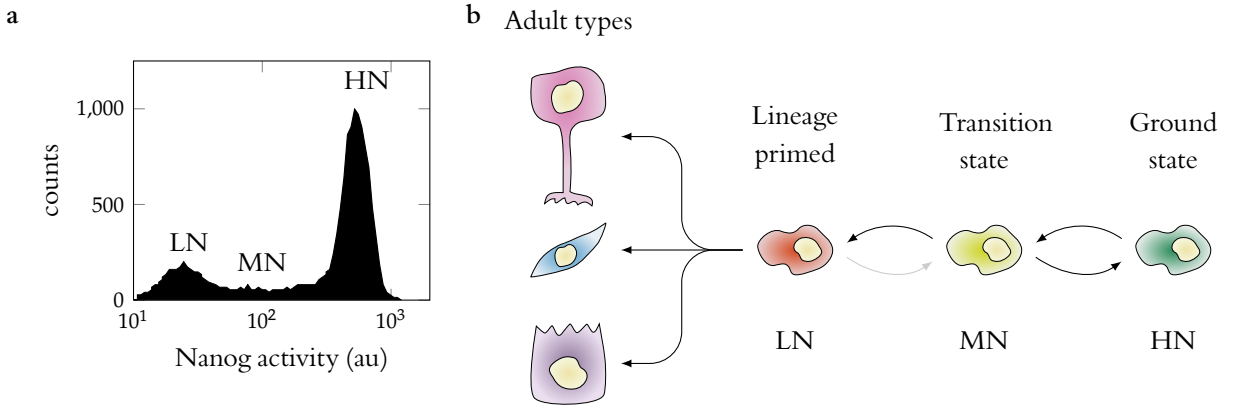

Figure 1.6: **Nanog heterogeneity and pluripotency.** (a) Stationary distribution of the reporter of *nanog* promoter activity as measured by FACS, adapted from Kalmar et al. (2009). (b) Schematic representation of the dynamic pluripotent state whereby cells in a high Nanog (ground) state, HN, dynamically transit to lower Nanog activity states (MN) and eventually prime for differentiation (LN) to adult cell types.

dynamic de-differentiation process was later observed by means of time-lapse fluorescence microscopy by Kalmar et al. (2009).

A simple view of these states is summarised in Figure 1.6b. As shown in this Figure, Nanog levels establish two main cell populations: HN, in which cells are in a robust or naive pluripotency state (also referred to as ground state); and LN in which cells are in a primed pluripotent state, ready to commit to lineage specification. Cells are known to jump between these states in a reversible manner, thus creating a third (dynamic) state of transition (MN).

In the light of these observations, Kalmar et al. (ibid.) proposed a mathematical model involving the pluripotency factors Nanog and Oct4, using interactions between these proteins previously elucidated in the experimental literature (see Figure 1.7), that supported the existence of a stable HN state from which transient drops of Nanog levels could be triggered via an excitable mechanism.

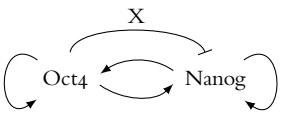

Figure 1.7: **Schema of the excitable circuit regulating pluripotency.** The circuit proposed by Kalmar et al. (2009) is centred around Oct4 and Nanog, which self-activate and activate each other and a third component X (presumably Tcf3) that inhibits Nanog at high doses of Oct4.

Alternative transcriptional models of regulation of the pluripotent state have been proposed. An earlier modelling attempt by Chickarmane et al. (2006) proposed a classic deterministic (bistable) switch between self-renewal and differentiation states as the underlying mechanism of pluripotency regulation. Another interesting work by Glauche, Herberg, and Roeder (2010) compares the stochastic mechanisms of transition to deterministic oscillatory transitions and proposes testable experimental setups.

The model by Kalmar et al. (2009) is of special interest, as it shows that fluctuations arising from the non-linear and stochastic interactions between Nanog and Oct4 (Figure 1.7) reproduce, at the population level,

the observed profile of Nanog expression (skewed and slightly bimodal) shown in Figure 1.6a. By the time the circuit model was presented, little was known about the specific molecular interactions. The existence of an unidentified factor, X, that was responsive to high levels of Oct4 and was an important element of the network was then suggested and its direct interaction with Nanog was also assumed. Posterior reports point out that this factor could be Tcf3, and its activity is now known to be at the transcriptional level rather than through direct protein interaction. Indeed Tcf3 is nowadays recognised as a repressor of the expression of Nanog under the control of Oct4 (Pereira, Yi, and Merrill, 2006).

Based on these recently elucidated facts, we have investigated the dynamics of transcriptional regulation in pluripotency using an updated version of the model by Kalmar et al. (2009). The proposed transcriptional circuit, which accounts for the transcriptional activity of Tcf3, shows a coexistence of the stable state HN of pluripotency with a second state at low Nanog levels which is also stable and that can be assimilated to the cellular lineage priming state, LN (Figure 1.6). That is to say, cells in this stable state might be primed in a reversible manner to specific lineage commitment, as opposed to irreversibly committed. The proposed model tries to link the biological state of the cells (their pluripotency phenotype, in terms of transcriptional activity) with the topology defined by the underlying dynamic system, and thus might reflect the dynamics of cell fate decision-making processes.

It is clear that the pluripotent state is associated with different levels of heterogeneity. It is also true that this state of pluripotency has been traditionally associated with large absolute levels of, among others, the transcription factors Nanog and Oct4. In a recent work (Muñoz Descalzo, Rué, Garcia-Ojalvo, et al., 2012), however, we have brought up experimental observations, using quantitative immunofluorescence (QIF), showing that pluripotency in mES cells is driven by specific ratios between Oct4 and Nanog rather than its absolute levels. These observations indicate that when cells are grown in robust pluripotency promoting conditions —in 2 inhibitors (2i) and LIF (2i+L) conditions, a medium with two inhibitors that confers pluripotency on cells— they exhibit uniform levels of pluripotency and this is associated with a high correlation between the levels of Oct4 and Nanog in individual cells. Indeed, changes in correlations also happen with other key pluripotency factors as the transcriptional repressor Tcf3 or even the multi-functioned

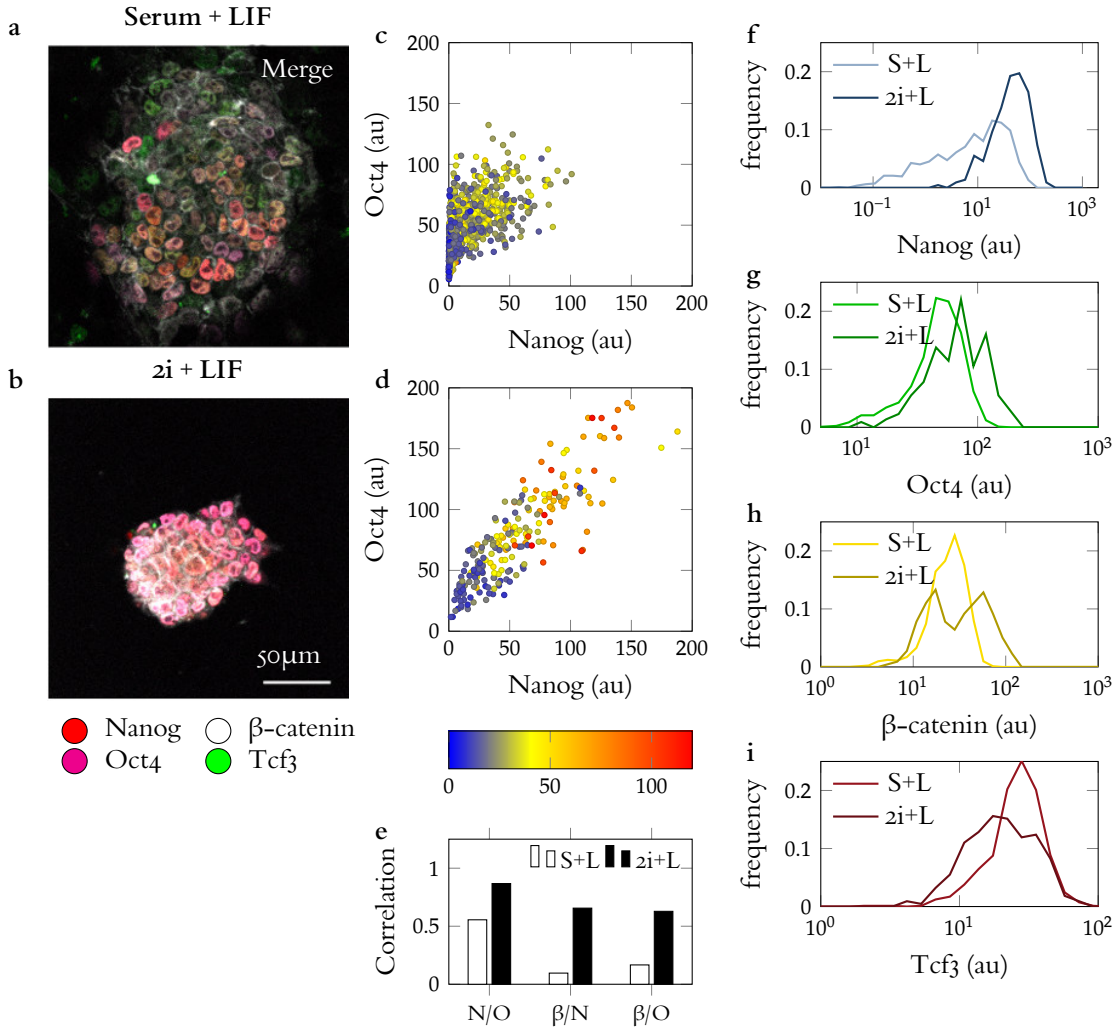

Figure 1.8: **Heterogeneity and correlations in the pluripotency network.** (a,b) Representative images of ES cells immunostained for Tcf3, Nanog, Oct4 and total β-catenin grown in S+L (a) and 2i+L (b). Images courtesy of Silvia Muñoz Descalzo. (c,d) Quantification of Nanog, Oct4 and β-catenin (in colour) from immunostaining of single-cells for naive pluripotency (2i+L,d) and standard (S+L,c) growth conditions. (e,f) Protein distributions for the pluripotency factors in the same conditions. All plots shown in adimensional fluorescence units (au).

protein β-catenin.

Panels a and b of Figure 1.8 show prototypical colonies of ES cells grown in standard (S+L) and naive (2i+L) pluripotency conditions. Fluorescent immunostaining of these cells and its quantification (Figure 1.8c,d) show that Nanog, Oct4, Tcf3 and β-catenin levels are much more inter-correlated in 2i conditions than otherwise (Figure 1.8e). There are some plausible explanations for the high correlations observed, but a strong one is the existence of a post-translational layer of regulation of the pluripotency factors. It is feasible that the existence of such correlations reflects stoichiometric conditions between protein complexes containing Oct4, Nanog, Tcf3 and/or β-catenin and that are optimal for pluripotency. Indeed some biochemical studies suggest the existence of some of these complexes and its role in the maintenance of pluripotency (Abu-Remaileh et al., 2010; Costa et al., 2013; Faunes et al., 2013; Kelly

et al., 2011; J. Wang et al., 2006).

To investigate this possibility and to analyse how the activity of those factors is regulated, we have combined single-cell QIF microscopy and gene expression analysis, together with computational modelling of these factors. In Sections 4.2 and 4.3 we propose a new model of post-translational protein interactions, which complements the well-established transcriptional regulatory layer. Specifically, we find that the activity of a network of protein complexes involving Nanog, Oct4, Tcf3 and β-catenin suffices to account for the behaviour of ES cells under different conditions.

The model, which assumes that protein stabilities are affected by the formation of complexes, reproduces the experimental distributions of mRNA and protein concentrations of the involved factors in a wide variety of mutant ES cell lines and medium conditions and, what is more important, suggests that the function of the network is to buffer the transcriptional activity of Oct4, which appears to be the main determinant of pluripotency. Specifically, the model inversely relates pluripotency to the levels of transcriptionally active Oct4, and envisages that the function of the protein network is to buffer these levels in a dynamic manner. This novel view of pluripotency is supported by a bioinformatic reanalysis we have performed on existing data of Nanog and Oct4 function and on their genetic interaction.

All in all, the post-translational model of pluripotency regulation investigated in Section 4.3, founded on solid quantitative experimental data, brings new insight to the molecular basis of mouse ES cell pluripotency.

1.2 *Relaxation and noise in cellular signalling networks*

Single-celled organisms and cells belonging to multicellular organisms sense external signals from their surrounding mircoenviroment and respond to them by adapting their metabolism and patterns of gene expression (Figure 1.9). A large part of the signals affecting the cell are small molecules (ligands) that are sensed at the plasma membrane by transmembrane proteins (receptors) such as receptor tyrosine kinases (RTKs) or G protein-coupled receptors (GPCRs). Upon binding to the receptors, intracellular signal transduction takes place by means of large chains of cytosolic biochemical reactions between proteins that convey the signal to the nucleus, where genetic programs are modulated in a signal-specific manner. Although intracellular signalling is a directed process —from cell surface to the nucleus— it is by no means linear and

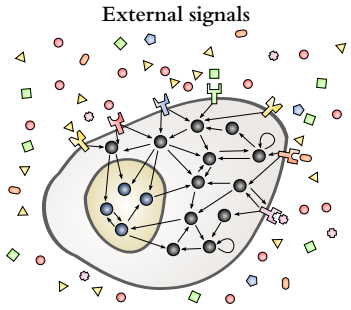

Figure 1.9: **Schema of a cell affected by external signals.**

involves a large, tangled network of protein interactions (Kholodenko, Hancock, and Kolch, 2010; Oda, Matsuoka, et al., 2005).

The analysis of isolated small genetic and protein-protein interaction modules, as the ones introduced above, is motivated by the growing evidence that most cellular processes are regulated by highly modular networks of interactions. Yet, most experimentally based protein-protein interaction and cellular signalling networks exhibit a high degree of connectivity among all their parts (Oda, Matsuoka, et al., 2005) and their study can only be addressed from a global perspective, as opposed to a reductionist, motif-based, approach.

One type of simplified description of regulatory networks that aims at maintaining only the essential ingredients of the biochemical interactions are Boolean networks (BNs). They were introduced by Kauffman (1969) as simple models to support the idea that life does not require precisely programmed units to generate complex and diverse behaviours, but rather random juxtaposition of simple units could achieve the same result. In this type of model, the network units (nodes) are genes or proteins that are either fully active or inactive. The activity of the whole system results from the dynamical interaction of the nodes dictated by a set of logic rules. The advantage of this type of model is that it is simple and requires much less parameters than kinetic-based models (or even none).

Most of studies dealing with BNs to date have considered random network topologies (the so called random Boolean networks (RBNs), Kauffman, 1993). Properties such as criticality (Derrida and Pomeau, 1986), robustness (Bornholdt and Sneppen, 2000; Bornholdt, 2008; Just, Shmulevich, and Konvalina, 2004; Klemm and Bornholdt, 2005a,b; Ribeiro and Kauffman, 2007), and scalability (Drossel, 2005; Samuelsson and Troein, 2003) have been studied in this type of BNs. However, real gene regulatory or signal transduction networks can be different from uniformly connected networks with random interactions. Furthermore, RBNs are autonomous and in the absence of random effects they evolve deterministically to a well-defined attractor from some specific initial conditions. This feature, which may hold for certain gene regulatory networks, is not fulfilled by signal transduction networks, which are subject to external signals which frequently vary in time. In fact, the main goal of signalling networks is to process the information carried by the input signals, which is usually dynamical, although this latter aspect has not been much considered so far.

Only recently, experimentally-based BNs have started proving to be a good paradigm for modelling certain biological processes. In fact, with the recent progress in the experimental characterisation of biological processes, BN models have been successfully applied to the study of gene regulatory networks (Covert et al., 2004; Davidich and Bornholdt, 2008; Faure et al., 2006; Mendoza, Thieffry, and Alvarez-Buylla, 1999), cellular differentiation (Huang and Ingber, 2000), evolution (Bornholdt and Sneppen, 2000) and signal transduction in the cell (Bauer et al., 2010; Helikar et al., 2008; Saez-Rodriguez et al., 2007; Samaga et al., 2009). Those studies provided evidence that sequences of events can indeed be reproduced by simple discrete dynamic models such as BNs.

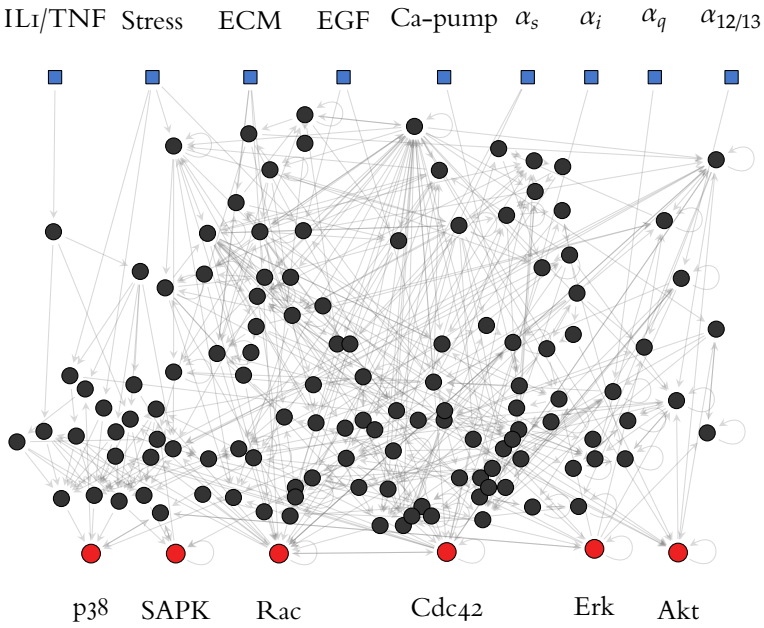

Figure 1.10: **Boolean model of the fibroblast signalling Network.** The network model used in this work consists of 130 internal nodes and 9 input nodes, which represent protein species being either active (1) or inactive (0), connected through 542 interactions. Input signals (blue squares on top) and output nodes (red circles at the bottom) are labelled with abbreviations that are defined in the text. A fully annotated version of this network is described in Figure A.1 and accompanying Table A.13

In this Thesis (Chapter 5) we explore the dynamics and the impact of environmental fluctuations on the information processing and integration capabilities of a cell signalling network as a whole.

To address these issues in a way that explicitly accounts for the complexity of the system under consideration, we use one of the most comprehensive dynamic models of cell signalling currently available in the literature, a recently published Boolean network for the human fibroblast that involves over 130 protein species (Helikar et al., 2008, see Figure 1.10).

The input nodes of this network (blue squares in the Figure) represent signals of varying nature, namely generic and specific stress signals (interleukin-1 and tumour necrosis factor (IL1/TNF-α)), a growth factor (epidermal growth factor (EGF)), an ion channel (calcium pump),

extracellular matrix (ECM) components and the four families of ligands that use GPCRs. These families are known as the α_s, α_i, α_q, and $\alpha_{12/13}$ ligand families, and contain some neurotransmitters, cytokines and hormones. Since the pattern of interactions within each of these families is conserved, the BN model contains a generic ligand/receptor pair for each of them.

Some of the 130 internal network nodes are related to a well-defined cellular response, namely, programmed cell death (Akt), gene transcription (Erk), cytoskeletal regulation (Rac, Cdc42), and cellular stress (SAPK, p38). Thus, these six proteins (shown as red circles in Figure 1.10) are considered to be output nodes of the network, even though they have outgoing edges to other intermediate nodes of the network (see Figure 1.10). Helikar et al. (ibid.) validated their model by comparing the predicted input/output behaviours of different pairs of nodes with existing experimental data, and the network was seen to agree with biological results.

We have investigated, from a physical/computational perspective, the relaxation dynamics of this experimentally-based network and observed that these are robust to environmental fluctuations (Section 5.1). In addition, given the ubiquity of periodic oscillations in cellular processes (Novák and Tyson, 2008), we also considered the case where an input signal was periodic in time (structured signal) rather than constant or erratic (unstructured). In this case, we found that the level of background noise (or chatter) shapes the response of the whole signalling network to external information carrying signals, and modulates the activity levels of the network outputs in a manner that is reminiscent of the well-known stochastic resonance effect (Chapter 5), an effect that seems to be caused by the singularities of the network topology, its functional elements, and the influence of chatter in the integration of signals (Chapter 6).

All considered, these findings indicate that cellular signal transduction and integration processes are dependent on the levels of environmental fluctuations, suggesting a trade-off between robust and context-dependent decision making in these networks. Thus, these results contribute, at least from a theoretical perspective, to the growing evidence that cells might have evolved to cope with and exploit noise and random fluctuations.

1.3 Spatiotemporal population dynamics in bacteria

Cellular self-assembly into three dimensional patterns is an essential biological process that occurs in developing multicellular organisms and microbial communities. Development of organisms and cellular communities depends on the ability of growing cell populations to self-organise in space and time. Such spatio-temporal behaviour appears to be universally governed by two concomitant processes: cellular genetic programs and mechanical forces. Researchers have identified genetic programs that guide the organisation of large amounts of individual cells during development. In addition, recent studies have also suggested that physical and mechanical processes can play a critical role during development of organs (Mammoto and Ingber, 2010).

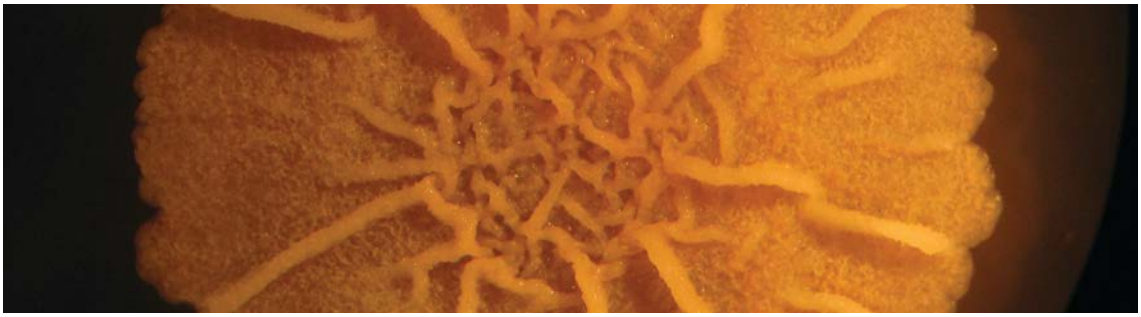

Figure 1.11: ***B. subtilis* biofilm.** A mature biofilm displaying the complex wrinkled spatial structures. Image courtesy of Munehiro Asally.

At the lowest level of biological organisation, unicellular organisms such as the soil bacterium *Bacillus subtilis* can assemble into multicellular convoluted structures called biofilms (Figure 1.11), even starting from a single cell (Branda et al., 2001; Webb, Givskov, and Kjelleberg, 2003). In *B. subtilis* colonies growing on flat surfaces, cells adhere each other by producing an external scaffolding mainly made of proteins and polysaccharides, known as ECM. ECM glues bacteria together to form a macroscopic aggregate (Webb, Givskov, and Kjelleberg, 2003), visible at naked eye, that provides the colony with protection against environmental extremes and chemical agents, as well as mechanical support (Kearns et al., 2005; Marvasi, Visscher, and Casillas Martinez, 2010). Rather than growing as flat pellicles, the external scaffold bends and folds onto itself generating intricately wrinkled structures (see Figure 1.11), which are also thought to aid in the transport of soluble nutrients and waste (Wilking et al., 2013) and prevent the community from permeation by fluids (Epstein et al., 2011).

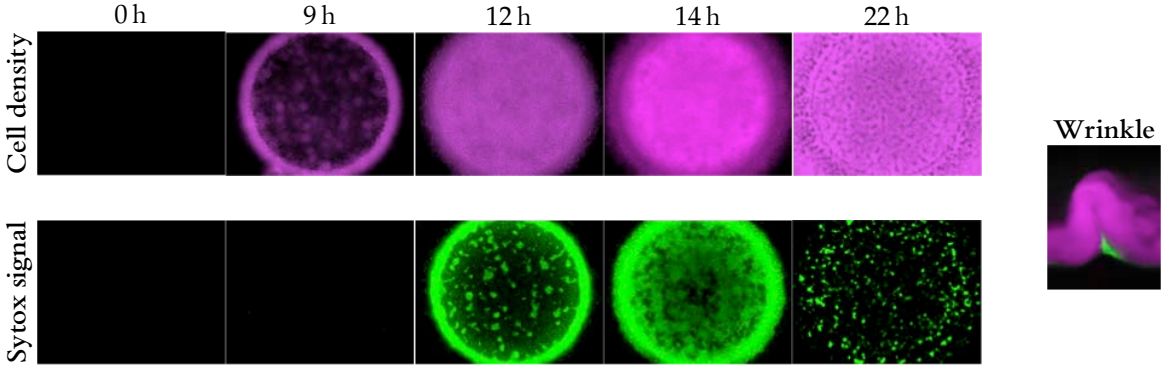

Figure 1.12: **Population dynamics during biofilm formation.** Film strip of a developing *B. subtilis* colony. The top panel shows (in magenta pseudocolor) a cell density marker and the middle panel shows the a cell death reporter Sytox (in green). Cross-section image of a 30 h old biofilm wrinkle.

We have recently unveiled a spatio-temporal pattern of massive bacterial death (Asally et al., 2012) which, in combination with mechanical forces exerted by colony growth and the stiffness conferred by ECM, drives wrinkle formation in developing biofilms. Indeed, using fluorescence time-lapse microscopy of *B. subtilis* biofilms that contained two fluorescent reporters: one for tracking dividing cells (and thus cellular density) and another specific of cellular death, we detected that cell death and wrinkles are spatially correlated (see Figure 1.12). Furthermore, cross-sections of biofilms revealed that cell death is localised at the bottom of biofilms, and more specifically at the centre of the folded structures (rightmost panel in Figure 1.12).

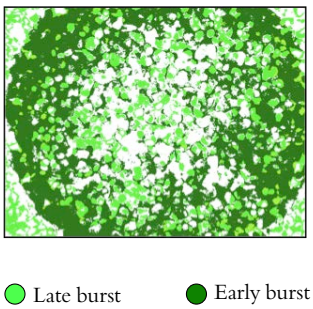

Figure 1.13: **Population dynamics during biofilm formation.** Spatio-temporal patterns of cell death consist of an early- and a late-burst of massive bacterial death. The early-burst appears in the form of isolated spots while the late-burst death occupies approximately the complementary space.

To establish the direction of causality of the observed cell death-wrinkle relation (namely, if wrinkles results in or from the cellular death), we measured the spatio-temporal dynamics of cell death during colony growth and biofilm development. These measurements uncovered a striking heterogeneous distribution of cell death that occurs before wrinkle formation (see Figure 1.13).

The death reporter uncovered a pattern of cellular disintegration that emerges early on during biofilm formation and changes over time (see bottom row in Figure 1.12). Indeed, two distinct cell death patterns (CDPs) can be recognised that are approximately separated in time and do not overlap in space (Figure 1.13). The first pattern (early bursts) is generated approximately 12 h into biofilm formation and includes a prominent ring of cell death that represents the edge of the drop of cells applied to the agar plate and appears due to the coffee ring effect (that is, faster liquid evaporation at the edge of the drop generates a volume compensation flow away from its centre that transports and accumulates

cells towards the boundaries (Yunker et al., 2011)). The second death pattern (late bursts) starts after the early bursts subside and are confined to small round regions (spots or patches) that are distributed throughout the colony, surrounded by regions where early bursts had previously occurred (Figure 1.13). These results indicate that cell death may be driven by local differences in cell growth and density that arise during biofilm formation. A similarly heterogeneous CDP is generated even when the biofilm is started from a single cell. Together, these data suggest that the replicating cell population during biofilm development gives rise to a heterogeneous CDP that may be genetically controlled. This connection between cell death and wrinkling is explained by our finding that localised cell death focuses mechanical forces in space —thereby converting lateral compressive stress into vertical buckling of the biofilm— and consequent wrinkle formation, as schematised in Figure 1.14. Therefore, macroscopic wrinkle structures do not seem to be caused by local growth as commonly assumed, but rather from massive, but localised cellular death.

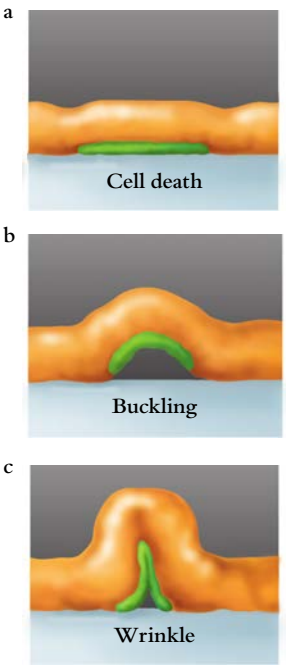

Figure 1.14: **Schema of the formation of a wrinkle in a *B. subtilis* biofilm.** Exhaustive analysis of biofilms —here shown from a cross-sectional view point— revealed accumulation of cellular death at the bottom of wrinkle formation areas, which triggers cell buckling and development of wrinkled protuberances. Reprinted from The Scientist with permission of Precision Graphics.

What still remains poorly understood is how these patterns of cellular death emerge from the mechanical and genetic interactions at the cellular level. Experimental results suggests that these macroscopic structures are originated at the mesoscopic level, where cellular growth and death events can give raise to spatial inhomogeneities in cell density. In Chapter 7 of this Thesis we propose a reaction-diffusion population dynamics model that includes both the effect of ECM production and accumulation of bacterial waste.

The model thus accounts for the basic spatio-temporal dynamics of the wild type *B. subtilis* CDP. Additionally, when matrix production is turned off in the model, the simulations generate a homogeneous CDP similar to that observed in a strain lacking *epsH*, a gene involved in the synthesis of extracellular polymeric substances and thus in ECM formation.

Strikingly, the lack of heterogeneity in the CDPs has a direct impact in the final biofilm structure, that appears with no perceptible wrinkles (cf. panels a and b of Figure 1.15) Therefore, this population-level model accounts for the observed heterogeneity of the CDP, and thereby provides a conceptual explanation for how the ECM can promote a localised pattern.

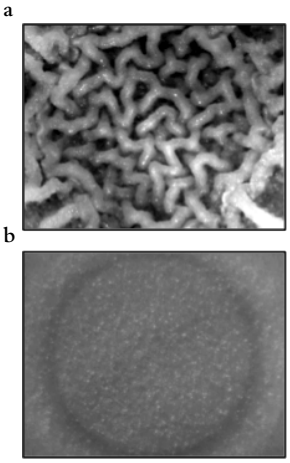

Figure 1.15: **Matrix components and biofilm morphology.** (a) Wild type cells generate the characteristic wrinkles. (b) Cells lacking a gene involved in extracellular matrix formation do not generate the wrinkles. Courtesy of Munehiro Asally.

II

Dynamics of transient stochastic cellular processes

2

Noisy excitable gene circuits with transient regulatory response

Gene expression has long been described as a process involving random bursts of messenger RNA (mRNA) synthesis and causing large variations of protein concentration among genetically identical cells. Unregulated protein bursts occur as a consequence of promoter activation kinetics and can, in some cases, create large long-lasting fluctuations in protein concentrations that can propagate through downstream genes and can even be inherited by daughter cells after cell divisions. More appealing is the idea that noise interacts with complex non-linear dynamic regulatory elements to generate constrained variability among otherwise identical cells. This is the case of genetic circuits with noise-driven excitable dynamics in which stochastic perturbations from the steady state can launch the system into large amplitude transient responses.

We devote this chapter to the study of different minimal genetic mechanisms of excitability, its dynamic properties and the appearance of non-trivial patterns of behaviour induced by noise. Given that the type of excitability determines the statistics of the noise-driven pulses, it is important to establish, in the context of gene regulation dynamics, the relationship between the gene circuit architecture and the phase-space topology resulting in excitability. A previous study by Guantes and Poyatos (2006) has linked the specific dynamical features of two-gene oscillators to the particular molecular implementation of one of the regulatory links, associating transcriptional regulation with type I dynamics, and post-translational regulation with type II. Here we show that both types of excitability are possible in both architectures, and add a third one that has been identified in bacteria, for which again both types of excitability are possible. Thus, our results indicate that circuit architecture does in principle not constrain the type of dynamics that the system exhibits.

The three circuits under consideration are described in Section 2.1,

followed, in the subsequent Sections (Sections 2.2 to 2.4) by detailed bifurcation analyses of the three systems for two parameter sets each, which lead to the two types of dynamics described above. Finally, in Section 2.5 we explore the parameter space to confirm that both types of excitable behaviours can be achieved for a wide range of parameter values.

2.1 *A catalogue of excitable genetic circuits*

We consider the three circuits shown in Figure 2.1. Circuit A in this Figure has been identified as the core module underlying transient differentiation into competence in *B. subtilis* cells when placed under nutritional stress (Süel et al., 2006). In this circuit, a protein X activates its own transcription and represses the transcription of a second protein Y. The bistable switch on X resulting from the positive auto-feedback loop is kept off by means of a third protein, P, with protease activity, that degrades X and also Y. This competitive degradation of Y and X—which is depicted in Figure 2.1 as Y-modulated repression of the P-mediated degradation of X— leads to excitable behaviour, as shown below in this Section. This simple circuit can be expressed in terms of biochemical reactions (see Appendix A.1 for details) and the macroscopic description of the circuit's dynamics can be expressed in terms of two coupled ordinary differential equations for the concentrations of X and Y (x and y respectively) given by:

$$\frac{dx}{dt} = a_1 + \frac{b_1 x^n}{k_x^{\ n} + x^n} - c_1 \frac{x}{1 + x/\ _1 + y/\ _2} - d_1 x, \tag{2.1a}$$

$$\frac{dy}{dt} = a_2 + \frac{b_2}{1 + \ x/k_y^{\ m}} - c_2 \frac{y}{1 + x/\ _1 + y/\ _2} - d_2 y. \tag{2.1b}$$

The details of the derivation of these equations from a set of reaction rates can also be found in Appendix A.1.

The other two gene circuits investigated in this work are standard activator-repressor systems, in which the activator species forms a direct positive feedback loop through its auto-regulation and an indirect negative feedback loop by means of the activation of its repressor (see circuits B and C in Figure 2.1). Specifically, in both circuits B and C an activator protein X binds to and activates its own expression, and the expression of a repressor protein Y. The repressor protein, in turn, inhibits the expression of the activator species in one of two ways:

i post-translationally, by inactivating directly the activator protein

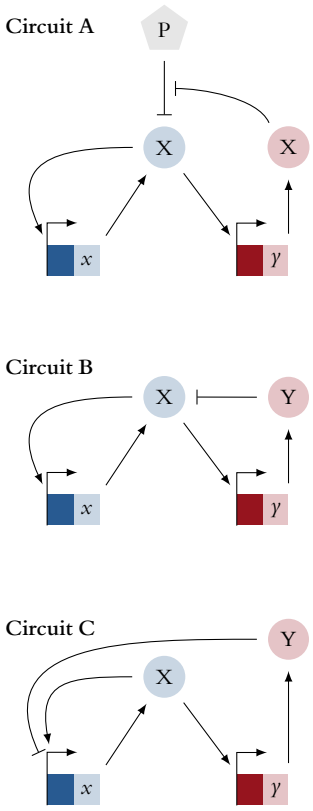

Figure 2.1: **Minimal gene circuits with excitable dynamics.** Three excitable gene circuits. In the circuit diagrams, uppercase letters and circles represents proteins, and lowercase letters and rectangles denote genes.

(degrading it, for instance), such as in circuit B, or

ii transcriptionally, by binding to the promoter of X and inhibiting its expression, which is the case of circuit C.

The deterministic equations of these two circuits are:

$$\frac{dx}{dt} = a_1 + \frac{b_1 x^n}{k_x{}^n + x^n} - \quad xy - d_1 x, \tag{2.2a}$$

$$\frac{dy}{dt} = a_2 + \frac{b_2 x^m}{k_y{}^m + x^m} - d_2 y, \tag{2.2b}$$

and

$$\frac{dx}{dt} = a_1 + \frac{b_1 x^n}{k_x{}^n + x^n + \ \gamma y^{\ p}} - d_1 x, \tag{2.3a}$$

$$\frac{dy}{dt} = a_2 + \frac{b_2 x^m}{k_y{}^m + x^m} - d_2 y, \tag{2.3b}$$

respectively. Again, details of the derivation of these two sets of equations can be found in Appendix A.1.

These activator-repressor modules are frequent in natural circuits. In particular, this is a generic motif for cell-cycle and circadian oscillators, which need to generate highly regular periodic activity. These two circuits have also been implemented synthetically. The post-translational activator-repressor module (circuit B) has been constructed synthetically in *B. subtilis*, where it has been shown to be functional in the development of competence (Çagatay et al., 2009), while the transcriptional activator-repressor module (circuit C) was built for instance in *E. coli* (Atkinson et al., 2003). Additional synthetic versions of this latter type of circuit have been built that either include a negative self-feedback of the repressor on itself (Stricker et al., 2008), or where the repressor action is mediated by quorum sensing (Danino et al., 2010).

In the next three Sections we analyse sequentially the potential dynamical regimes that these circuits exhibit, for different parameter values. As we show below, all three circuit architectures are able to display the two types of excitable dynamics represented in Figure 1.3. To verify this, we show the bifurcation diagrams leading to excitability in all cases, using the unregulated, basal expression of protein X (a_1) as bifurcation parameter. A qualitative comparison between the excitable time traces generated by the discrete stochastic simulation of these models in the three circuits is also made.

2.2 Circuit A: a competitive degradation circuit

Type I excitability close to a saddle-homoclinic bifurcation

For the set of parameters given in Table A.4 (determined by Suel et al. (2007) by comparison with experimental observations), the competitive degradation circuit libelled as circuit A in Figure 2.1 exhibits excitable dynamics. Figure 2.2a shows the phase portrait of the deterministic model of that circuit (see Equation (2.1)) in the excitable regime.

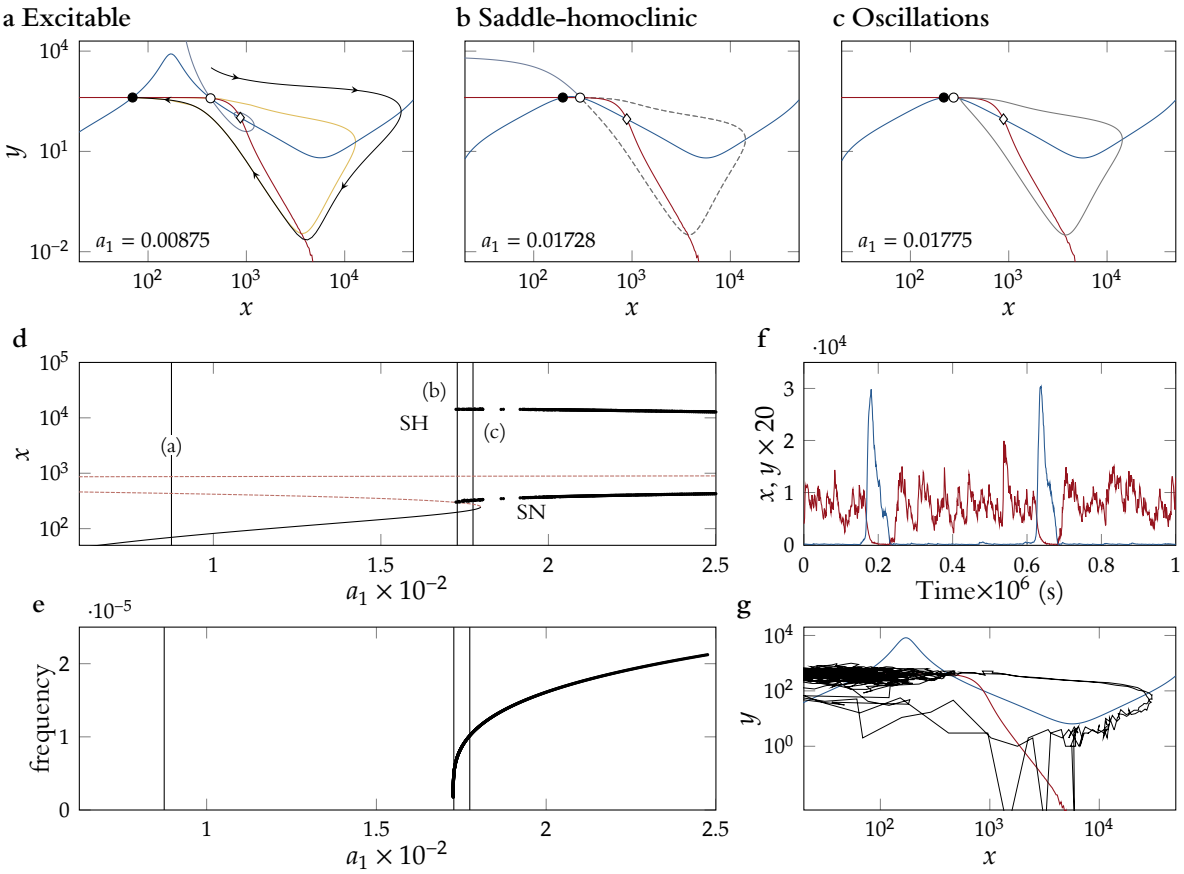

Figure 2.2: **Type I excitability close to a saddle-homoclinic bifurcation in circuit A.** (a-c) Phase portraits for three values of the parameter a_1. Nullclines of x (blue) and y (red) and the stable (grey) and unstable (yellow) manifolds of the saddle point are represented. (d) Bifurcation diagram of the system for varying a_1. Vertical lines correspond to values shown in plots (a-c). (e) Frequency of the oscillations versus a_1. (f) Time traces of protein levels for both species, as generated from stochastic simulations of its constituent reactions. (g) Corresponding stochastic trajectory in the phase plane. The values of the parameters used here are given in Table A.4.

The plot depicts the x and y nullclines (in blue and red, respectively, in plots a-c and g), the stable and unstable manifolds of the saddle, and a sample deterministic trajectory (thin black line, see figure caption for a description of the other lines used). The excitable regime is characterised by a stable fixed point with low levels of X (black circle) that coexists with two unstable fixed points, a saddle (white circle) and a spiral (white diamond). The upper part of the stable manifold of the saddle (grey curve) is the excitability threshold: any trajectory starting at the right of that manifold is forced to go through a large excursion in phase space (note the logarithmic scale in the plot's axes) towards increasing x, leading to an excitability pulse.

As the basal expression a_1 of X increases, a saddle-homoclinic (SH) bifurcation takes place, at which the saddle point collides with a stable limit cycle that surrounds the unstable spiral point (Figure 2.2b), giving rise to a homoclinic orbit (dotted grey line in panel b). For a small parameter range the limit cycle coexists with the stable node and the saddle (Figure 2.2c), until both collide in a Saddle Node (SN) bifurcation, after which the limit cycle (solid grey line) is the only attractor of the system. This can be seen in the bifurcation diagram shown in Figure 2.2d, which represents the limit cycle as a thick black line, and the stable and unstable fixed points as thin solid black and dashed red lines, respectively.

The corresponding behaviour of the frequency of the oscillations with respect to the parameter a_1 is represented in Figure 2.2e. The plot shows that the oscillation frequency at the right of the saddle-homoclinic bifurcation tends to zero as the bifurcation is approached, which is a clear hallmark of type I excitable behaviour (the frequency values close to zero are not shown due to numerical constraints). Time traces generated via stochastic simulations of the set of reactions forming the circuit (described in Appendix A.1), using the standard exact discrete stochastic simulation method (Gillespie, 1977), are displayed in Figure 2.2f. The Figure shows that the two species participating in the dynamics are anticorrelated, with wide pulses of activity of X leading to dips in the amount of Y protein present in the cell. The corresponding trajectory in phase space is shown in Figure 2.2g. Previous work by (Çagatay et al., 2009) has revealed that this circuit design leads to highly variable pulse durations, which have the physiological function of allowing cells to cope with the natural unpredictability associated with low extracellular DNA levels to which *B. subtilis* cells are subject in the presence of nutritional stress.

Type II excitability close to a supercritical Hopf bifurcation

The same circuit A with a slightly different parameter set (see Table A.4) leads to a different bifurcation scenario that also leads to excitability. The phase portrait of the system for the excitable regime in this case is shown in Figure 2.3a. The nullclines now cross only once, corresponding to a stable spiral point which is the rest state of the excitable dynamics. Even though there is no saddle here that helps us define an excitability threshold, a large enough perturbation away from the stable fixed point is able to elicit a large excursion in phase space, an example of which is shown as a thin black line in Figure 2.3a.

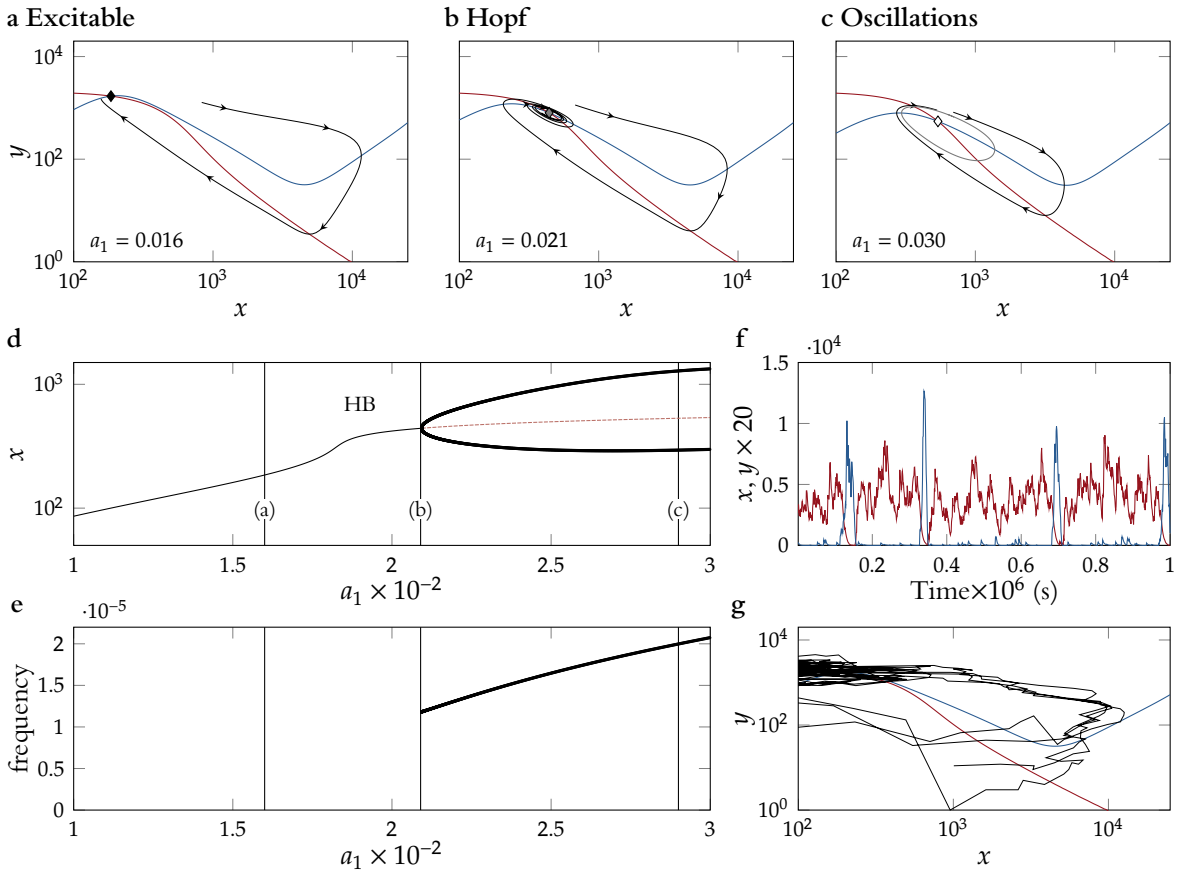

Figure 2.3: **Type II excitability close to a supercritical Hopf bifurcation in circuit A.** The meaning of the plots is the same as in Figure 2.2.

As the value of a_1 increases, the system undergoes a supercritical Hopf bifurcation (HB) —see Figure 2.3b that destabilises the spiral point and generates a stable limit cycle around it, which is born with zero amplitude, as can be seen in the bifurcation diagram of Figure 2.3d. Beyond the HB, the system exhibits a well-developed stable limit cycle (Figure 2.3c). Certain perturbations of this limit cycle evoke large-amplitude excursions in phase-space, similarly to what happens in the standard excitable regime shown in Figure 2.3a. An example of such an excursion is shown by a thin black line in Figure 2.3c. This regime can be interpreted as an excitable dynamics with subthreshold oscillations, something that is well-known to exist in the context of neural systems (Sancristóbal, Sancho, and Garcia-Ojalvo, 2010).

The frequency response of the circuit to a variation in a_1 is shown in Figure 2.3e. In this case one can see that the frequency remains finite at the bifurcation, which is characteristic of type II excitability, as mentioned above. This behaviour is a direct consequence of the fact that the underlying bifurcation bridging excitability and oscillatory dynamics is

a supercritical HB bifurcation. Typical time traces of the protein levels of X and Y are shown in Figure 2.3f. Again, as in the previous Section, there is anticorrelation between the two proteins during the excitation pulse, although in contrast with the previous case, there is a large variability in the pulse amplitude. This is due to the fact that the pulse trajectories in this case depend much more strongly on the initial conditions of the pulse than in the case of the previous Section (cf. Figure 2.2g and Figure 2.3g).

2.3 Circuit B: a post-translational activator-repressor circuit

Type I excitability close to a SNIC bifurcation

We now turn to the activator-repressor circuits, beginning with the post-translational one, in which protein Y represses X by interfering directly, at the protein level, with its activity (for example by enzymatically degrading it). For a set of parameters (given in Table A.4) that successfully reproduce experimental observations (Çagatay et al., 2009), this system also behaves in an excitable manner. The corresponding phase portrait is shown in Figure 2.4a. Again, as in the case of type I dynamics for the circuit A, the system has three fixed points, only one of which is stable. Perturbations away from this stable fixed point, that take the trajectory past the stable manifold of the saddle (which acts again as an excitability threshold, see black line in Figure 2.4a), will cause a large excursion in phase space surrounding the two unstable fixed points, as represented by a thin black line in the Figure.

As the value of the parameter a_1 increases, a global bifurcation will again generate a limit cycle with a finite amplitude (Figure 2.4b). In contrast with the type I case in circuit A however, the bifurcation is now a saddle node on invariant circle (SNIC), since the stable node and the saddle that exist in the excitable regime collide with each other and annihilate exactly on top of the newly born limit cycle (Figure 2.4b). The resulting fully developed limit cycle that exists after the bifurcation is represented in Figure 2.4c, and the corresponding bifurcation diagram is shown in Figure 2.4d. The SNIC bifurcation is revealed in this plot by the simultaneous occurrence of the SN bifurcation, through which the stable and unstable fixed-point branches annihilate with each other, and the appearance of the stable limit cycle, at exactly the same value of a_1.

As summarised in the scheme of Figure 1.3, SNIC bifurcations are associated with type I excitable dynamics, for which the frequency tends

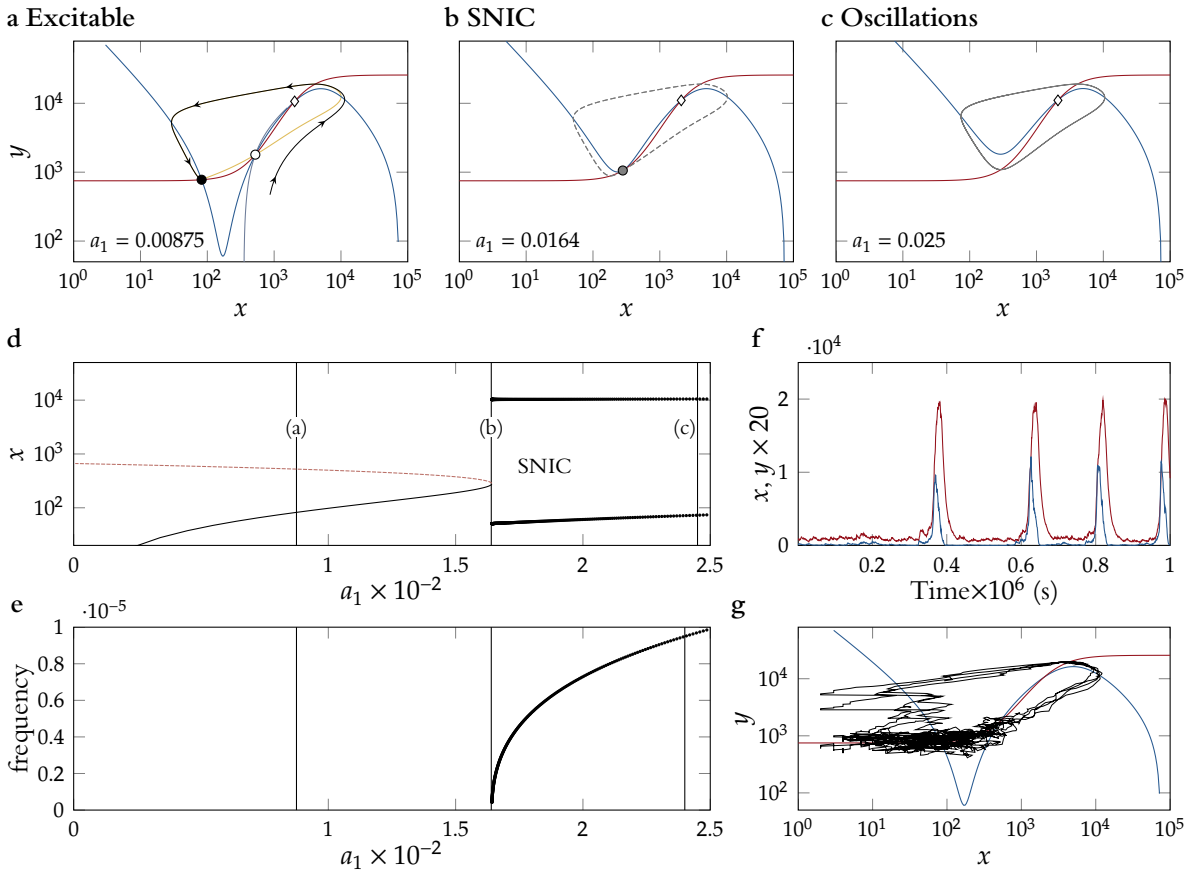

Figure 2.4: **Type I excitability close to a SNIC bifurcation in circuit B.** The meaning of the plots is the same as in Figure 2.2.

to zero at the bifurcation point. This fact is shown in Figure 2.4e. Time traces and phase-plane trajectories corresponding to the excitable regime are presented in Figure 2.4f and g. Note that in this case there is a positive correlation between the two protein species of the circuit, with the repressor slightly tailing behind (shutting off, in fact) the activator. The positive correlation is a reflection that the trajectory in phase space is counterclockwise (see Figure 2.4a). Again, as in the case of circuit A (Section 2.2), there is little dispersion in the pulse amplitudes. In contrast with that other circuit, however, the dispersion in the pulse widths is known to be here much smaller (Çagatay et al., 2009).

Type II excitability close to a subcritical Hopf bifurcation

The post-translational activator-repressor circuit represented as circuit B in Figure 2.1 can also exhibit type II dynamics. The situation is represented in Figure 2.5, for which the corresponding parameters are again given in Table A.4.

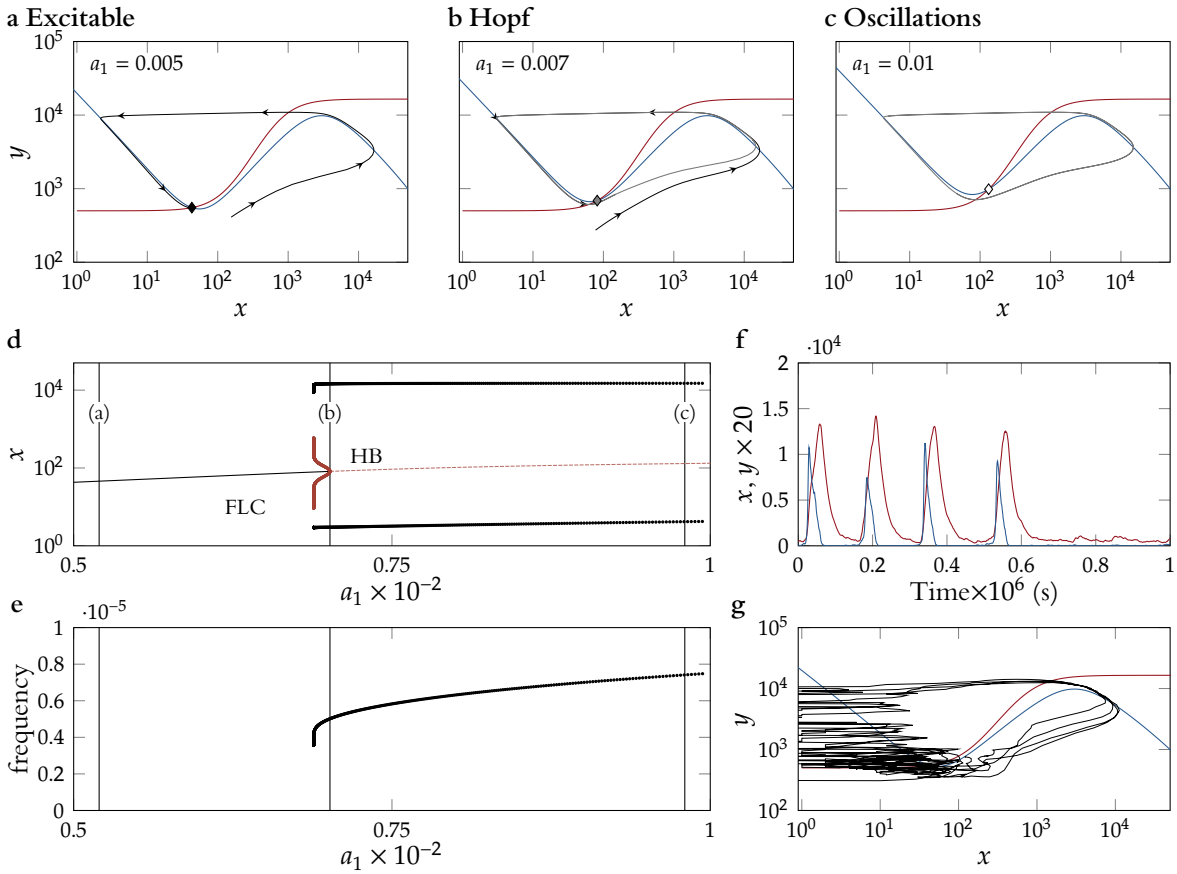

Figure 2.5: **Type II excitability close to a subcritical Hopf bifurcation in circuit B.** The meaning of the plots is the same as in Figure 2.2.

As shown in Figure 2.5a, the excitable regime is characterised by a single fixed point. The cubic nullcline of the activator and the monotonically increasing nullcline of the inhibitor make this system analogous to a FitzHugh–Nagumo model (Figure 2.5). As in the FitzHugh–Nagumo model, the dynamics exhibit separation of time scales between the two species, in such a way that the dynamics of the activator is much faster than that of the inhibitor. This is reflected in a quick entry into the excitable pulse and a slower exit, the latter governed by the final part of the trajectory in which the system approaches the fixed point again by slowly sliding down the left branch of the activator nullcline (see Figure 2.5a). In this case, however, the separation of time scales is not imposed by tuning the parameters (both species have the same linear degradation terms), but effective and due to the post-translational inactivation of X by Y. As

the parameter a_1 increases, excitability is once more lost, this time via a subcritical HB bifurcation in which the fixed point loses stability via a collision with an unstable limit cycle (Figure 2.5b). This unstable invariant set was previously generated via a fold limit cycle (FLC) bifurcation, together with a stable limit cycle which becomes the only stable attractor (Figure 2.5c) after the subcritical Hopf bifurcation occurs. The situation is summarised in the bifurcation diagram shown in Figure 2.5d where the thick red line represents the unstable limit cycle branch.

The subcritical Hopf bifurcation described above leads to type II dynamics, which we remind is characterised by the fact that the frequency remains finite at the bifurcation. This is shown in Figure 2.5e. Finally, Figure 2.5f and g shows sample trajectories of the system as generated from stochastic simulations of the set of reactions. In this case, as in the type II case of circuit A, the variability of pulse amplitudes is slightly higher than for type I excitability.

2.4 Circuit C: a transcriptional activator-repressor circuit

Type I excitability close to a SNIC bifurcation

We now substitute the post-translational repression of X by Y for a transcriptional regulation, in which the repressor competitively inhibits expression of the activator by binding to its promoter (circuit C in Figure 2.1). In this case the system can also behave as an excitable system, as shown in Figure 2.6a. As in the previous examples of type I excitability discussed above for circuits A and B, this system has three fixed points, only one of them being stable. Once more the stable manifold of the saddle defines the excitability threshold, beyond which a pulse of excitability is triggered. An increase in a_1 leads to a SNIC bifurcation in which the stable node and the saddle collide and annihilate at the same location at which a limit cycle appears (Figure 2.6b). The corresponding bifurcation diagram is shown in Figure 2.6d.

As expected, the frequency of the limit cycle (shown in Figure 2.6c) decays to zero as the bifurcation is approached, as displayed in Figure 2.6e. The time traces (Figure 2.6f) of the stochastic simulations and corresponding trajectory in phase space (Figure 2.6g) show that the two species are positively correlated, and that there are slight variations in the pulse amplitudes (in contrast with the other examples of type I dynamics discussed above, where the variability in amplitude was basically absent).

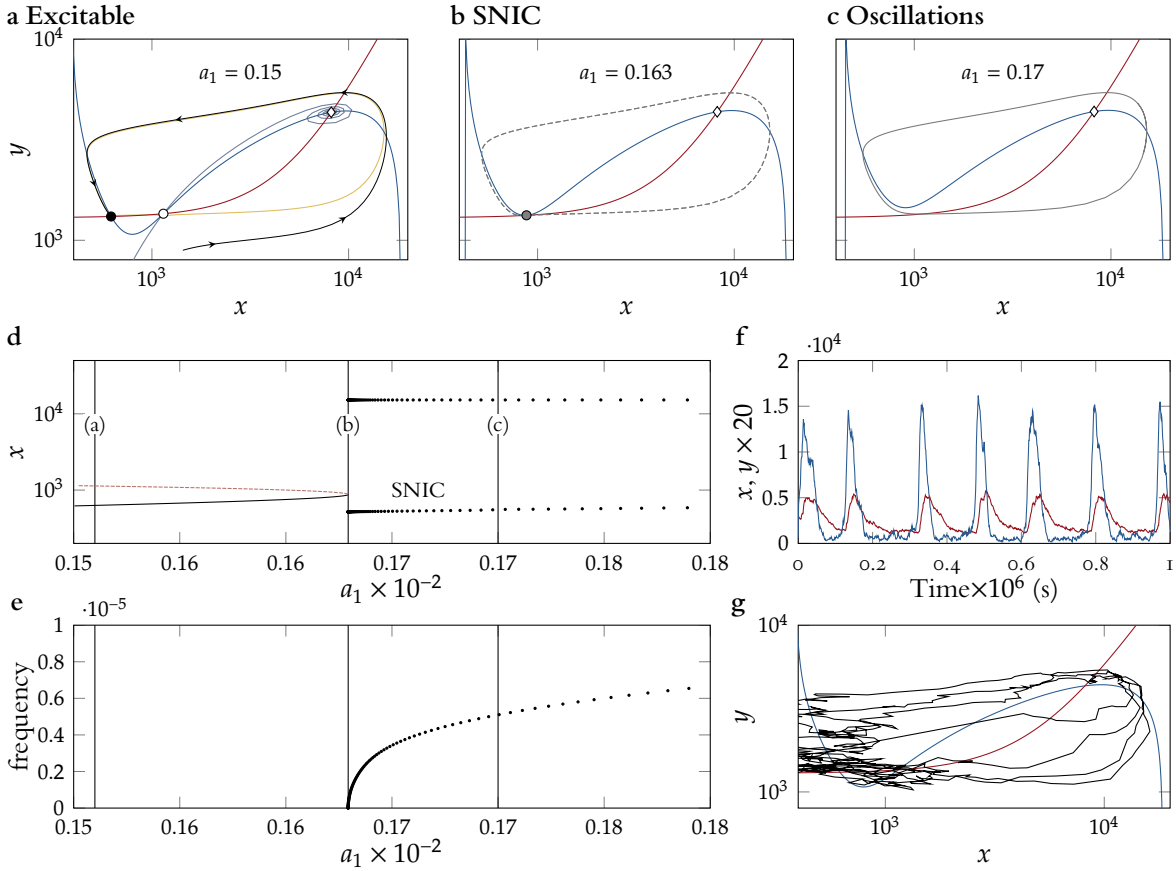

Figure 2.6: **Type I excitability close to a SNIC bifurcation in circuit C.** The meaning of the plots is the same as in Figure 2.2.

Type II excitability close to a subcritical Hopf bifurcation

Finally, we show that the transcriptional version of the activator-repressor circuit can also display type II excitable dynamics. Figure 2.7 shows the behaviour of the circuit for the parameters given in Table A.4. As in the FitzHugh–Nagumo model, we have to assume a clear separation of time scales between the activator and repressor, the former being much faster than the latter. The phase portrait corresponding to the excitable regime, shown in Figure 2.7a, is again analogous to that exhibited by a standard FitzHugh–Nagumo model. No well-defined excitability threshold exists, but perturbations sufficiently far away from the rest state take the system to the excited branch of the activator nullcline (right branch). The bifurcation is similar to the one exhibited by the post-translational activator-repressor circuit B in the type II excitable regime. As the parameter a_1 increases, the rest state loses stability via a subcritical Hopf bifurcation (HB, Figure 2.7b) that generates a stable limit cycle (Figure 2.7c).

The bifurcation plot (Figure 2.7d) reveals that the limit cycle is born

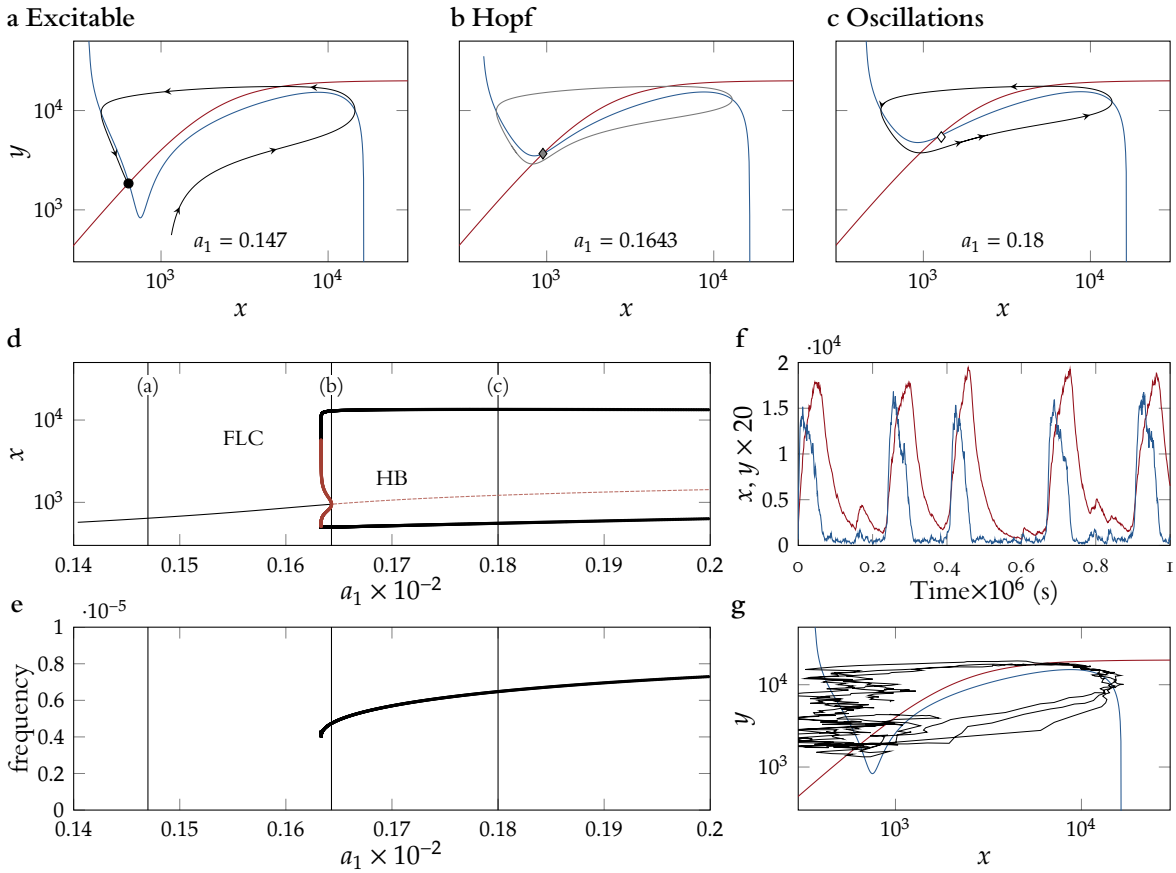

Figure 2.7: **Type II excitability close to a subcritical Hopf bifurcation in circuit C.** The meaning of the plots is the same as in Figure 2.2.

via a FLC that is closely followed by the above-mentioned subcritical Hopf bifurcation, after which the limit cycle is the only stable attractor of the system. Excitable dynamics occurs before the FLC bifurcation. The period behaves in a manner consistent with a type II dynamics, remaining different from zero up to the bifurcation point. Typical time traces and phase-space trajectories are shown in Figure 2.7f and g, respectively.

2.5 *Excitability in parameter space*

We have shown that all three circuits can exhibit both types of excitability regimes depending on specific values of the kinetic parameters. In this Section we show that the parameter values used in each case (Table A.4) are not exceptional and that both types of excitability are expected in these circuits. For each of the three circuits, we partially explore parameter space between the parameters sets used in the previous Sections (see Table A.4). Specifically, if and are the parameter vectors —i.e., they are vectors containing all the model parameters— used for type I and type II excitability, we can map a control

parameter α to the linear combination of these two parameter vectors, α $(\alpha) = (1-\alpha)\cdot$ $+\alpha\cdot$. This new parameter α allows us to track the transition from type I excitability at $\alpha = 0$ to type II excitability at $\alpha = 1$, and will give us a sense of how general these regimes are in each of the circuits studied here. We have restricted the parameter space scan to values of α for which all the parameters were strictly positive. This restriction defines the range of values that α can take to $(-0.25, 1.667)$ for circuit A, $(-0.19, 1.429)$ for circuit B, and $(-27.57, 1)$ for circuit C. Figure 2.8 shows a two-dimensional bifurcation diagram for the parameters α and a_1 for the three circuits. The transition from type I to type II excitability in circuit A is shown in Figure 2.8a. Type I excitability is maintained as we move from the set of type I parameters, $(\alpha = 0$, black circle), towards the set of parameters used type II dynamics, $(\alpha = 1$, white square). In this Figure, the left-most SN branch is associated with type I excitable dynamics (Figure 2.2). As α increases, the complementary SN branch appears and both eventually collide in a cusp point. Just before this codimension-2 bifurcation occurs, a Bodganov–Takens (BT) bifurcation takes place in the lowest SN branch (α 0.865) and gives rise to the Hopf bifurcation leading to type II excitability. Thus, in this circuit excitability is preserved for all values of α, and both types of excitable dynamics are possible in a wide range of parameters. The same is true for circuit B (Figure 2.8b) where the SNIC bifurcation responsible of type I excitability is annihilated in a cusp point at α 0.1748 and the nascent subcritical Hopf bifurcation (notice the BT bifurcation detail in the inset of Figure 2.8b) allows again for type II excitability. Finally, in Circuit C, type I dynamics occupies almost the whole scanned parameter range (see Figure 2.8c), while type II dynamics is relegated to a marginal region. In our opinion, this is a limitation of this parameter scanning procedure rather than an imprint of the exceptionality of the type II regime in this circuit. The separation of time scales required for the excitability in this regime plus the large difference in some of the parameter values between the two regimes (e.g., k_y), makes it difficult to stay in type II excitability while decreasing from $\alpha = 1$ to 0. We have observed that other parameter sets lead to type II dynamics in this circuit (data not shown), but this parameter scanning strategy fails in those cases.

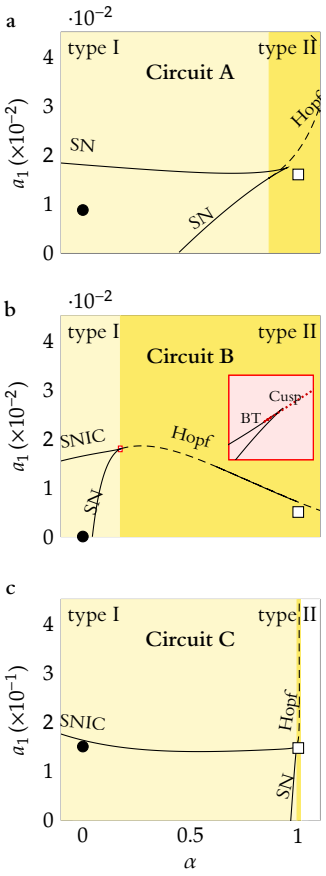

Figure 2.8: **Bifurcation lines in the parameter space formed by α and a_1 for the three circuits.** Faint yellow (dark yellow) areas correspond to the values of α for which the circuits exhibits type I (type II) excitability. Black circles denote the sets of parameters used in this text for type I dynamics, and white squares denote parameters for type II behaviour. The inset in the middle panel (circuit B) shows a zoom-in of the codimension-2 Bogdanov-Takens bifurcation.

2.6 *Conclusions and discussion*

Relating structure and function is an outstanding problem in biology. The design principles underlying such a mapping can be revealed by

means of design space representations (Atkinson et al., 2003). In some cases, multiple circuit architectures can be associated with a single cellular function (Cotterell and Sharpe, 2010; Ma, Trusina, et al., 2009). In this Chapter we have presented a collection of three gene-circuit designs that are able to exhibit excitable dynamics. All three architectures are known to exist in nature, and in some cases excitable dynamics has been observed experimentally. We were interested in determining whether a specific circuit type was necessarily associated with a given type of excitable dynamics, given that two distinct types are known to exist with different dynamical and statistical properties of the generated pulses. Our study shows that this is not the case, but that rather all types of circuits can generate both types of excitable behaviour for a wide range of parameter values (at least in circuits A and B). These results agree with other similar works (Conrad et al., 2008; Munteanu et al., 2010).

This type of design degeneracy can be expected to be very common in natural systems. Theoretical studies have found this to be the case in noise-buffering circuits (Hornung and Barkai, 2008) and pattern-forming circuits in development, for functions such as segmentation (Ma, Lai, et al., 2006) and single-stripe formation (Cotterell and Sharpe, 2010). Recent studies (Çagatay et al., 2009; Kittisopikul and Süel, 2010) have shown that such a dynamical degeneracy can be lifted by noise, provided the different circuit architectures involved are associated with different phase relations between the circuit components: if the circuit topology is such that the proteins are anti-correlated, one of them will be always occurring in low numbers and will thus be subject to large fluctuations, so that the overall amount of noise in the system will be high. On the other hand, if the topology is such that the proteins are correlated, the levels of both proteins will be high or low simultaneously, and therefore the amount of noise will fluctuate in time inversely with the protein levels. In that way, noise propagates in different ways through the different circuit designs, and provides the system with specific functional properties that most likely act as fitness determinants in the evolution of a given molecular circuitry to perform a desired cellular function (Çagatay et al., 2009; Kittisopikul and Süel, 2010).

On the other hand, the fact that a given circuit can have different dynamical behaviours would allow evolutionary pressures to shape the specific dynamical response of the cell —for instance in terms of their noise level— without having to modify the circuit's topology. Hopefully the systematic dynamical study presented here could help in establishing

how such these selection processes take place in other excitable cellular functions.

3

Noise-induced polymodal genetic oscillators

Using the naturally occurring polymodal oscillations described in the Introduction (Part I) as motivation, here we address the general question of how a limit cycle behaviour with polymodal period distribution can arise in minimal oscillator models. To that end we consider one of the most basic oscillator architectures, namely a two-component activator-inhibitor system operating in an excitable regime (close to the oscillatory region) and subject to noise, as already described in the previous chapter. Turcotte, Garcia-Ojalvo, and Süel (2008) recently showed that, for certain parameter values, the circuit B described in Section 2.3 above in the type I regime, exhibits noise-induced stabilisation of an unstable spiral state.

Due to its excitable nature, this model system displays noise-triggered excursions away from the stable (rest) state, during which the cell passes through a region near the unstable spiral (see Figure 2.4). The stabilisation mechanism consists in the appearance of oscillations around the unstable state, due to the stochastic fluctuations (see Figure 3.1a). These noise-induced oscillations around the spiral occur when the system parameters are such that the stable and unstable manifolds of the saddle point are close enough. In this setting, noise can drive the system, which otherwise would follow excitable orbits close to the unstable manifold (yellow line), across the stable manifold (grey line) and force it to turn around the spiral point. The proximity of the invariant manifolds of the saddle in this case can be explained by the fact that the system is close to a global bifurcation to a stable limit cycle that appears as a homoclinic orbit at the saddle point (SH bifurcation, see Figure 3.1b).

As a consequence of these noise-induced oscillations, the distribution of excursion times away from the rest state exhibits a marked polymodality: each noise-induced oscillation around the unstable state introduces

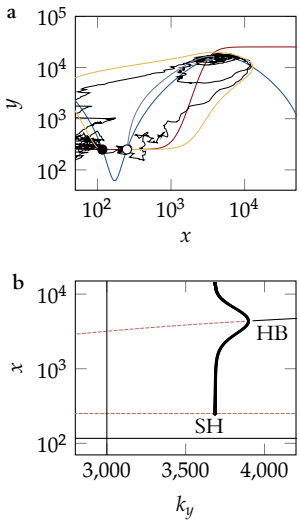

Figure 3.1: **Multipulse dynamics in an excitable circuit.** The activator-repressor Circuit B from Figure 2.1 displays multipulsed dynamics. (a) Noise-triggered excursions out of the stable state get temporarily trapped around the unstable focus, generating an integer number of consecutive cycles around it. Grey (yellow) line indicates the stable (unstable) manifold of the saddle point. Adapted from Turcotte, Garcia-Ojalvo, and Süel (2008). (b) Multipulse dynamics in the excitable regime appear beneath a bifurcation to a short limit cycle (SNIC) that rapidly ceases at a Hopf-stabilising point (HB).

a well defined delay (the oscillation period) in the pulse duration. Thus noise can explain the polymodality of pulse duration distributions in certain conditions. However, the excitable pulses are triggered by noise to begin with, and thus they are far from occurring periodically, which would be necessary if this mechanism is to explain the polymodal cell-cycle duration distributions mentioned above.

Coincidentally, however, systems with excitable dynamics are known to exhibit enhanced periodicity, or coherence, for an optimal amount of noise: too little noise will elicit pulses only sparsely, and therefore irregularly, while too high noise will lead to a strong disorder in the dynamics. A moderate level of noise, on the other hand, is able to evoke pulses frequently, as soon as the refractory time following the previous pulse —characteristic of all excitable systems— has elapsed, and thus leads to an approximately periodic behaviour, with a period basically given by the refractory time. Such somewhat counter-intuitive effect of noise has been termed coherence resonance (CR) or, more appropriately, stochastic coherence (Gang et al., 1993; Lindner et al., 2004; A. S. Pikovsky and Kurths, 1997). The goal of this Chapter is to show that stochastic coherence can be invoked, together with the noise-induced stabilisation effect discussed above, to provide a minimal mechanism for the generation of polymodal distributions of cycle lengths in an otherwise periodic behaviour. The mechanism requires only a simple genetic activator-repressor motif and an optimal amount of random fluctuations. In our setting, the effect of intrinsic molecular noise is again characterised by using discrete stochastic simulations.

The level of intrinsic noise is controlled by the cell volume, whose increase (together with the gene copy numbers) effectively scales up the numbers of molecular species —thus reducing the noise—, while maintaining the concentrations constant (Kampen, 2007). This approach was recently introduced experimentally in *B. subtilis* (Suel et al., 2007), and has been subsequently used in *E. coli* as well (Fischer-Friedrich et al., 2010). From the theoretical side, system-size effects have been seen to lead to stochastic resonance (SR) (A. Pikovsky, Zaikin, and Casa, 2002; Schmid, Goychuk, and Hänggi, 2001) and CR (Toral, Mirasso, and Gunton, 2003) through their control of the effective noise intensity perceived by the system.

Our results show that noise, besides enhancing the regularity of the pulse activations, also optimises polymodality in the system's response. Furthermore, optimisation of periodicity and polymodality are achieved

when noise levels are comparable. Thus, when the coherence of the excitable pulses is maximised, so is the probability that the pulses undergo oscillations around the unstable spiral state. There is a range of noise levels for which optimisation holds. Together, these results show that noisy activator-repressor genetic circuits can naturally behave as polymodal oscillators.

We focus on the circuit C described in the previous chapter (Section 2.4) working in a type I regime, although the results obtained here are applicable to almost any activator-inhibitor system. The genetic circuit (see Circuit C in Figure 2.1) is a simple two-gene system where an activator protein, X, binds to and activates its own promoter, P_X, and the promoter of a repressor species Y, P . The repressor component, Y, in turn, inhibits the expression of the activator species by competitively binding to the promoter P_X. As extensively described in the previous chapter, this system is a prototypical transcriptional activator-repressor genetic circuit where the activator species forms a direct positive feedback loop through its auto-regulation and an indirect negative feedback loop by means of the activation of its own repressor (Stricker et al., 2008; Tigges et al., 2009; Tsai et al., 2008).

The dynamics of this system is highly non-linear, due to the cooperative nature of the regulated transcription processes. All modelling details can be found in Appendix A.1 and Chapter 2. The values of the reaction rates used in the simulations shown below are given in Table A.5. These values are within reasonable biological ranges for a gene regulation circuit. In particular, the values chosen for the transcription, translation, and degradation rates, and for the activation and inhibition threshold concentrations, are of the same order of magnitude of previous studies that involved qualitative comparison and careful validation with experimental measurements (Çagatay et al., 2009; Süel et al., 2006; Suel et al., 2007). The rates of the transcription reactions are proportional to a factor Ω. This parameter is a global scaling factor that depends on the size of the cell. Specifically, $\Omega =$ $\cdot N_A$, where is the cell volume and N_A is Avogadro's number (see Appendix A for a detailed explanation, and Kampen, 2007). In that way, Ω relates the species concentrations with the number of molecules: if X is the number of molecules of X and x its concentration, then $X = \Omega x$. We will also consider that the transcription strengths are proportional to Ω, which is a generalisation to the continuous of the assumption that the gene copy number increases with cell size.

This happens when a cell is prevented from dividing but not from replicating its DNA, as happens in certain bacterial mutants (Suel et al., 2007). Under those assumptions, the level of molecular noise increases monotonically with the inverse of the system size, Ω^{-1} (Gillespie, 1977). In the results presented below, we will use the parameter Ω^{-1} to characterise the levels of noise in the system.

In this Chapter, as in the previous one, the model dynamics is obtained by exact discrete stochastic simulation of the chemical kinetics (ibid.) described in Table A.3. We also make use of the continuous approximation of this system, already detailed in Equation (2.3), to analyse the underlying deterministic dynamics. This deterministic model does not depend on the system size, and thus the deterministic dynamics will be unchanged as noise levels vary. This is due to the fact that the system size factor Ω rescales the levels of all proteins —and thus the absolute activation and inhibition thresholds— in the same manner as the transcription rates, while keeping the concentrations unchanged. In that way, the average concentration dynamics of the model does not depend on noise, but only the variances of the concentrations do.

The rest of the Chapter is organised as follows: in Section 3.1 we review the dynamics of the system working in the type I regime, the robustness of this regime against parameter changes and the effect of molecular noise in the generation of excitable pulses. In Section 3.2 we analyse the effect of noise in the generation of multi-pulses via the noise-induced stabilisation effect. We next address, in Section 3.3, a second noise-induced effect, namely coherence resonance (CR), and analyse its coexistence with the polymodal pulses (Section 3.4). Finally, in Section 3.5 we show that both noise-induced effects can robustly coexist for a broad range of noise amplitudes.

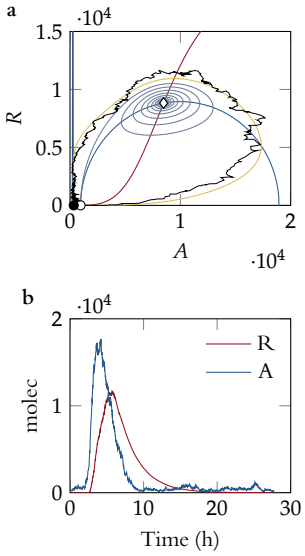

Figure 3.2: **Simulation results of the minimal genetic circuit** (a) Phase portrait representing the dynamics of the system of equations Equation (2.3). The nullclines for X and Y are represented by blue and red lines, respectively. The system has three equilibrium points: a stable node (black circle), a saddle point (white circle) and an unstable focus (white diamond). The stable manifold of the saddle point (grey line) introduces a threshold of excitability. Near the spiral, the unstable manifold of the saddle point (yellow line) approaches the stable manifold, as in the case shown in Figure 3.1. A trajectory obtained by discrete stochastic simulations (Ω = 1 molec nM^{-1}) is shown in black. The random fluctuations of this system make it possible to cross the separatrix and initiate a large excursion in the phase plane. (b) The same trajectory is plotted as a time course of the protein species. When the threshold of excitability is crossed, a transient pulse in the number of activator and repressor molecules is produced.

3.1 Deterministic excitable dynamics and effects of molecular noise

The interplay between the auto-activating positive feedback loop of X on itself and the negative feedback loop formed by the activator and the repressor allows for a wide range of rich dynamics. In particular, for the set of parameters given in Table A.5 and Equation (A.4a) the system is, once more, excitable. The phase portrait depicting the nullclines of this system for those parameter values is shown in Figure 3.2a.

The system is topologically equivalent to that in Figure 2.7, with three equilibrium points: a stable node (black circle in Figure 3.2a), a saddle point (white circle) and an unstable focus (white diamond), and therefore

obeys type I excitability. In the absence of noise, the system rests in the only stable state, which in this case corresponds to low numbers of both activator and repressor molecules. Small perturbations from this stable point vanish exponentially and the system quickly recovers the rest state. However, the stable manifold of the saddle point (the separatrix, grey spiral line in Figure 3.2a) acts as an excitability threshold beyond which perturbations evoke a large excursion through phase space, passing around the unstable focus and back to the stable point avoiding to cross the separatrix (Izhikevich, 2006). The occurrence of this excitability cycle can be understood as follows: a sufficiently large initial amount of activator molecules triggers the positive feedback loop and leads to a large pulse of activator molecules. The increasing levels of activator switch on the production of repressor proteins (repressor pulse) which, in turn, shut down the production of activator. Finally, the amounts of activator and repressor decay due to linear degradation/dilution. This transient response is characterised by a refractory time, which is the duration of the cycle from the triggering event to the recovery of the rest state.

Stochastic fluctuations due to intrinsic noise can destabilise the rest state by randomly crossing the excitability threshold and hence generating pulses of activator and repressor protein levels. Figure 3.2a shows an excursion in phase space triggered by stochastic fluctuations, and Figure 3.2b shows the corresponding time course for both the activator (blue line) and the repressor (red line).

The excitable regime in which this system operates arises when the system is close to a bifurcation point beyond which the dynamics has the form of a limit cycle —in this case a SNIC bifurcation analogous to the one analysed in Section 2.4. In addition, the noise-induced stabilisation of the unstable state emerges close to a Hopf bifurcation (Turcotte, Garcia-Ojalvo, and Süel, 2008) below which the system becomes bistable (see Figure 3.3a and compare with Figure 3.1a). The bifurcation diagram corresponding to the parameter controlling the threshold of repressor activation (k_y in Equation (2.3)), which is depicted in Figure 3.3a, makes this fact evident: the system is working in an excitable regime close to a fold limit cycle (FLC) bifurcation —whereby a stable and an unstable concentric limit cycles around the unstable focus are generated— and shortly after the focus point is Hopf-stabilised (HB). Thus, the system operates close to the SN(IC) bifurcation studied in Section 2.4, that allows noise to easily push the system across the separatrix and generate regulated pulses; and also close to a Hopf-stabilising bifurcation that

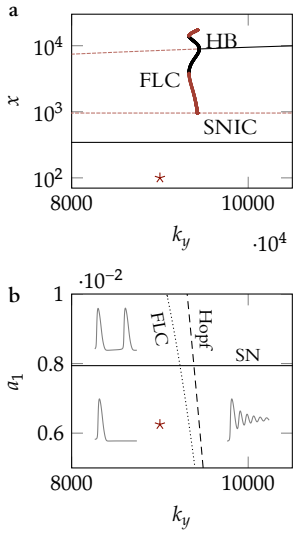

Figure 3.3: **The system operates close to bifurcations to a limit cycle.** (a) Bifurcation diagram for the threshold of repressor activation, k_y in Equation (2.3). This diagram reveals a bifurcation to a double limit cycle (FLC), the stable limit cycle (thick black line) fades in a Hopf (HB) and the unstable cycle (thick red line) collides with the saddle (SNIC). The red star symbol indicates the parameter value used in the excitable regime. (b) The 2d diagram of K_Y and a_1 shows that the system is close to another SNIC bifurcation, apart from the FLC in (a). Both bifurcations lead to different limit cycle oscillations but the FLC is immediately followed by a Hopf-stabilising bifurcation. The red star symbol denotes the operating parameters. For each regime a characteristic time-trace is depicted.

increases the chances of the system spiralling back and generating multiple pulses. Contrary to the case presented by Turcotte, Garcia-Ojalvo, and Süel (2008), where the Hopf bifurcation was preceded by a SH bifurcation, here the limit cycle is achieved by the FLC mentioned above. Despite this distinction in the bifurcation structure, the effect is the same: the stable and unstable manifolds of the saddle point are close in the surroundings of the spiral point, as depicted in Figure 3.2a. Figure 3.3 shows a 2-parameter bifurcation diagram illustrating this fact with example time traces of the different working regimes.

Despite these constraints in the parameter values, the activator-repressor system presented here is robust to parameter changes, as shown in Figure 3.4. This Figure depicts the main bifurcations from excitability that the system undergoes as the kinetic parameters are varied one by one. In particular, the codimension-one bifurcations found are

i the stabilisation of the unstable spiral via a Hopf bifurcation;

ii the collision of the saddle and the node defining the resting state (saddle-node 1 in the Figure 3.4); and

iii the collision of the saddle and the unstable state (saddle-node 2).

This analysis shows that the more sensitive parameters are the activator degradation rate, d_1, for which the excitability is maintained for a global range of variation of 27.8 % and the maximum regulated transcription rate, b_1, for which a global variation of 36.7 % is possible without losing excitability.

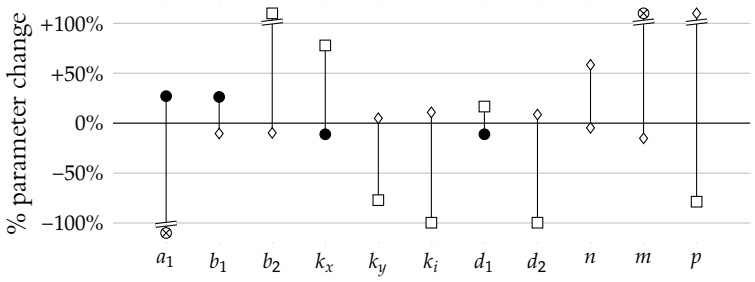

Figure 3.4: **Sensitivity of the model to single parameter changes.** Vertical lines show the range of parameter values for which excitability is maintained (in % change from the values in Table A.5). The symbols at the end of the vertical lines indicate the type of bifurcation leading to loss of excitability: Hopf (diamond), saddle-node 1 (white square), saddle-node 2 (black circle); or whether the range is unbounded (crossed circle). Broken bars indicate much larger ranges than the one indicated by the vertical axis.

3.2 *Polymodality in the cycle duration depends on the level of intrinsic noise*

As already mentioned, the sporadic generation of pulses of activity is not the only effect produced by molecular noise in this system. As shown in Figure 3.5a and Figure 3.5b, noise also stabilises the unstable state and generates bursts, or cycles of multiple pulses. This is due to the fact that noise causes the trajectories travelling around the unstable spiral point —which is close to a FLC and a Hopf stabilising bifurcation— to cross over the stable manifold of the saddle (grey spiralling line in Figure 3.5a). This leads to the trajectory getting trapped orbiting around the unstable fixed point for an integer number of cycles, thus generating a polymodal distribution of pulse durations (Turcotte, Garcia-Ojalvo, and Süel, 2008).

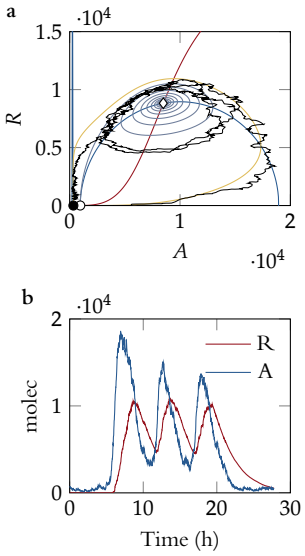

Figure 3.5: **Molecular noise induces bursts with multiple pulses.** (a) A burst trajectory obtained with discrete stochastic simulations ($\Omega = 1\,\text{molec}\,\text{nM}^{-1}$) is depicted in the phase plane. Fluctuations due to molecular noise synergistically interact with the dynamics around the focus and can temporarily trap the system in an area around the otherwise unstable state. (b) Time course for the number of activator and repressor proteins for the burst shown in (a). The stabilisation of the active state is characterised by the oscillations in the numbers of molecules.

In order to characterise this noise-induced polymodality, we have computed the cycle durations for varying levels of molecular noise (which increase as the system size Ω decreases). Figure 3.6a shows, in colour code, the normalised histogram of burst durations for different values of Ω^{-1}. For large system sizes (small Ω^{-1}, small molecular fluctuations), the histogram shows multiple modes with defined peaks. Each of the modes of the histogram obviously corresponds to a class of burst with a definite number of pulses. The single pulse cycle (corresponding to the first mode in the histogram) is, by far, the most probable case. For intermediate noise levels, the secondary peaks, corresponding to cycles with multiple pulses, get both higher and wider. If noise is increased further, the polymodal character of cycle durations is lost and just a single wide peak remains.

These results insinuate a new resonance-like effect in which growing levels of noise increase the polymodality in the pulse-duration distribution, until it reaches a maximum and starts declining again. In order to assess this resonant effect, we have computed the probability of having more than one pulse in one burst (Figure 3.5b), together with the number of pulses per activation burst (Figure 3.5c) as a function of increasing noise levels. Here, the pulses per cycle are computed by counting the number of complete cycles around the unstable focus. This method was found to be a robust way to compute the number of pulses in the presence of large random fluctuations. Both plots show the predicted resonance effect, with the optimal degree of polymodality arising at a value of Ω 1 molec nM^{-1}.

The reasoning behind this resonant effect can be stated as follows. For large system sizes (small noise), noise is large enough to trigger excitable pulses, but it is too small to easily induce crossings of the trajectory beyond the stable manifold of the saddle. Thus the fraction of bursts in which there is more than one cycle is small —and the average number of cycles is close to one. On the opposite side, for small system sizes, the random fluctuations are too large to maintain the coherence of the oscillations around the unstable spiral, and the burst duration is no longer quantised, but is widely variable and with small mean, since it is easy for the trajectory to escape the area near the unstable spiral and relax back to the neighbourhood of the rest state. For intermediate system sizes, on the other hand, noise is large enough to induce frequent crossings of the spiral's stable manifold, but small enough to maintain the coherence of the noise-induced oscillations, and thus polymodality is maximised.

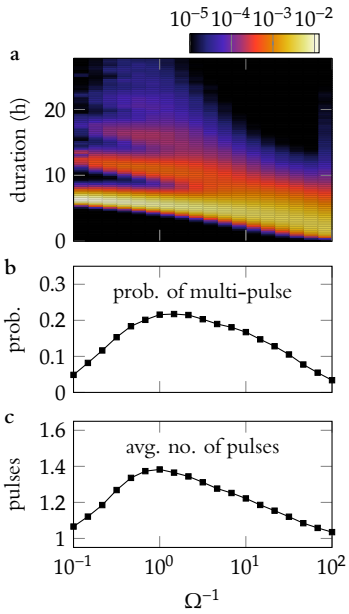

Figure 3.6: **Molecular noise induces polymodal pulsing in the gene circuit.** (a) Histograms of the burst durations for varying levels of the intrinsic noise (Ω^{-1}). Colour is coded in logarithmic scale to emphasise the existence of polymodality. (b) Probability of generating a cycle with multiple pulses plotted against the noise level. For intermediate values of the noise strength the system reaches a maximum probability of generating bursts with more than one pulse. (c) Average number of pulses per burst of activity as a function of Ω^{-1}.

3.3 Noise modulates the regularity/coherence of the oscillations

Let us now turn our attention to the ability of the system to generate regular cycles. Not being a genuine genetic oscillator but an excitable system, pulses in this circuit are in principle randomly generated by noise. In this scenario, we want to establish whether the level of noise has an impact in the regularity of pulse initiation. This effect is already made evident by visual inspection of the time traces of the activator species for different values of Ω, as shown in Figure 3.7a. The three panels in this Figure display time traces of the activator species for three different levels of molecular noise (increasing noise from top to bottom). The dynamics of the system in these three panels are qualitatively different. While the top and bottom panels show bursts of activity at irregular time intervals, the middle panel presents a regular pattern of cycles.

This plot already hints at a second noise-dependent resonance-like effect, according to which the regularity of excitable pulses is maximised for an intermediate amount of noise, what is known as coherence resonance or stochastic resonance (Lindner et al., 2004; A. S. Pikovsky and Kurths, 1997). In order to quantify this effect, we compute the Coefficient of Variation (CV) of the time between bursts of protein concentration, a reliable regularity measure. The CV is defined as the standard deviation normalised to the mean

$$c_v = \frac{\sqrt{\langle\ ^2\rangle - \langle\ \rangle^2}}{\langle\ \rangle},$$

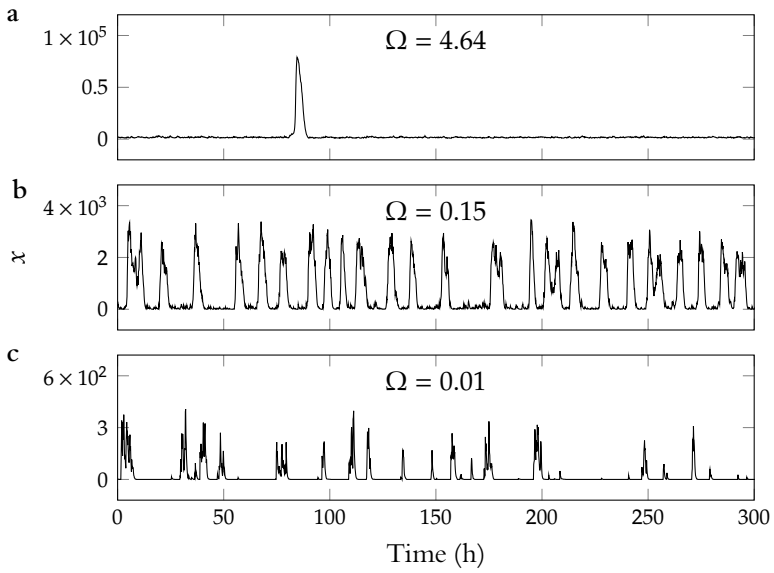

Figure 3.7: **Temporal evolution of number of activator molecules for three different values of the system size.** Note how the peak number of molecules decreases as the system size decreases.

where the random variable denotes the duration between bursts of activity. This measure is routinely used as a quantifier of noise-induced regularity in CR studies (Lindner et al., 2004). In a perfectly periodic regime where the pulses were equally spaced in time, the coefficient of variation would be exactly zero —as the variance of the distribution of periods would be zero. Conversely, in a totally random regime with pulses following poissonian statistics —with exponential waiting times—, the coefficient would take the value of one —the standard deviation of an exponential distribution being equal to its mean— or even higher — hyper-exponential waiting times. Thus, the smaller the value of the CV, the closer the system is to operate in a coherent regime. Other quantifiers of CR, such as the correlation time of the dynamics, can be used, leading to the same conclusions (A. S. Pikovsky and Kurths, 1997). Here, the CV is estimated from the simulated time courses using a robust thresholding method.

The coherence resonance is revealed in Figure 3.8a, where the CV is plotted against the system size. For small amounts of intrinsic noise the system mainly remains in the basal stable state, with some sporadic pulses appearing randomly in time (see also Figure 3.8b, which shows that the pulsing rate approaches 0 for small noise). This results in a CV of around 1. As the system size is decreased, the effect of intrinsic noise increases and the system pulses at a higher pace. Here is where the refractory time enters the game, as it poses a limit in the pulsation frequency —the system cannot undergo a new cycle if it has already started one. Thus,

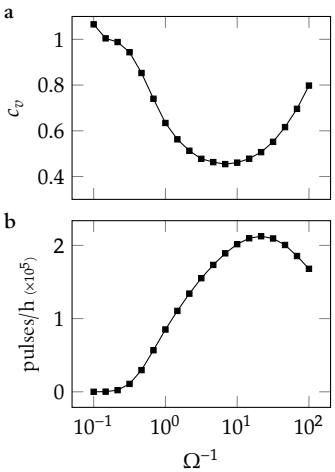

Figure 3.8: **Stochastic coherence effect in the minimal gene circuit.** (a) Coefficient of variation of the intervals between consecutive cycle initiations, plotted versus Ω^{-1}. (b) The average rate of cycle initiation as a function of the noise level. For large system sizes, the fluctuations are not large enough to initiate the cycles. For small system sizes the decay in the pulsing rate is mainly due to the discrete nature of the chemical reactions and the small amount of mRNA species.

temporal regularity appears as a synergistic effect involving the dynamics of the system and the intrinsic noise. In particular, the maximum regularity in the oscillations (a minimum CV) appears at $\Omega = 0.147$ molec nM^{-1}. Further reducing the system size causes a reduction in the regularity of the oscillations. This loss of coherence is due to two main causes: first, large amounts of noise destroy the excursion paths, thus generating a variety of incomplete pulses and eliminating the system's eigenfrequency dictated by the refractory time (Lindner et al., 2004). The second cause of coherence loss, which is not common in standard coherence resonance, is the appearance of periods of silencing where the activity of the system is shut down (see bottom panel in Figure 3.7a). These periods of silencing are due to the fact that when the cell size is small, the number of molecular species is small. In particular, the number of mRNA molecules falls frequently to zero, resulting in the total absence of protein expression during large time intervals. Such low levels of mRNA are not uncommon in cells, as has been recently shown experimentally in *E. coli* (Taniguchi et al., 2010).

3.4 Polymodality affects regularity

We have shown that this circuit displays two apparently opposing effects caused by intrinsic noise. On one side, noise increases the variability in the duration of the cycles in a quantised manner. On the other side, it reduces the variability in the cycle initiation times. Thus, it is reasonable to think that these effects might affect one another. Here we show how polymodality in the duration of the cycle poses a limit in the temporal coherence attained by the genetic oscillator. For this purpose, we first assume that for a given range of noise we reach a perfect timing of the cycles, where each cycle follows the next without delay. In this hypothetical case we can estimate the loss of temporal coherence (in the CV sense) attributable to the cycle duration polymodality by taking into account the probabilities of obtaining a cycle with a particular number of bursts. In this case, in which a new cycle starts just after the previous one, the average time between cycles is

$$\langle \quad _{\mathrm{pm}} \rangle = \sum_{i \ 1} p_i \quad _i ,$$

where p_i is the probability of getting a burst with exactly i pulses ($i > 0$) and $\quad _i$ is the average duration of a burst with i pulses. We can also estimate the variance by further assuming that all the cycles with i pulses

have a length of exactly $_i$ (zero variance among them):

$$_{\rm pm}{}^2 = \langle \ _{\rm pm}{}^2\rangle - \langle \ _{\rm pm}\rangle^2 = \ _i \ _1 \ p_i \ _i{}^2 \ - \langle \ _{\rm pm}\rangle^2.$$

This allows us to calculate the contribution of the polymodality, $c_v^{\rm pm}$, to the total coefficient of variation of the time between initiations.

Figure 3.9 compares $c_v^{\rm pm}$ (white circles) with the total CV (black squares, see also Figure 3.8a) for varying noise levels.

In this Figure, $c_v^{\rm pm}$ has been computed using values for the probabilities p_i that were estimated from the time traces of the stochastic simulations (see Figure 3.6b,c and accompanying text). In addition, the values for $_i$ have been fitted to the formula $_i/\ _1 = 1 + (i-1)$ (resulting in $= 0.75$). A comparison between the white circles in Figure 3.9 and the result of Figure 3.6c shows that, for continuous pulsing —noise generates a pulse as soon as the refractory time from the previous pulse is over—, the regularity of the pulses drops ($c_v^{\rm pm}$ increases), as the polymodality is enhanced. Thus, in the regime of continuous pulsing there is an evident trade-off between cycle length polymodality and temporal coherence. Finally, Figure 3.9 also shows that around the minimum CV the system is close to the regime of constant cycling. Thus in that case basically all the remaining irregularity —note that the CV does not decay to zero— is due to the polymodality. Therefore, polymodal behaviour establishes an upper bound for the regularity of the system's dynamics.

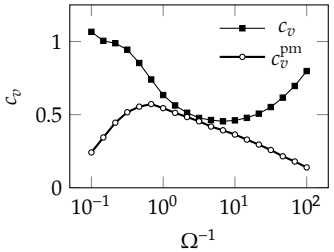

Figure 3.9: **The gene circuit approximately achieves optimal coherence** Coefficient of variation of the inter-burst times due to polymodality, $c_v^{\rm pm}$ (white circles), compared to the total CV, c_v, of the dynamics (black squares, from Figure 3.8a). For intermediate values of noise the CV of the oscillator becomes close to $c_v^{\rm pm}$.

3.5 *Noise induced polymodality and regularity robustly coexist*

An important question is how the noise levels that optimise coherence and polymodality compare to one another. A comparison between Figure 3.6b and Figure 3.8a reveals that, for the parameter values chosen, there is one order of magnitude difference between the two optimal noise levels. However, we can still say that the two effects displayed by this simple genetic model take place for overlapping ranges of molecular noise. This can be argued from Figure 3.10, which plots together the measures of polymodality (in the x-axis) and regularity (in the y-axis) for varying amounts of noise (milestones labelled in some data points in the figure).

The Figure shows that as noise increases, both the regularity and polymodality increase, and there is a range of noise levels spanning over an order of magnitude (black squares in the figure), for which both magnitudes are equally high, before decreasing again as noise is further

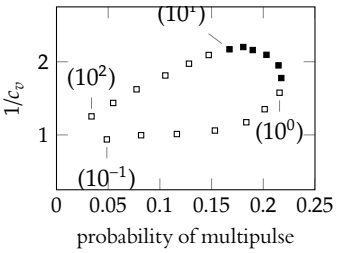

Figure 3.10: **A range of noise levels optimise regularity and polymodality almost simultaneously.** Regularity ($1/c_v$, from Figure 3.8) is plotted against polymodality (probability of multiple pulses in a burst, from Figure 3.6) for varying levels of molecular noise. Labels indicate the inverse of the system sizes (Ω^{-1}) for some of the points. Black squares correspond to noise levels for which the system is both polymodal (probability $\ 0.15$) and regular ($c_v < 0.6$).

increased. Thus, one can say that an optimal level of noise maximises almost simultaneously, and for a wide range of noise amplitudes, both the regularity and the polymodality of the dynamics of the activator-repressor module.

3.6 Conclusions and discussion

Cellular processes regulated by genetic components are subject to large amounts of random fluctuations. In the face of this fact, it is appealing to conjecture that, rather than simply trying to filter out noise, certain cellular mechanisms have evolved to cope with random fluctuations, and in some cases even rely on them for function. In the last decades many noise-induced phenomena in physical systems have been described, both theoretically and experimentally. Strikingly, noise can in some cases increase order in the dynamics (Sagués, Sancho, and Garcia-Ojalvo, 2007) and play a constructive role in non-linear systems. It is also becoming evident in recent years that molecular noise has an impact on the dynamics underlying many biological processes (Alonso, McKane, and Pascual, 2007; Çagatay et al., 2009; Eldar and Elowitz, 2010; Kittisopikul and Süel, 2010; McKane and Newman, 2005).

Here we have shown that intrinsic noise is able to turn a simple activator-repressor genetic circuit into an oscillator with non-trivial statistical properties, reflected in a polymodal distribution of cycle durations embedded in a relatively strongly periodic sequence. A similar effect has been reported in coupled excitable elements (Koseska et al., 2007; Volkov et al., 2003). Here, in contrast, we show that the phenomenon can arise in single excitable elements. The role of noise in our system is two-fold. On the one hand, it stabilises an unstable spiral point by inducing oscillations around it, which increases the duration of phase-space excursions in a quantised manner. Furthermore, the resulting polymodal character of the dynamics is enhanced for an intermediate noise level. The second role of noise is to enhance coherence in the pulse initiation times, which occurs via a standard coherence resonance effect, characteristic of excitable systems subject to noise (Lindner et al., 2004). This double optimisation provides a relatively simple mechanism for the emergence of polymodal behaviour in genetic oscillators.

4

Heterogeneity and correlations in the pluripotency network of Embryonic Stem cells

Transient cellular programs pervade most processes throughout the development of multicellular organisms. During mammalian development, for instance, individual cells undergo multiple and fast transitory fates that eventually lead to specific adult cellular types. In one of these stages, the blastocyst stage, some cells take a particular fate characterised by their self-replication ability and their pluripotency (Section 1.1). Despite the self-renewal condition, these cells (called Embryonic Stem cells) can only maintain this fate for a short period during development.

This Chapter is devoted to the analysis of the molecular mechanisms underlying pluripotency from two complementary points of view: on one side we study the transcriptional interactions that regulate the large fluctuations observed in the levels of the core pluripotency factor Nanog (Section 4.1) and, on the other side, we propose a novel mechanism of pluripotency regulation based on post-translational interactions that explains the observed correlations between four key factors in pluripotency maintenance (Sections 4.2 and 4.3).

4.1 Transcriptional regulation of pluripotency: a model with three states

Over recent years, some models have been proposed to describe the dynamics underlying pluripotency regulation in ES cells. Of particular interest to us is the mechanism proposed by Kalmar et al. (2009). In their model, Kalmar and coworkers proposed a core transcriptional circuit that describes the experimentally observed heterogeneities in terms of regulated excitable excursions from the high Nanog expression state (HN) to lower expression levels (LN state). Similarly to the genetic circuits studied in Chapter 2 of this Thesis, Nanog excursions are triggered by molec-

ular noise. At the time that model was presented, only a small number of transcriptional interactions were known and some assumptions were made in the model. These interactions (summarised in Figure 1.7) comprise mutually and self-regulated activation of Nanog and Oct4, and also a negative feedback on Nanog expression activated by high levels of Oct4 via an unknown factor X.

As knowledge on the transcriptional details of the pluripotency network has been broadened, we decided to update that model. On a first approach, we identified the factor X with the transcriptional repressor Tcf3. Experimental data supports that Tcf3 is activated by high levels of Oct4 and directly represses activity of Nanog promoter (Pereira, Yi, and Merrill, 2006). Although the role of Tcf3 was already predicted by the original model from Kalmar et al. (2009), its interaction form with Nanog was assumed to be enzymatic —via direct chemical interaction at the protein level— rather than transcriptional. In addition, the original model neglected the activation of Nanog mediated by Oct4 and assumed non-saturable activity of the Oct4 promoters.

Another important observation refers to the structure of the ES cell population. As commented in the Introduction, in addition to the High (HN) and Low (LN) Nanog populations, an ES cell culture also contains a population with intermediate levels of Nanog (MN, for a dynamical/statistical treatment of these three states see Luo et al., 2013). This population which seems to be unstable, includes cells that can either up or down-regulate their levels of Nanog and could represent a transition state between HN and LN.

In view of these observations we updated the original model to incorporate them and see the effect that they would have in the dynamics of the network. In its simplest form, the resulting gene circuit is represented by the following coupled differential equations, which describe the temporal evolution of Nanog (N) and Oct4 (A):

$$\frac{dN}{dt} = a_n + \frac{b_{nn}(N/K_{nn})^{p_n} + b_{na}(A/K_{na})^{p_a}}{1 + (N/K_{nn})^{p_n} + (A/K_{na})^{p_a} + (A/K_{nb})^{p_b}} - d_n N, \qquad (4.1a)$$

$$\frac{dA}{dt} = a_a + \frac{b_{aa}(A/K_{nn})^{q_a} + b_{an}(N/K_{an})^{q_n}}{1 + (A/K_{aa})^{q_a} + (N/K_{an})^{q_n}} - d_a A. \qquad (4.1b)$$

Here, the Nanog input function takes a complex form, integrating three signals:

i Nanog self-activation, with cooperativity p_n and binding constant K_{nn};

ii direct Oct4-activation, with cooperativity p_a and binding constant K_{na}; and

iii indirect Oct4-regulated repression via Tcf3, with an effective binding constant K_{nb} and cooperativity p_b.

The Oct4 input function is also more complex than in the original model, as it also includes self-activation (K_{aa}, q_a) as well as saturable Nanog-mediated activation (K_{an}, q_n). Specific details on the model derivation and the parameter values used in the simulations below can be found in Appendix A.2.

As a planar system, a qualitative sense of its dynamics can be accounted for by means of the phase plane analysis, shown in Figure 4.1. In this Figure, blue and red lines correspond to N and A nullclines, respectively. These cross thrice, generating three equilibrium points, two of which are stable (black circles) and the third is of saddle type (white circle). The two stable points —identified as LN (left most) and HN (right most)— generate two wells of attraction that are fenced off by the separatrix originated in the saddle (yellow line).

Although this is, in principle, a standard bistable system, it is interesting to note that the S-shape of the upper part of the N nullcline induces excitability on the HN state —its topology is locally equivalent to a type II FitzHugh–Nagumo, but with inverted axes. Therefore, any point far enough of the HN state will undergo a large excursion to low levels of Nanog and Oct4 prior to relaxing back to the HN state. In the deterministic regime, however, these excursions are limited to the HN attraction well —limited by the (yellow) separatrix.

To give a sense of these points that temporarily escape from HN, we display the background colour map in Figure 4.1. In this colour map, colour approximately indicates the relaxation time from each point of the plane to the corresponding equilibrium state (HN or LN according to the regions defined by the yellow line). Here, black/dark green indicates fast relaxation whereas clear green/white implies long-lasting relaxation excursions. In this phase plane there is another clear boundary, above the HN state, delimited by the upper strong manifold of this point (white line). Points above this boundary relax back through a long dip in N and A levels —in the Figure, short black lines indicate the field of the system and thus, the direction of these excitable dips.

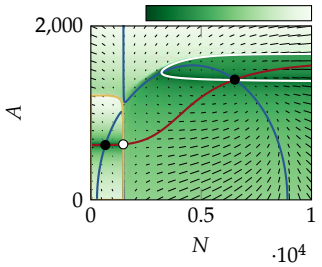

Figure 4.1: **Dynamics of the transcriptional circuit of pluripotency** Phase space of the system described by Equation (4.1). See main text for details.

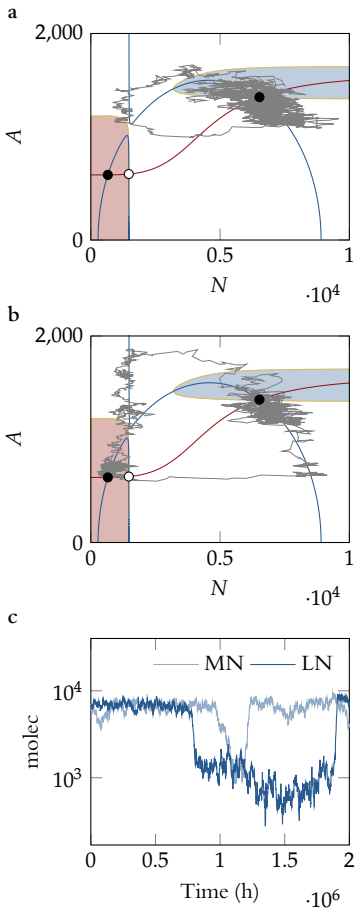

Figure 4.2: **Stochastic dynamics of the transcriptional circuit of pluripotency** (a) A simulated stochastic trajectory where the system escapes the HN state and undergoes an excitable excursion. (b) During excitable excursions, noise can bias the system towards the LN state. (c) Time traces for N of the simulations shown in (b, faint blue) and (c, dark blue).

Analogous to the cases studied in Chapters 2 and 3, random fluctuations due to molecular noise can trigger large excursions in this system,

which we identify here with cells in the MN state. Panel a in Figure 4.2 shows one such excursion in the phase plane and panel c the corresponding time trace for Nanog. These regulated excursions from HN to lower levels of Nanog triggered by molecular fluctuations —simulated with the standard Gillespie (1977) algorithm,see Appendix A.2 for details of the implementation— are equivalent to the ones described in the original model by Kalmar et al. (2009), and demonstrate that the regulated fluctuations do not really depend on the functional form of the network-level repression of Nanog. Indeed, computational analysis of the circuit indicates that these models robustly sustain the excitable regime with downward pulses (drops) of Nanog levels for a broad range of parameter values (data not shown).

In this particular model, however, the LN state can be identified with a stable state at lower Nanog and Oct4 levels. Thus, cells can steadily remain in the LN state unless randomly expelled by noise. This is a qualitative change from the original model —whereby LN was considered a purely transient state— that makes a closer agreement with the biological notion of differentiation priming attributed to the LN state. Besides, the transient excursions from HN to the MN state get close to the separatrix and, in some cases can cross to the attraction well of the LN state (shown in faint red in Figure 4.2b).

In a small fraction of cases, these orbits get trapped by the LN state for long time periods, as illustrated by the sample trajectory depicted in Figure 4.2b, and its Nanog time trace, is shown as a dark blue line in panel c of the same Figure. This larger stability of the LN state as compared to the same state in the original model, allows for longer residence times in this state, and permits a fraction of these cells to irreversibly commit to differentiation.

The topological traits of this model, summarised in Figures 4.1 and 4.2, suggest a partition of the phase space into two states that roughly approximate the biological HN and LN states —in a wide sense, not as strictly mathematical states. Specifically, the LN attraction well defined by the saddle separatrix (faint red area in panels a and b, in Figure 4.2) defines a LN state whereby cells might be primed and ready for differentiation; and the strong manifolds of the HN equilibrium point define a region (faint blue area in Figure 4.2a,b), that might be associated with robust pluripotency.

A particularity of this system is that it lives close to a qualitative change of the N nullcline that does not change the equilibrium points,

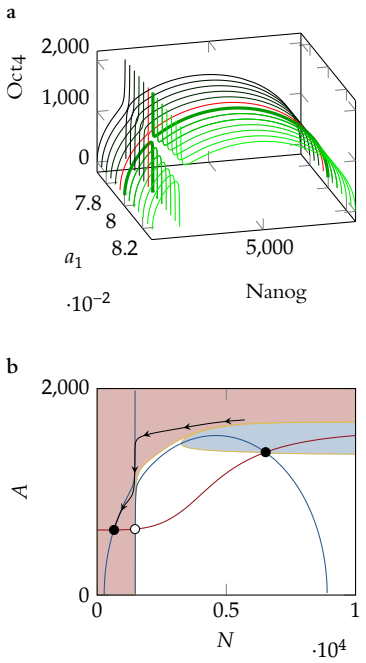

Figure 4.3: **Dynamics of the system beyond the critical point** (a) As the parameter a_1 is reduced the two branches of the N nullcline (green curves) approach and collide at a critical point (red curve). Below that point the nullcline suffers a qualitative change (black curves), with swapped semi-branches. (b) Below the critical point excitability is lost and large enough perturbations from the HN state end up in the LN state.

nor their stability. In fact, for a small change of the parameter a_1, the branches of the N nullclines cross and redefine the attraction wells of the stable states. The shape of the nullcline is shown in Figure 4.3a for different values of a_1. In this Figure, as the value of a_1 is reduced, the nullcline gradually morphs its original form (green curves) by approximating the two branches until a critical parameter value is reached where the branches collide at a single point (red curve). At this critical value four nullcline semi-branches are defined. Further decreasing the parameter, the semi-branches change partners (black curves) and define a new topology in the system. This new topology does not change qualitatively the three equilibrium points, but induces a large change in the LN attraction well, as can be observed in Figure 4.3b (red area). Under these conditions, excitability is lost as all perturbations of the HN state beyond its strong manifold enter into the LN attraction well. All in all, this proximity of the system (with all its original parameters) to the qualitative change in the nullclines generates a bottleneck –between the separatrix and the strong set of the HN state— that is visited during excitable excursions. While in this bottleneck, noise can make the system jump between the two attraction wells, somehow anticipating the dynamics beyond the critical point. This effect can thus be considered another noise induced phenomenon.

In this scenario, with the circuit generating noise-triggered excursions from the HN state (to the MN state); and with a small fraction of these being attracted by the LN state, we can estimate the resulting Nanog profile. If we assume ergodicity we can approximately recover the stationary profile of Nanog expression by simulating long trajectories of the model dynamics. This way, not only we can reconstruct the experimental histogram of Nanog (Figure 1.6), but also identify the HN and LN subpopulations according to the phase-plane division described above. Figure 4.4 shows this histogram, with the HN and LN cell populations coloured in red and blue, respectively, with the rest of cells (green area) also containing the transiting MN cells.

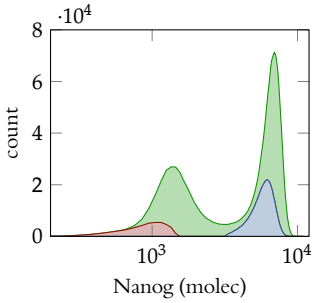

Figure 4.4: **Simulated Nanog profile with 3 subpopulations identified** Histogram for N of the stationary state of the system, decomposed in the three subpopulations (piled up): LN (red), HN (blue), and the remaining cells, which contain the MN population.

Excitability as a robust mechanism of Nanog regulated pulses

In the model proposed above, we have shown how some of the most well-established interactions between three central factors (Nanog, Oct4 and Tcf3) explain the observed profiles of Nanog expression through a bistable-excitable mechanism arising from the interplay of Nanog with Oct4 and Tcf3 and, in addition, relates well-defined areas of the phase space to observed phenotypic cell traits. Notwithstanding the usefulness of the predictions and its mathematical interest, the model is remarkably sensitive to variations of parameter values. Indeed, the exceptional shape of the N nullcline is rapidly lost as some parameters are varied, either through the bottleneck effect depicted in Figure 4.3 or, more frequently, via a Saddle Node bifurcation that annihilates the stable LN state.

In the next few paragraphs we argue that excitability is a robust trait of transcriptional regulation of Nanog. Indeed, below we intend to simplify the model, by relaxing some of the assumptions and reinterpreting others, to bring a minimal circuit that can account for regulated Nanog fluctuations by means of inverted excitable pulses.

Over the past years, a number of functional studies (Campbell et al. (2007); Kim et al. (2008); Y.-H. Loh et al. (2006); Lu et al. (2009); Matoba et al. (2006); Nishiyama, Xin, et al. (2009); A. Sharov, S. Masui, et al. (2008) ; and more recently Nishiyama, A. A. Sharov, et al. (2013); A. Sharov, Nishiyama, et al. (2011)) have gathered information on the transcriptional control of pluripotency.

As already pointed out, these studies point to a small core network of transcription factors involving Oct4, Nanog, Sox2, Tcf3, and also Esrrb.

The main currently accepted transcriptional interactions among the cited pluripotency factors are outlined in Figure 4.5. The edges of the network in this Figure do not necessarily represent simple direct regulatory events, and are based on reproducible observations. Two transcription factors, Oct4 and Sox2, activate each other and both regulate Nanog positively (Chew et al., 2005; S. Masui et al., 2007; Rodda et al., 2005; A. Sharov, S. Masui, et al., 2008). In turn, Nanog activates Esrrb expression (Festuccia et al., 2012), and probably represses its own transcription (Navarro et al., 2012); this leads to an indirect feedback on Oct4, as Essrb participates in Oct4 auto-regulation (X. Zhang et al., 2008). Oct4 activates the expression of Tcf3, though how direct is this event is a matter for discussion (A. Sharov, S. Masui, et al., 2008; Yi, Pereira, and Mer-

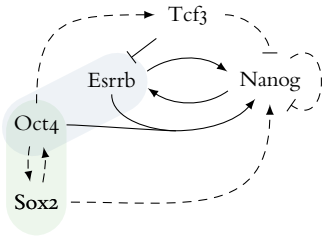

Figure 4.5: **Schema of the core network maintaining pluripotency.** Solid lines show direct interactions and dashed lines indirect connections.

rill, 2008). Tcf3, finally, represses both Nanog and Esrrb (Martello et al., 2012; Pereira, Yi, and Merrill, 2006). Indeed, experiments support the notion that this network forms a hub for pluripotency and regulates a number of secondary networks. In order to build a minimal circuit we simplify these interactions by making some experimentally supported assumptions:

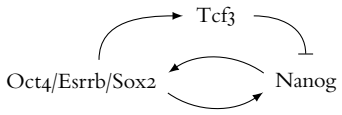

Figure 4.6: **Simplified schema of the core network maintaining pluripotency.**

i we consider that Oct4 and Sox2 behave as a single (dynamic) element (green region in Figure 4.5), since they share a common role, positively regulate each other and thus can be expected to have similar activity within the network (Chew et al., 2005; Rodda et al., 2005);

ii we can also bind Oct4 and Esrrb together (blue region in Figure 4.5) because of their similar effect on Nanog and of the observation that they operate together in this activity (Berg, Snoek, et al., 2010; Berg, W. Zhang, et al., 2008);

iii finally, we can ignore the negative feedback of Oct4 on itself via Tcf3, and the direct auto-repression of Nanog (Navarro et al., 2012).

As a result of these simplifications, we are left with the transcriptional regulatory network schematised in Figure 4.6. In this Figure, Oct4, Esrrb and Sox2 are combined into a single dynamical node, whose interactions with the other circuit elements (Nanog and Tcf3) correspond to the sum of all interactions affecting all three proteins. This network is topologically equivalent to the one we have discussed before and which can account for the fluctuations of Nanog expression (Kalmar et al., 2009). Assuming that the dynamics of Tcf3 adapts instantaneously to the levels of the other two proteins —adiabatic elimination—, the dynamics of the circuit shown in Figure 4.6 is described by the following two-dimensional differential-equation model:

$$\frac{dN}{dt} = a_n + \frac{b_n N^n}{K_n{}^n + N^n + \quad A^{\,p_A}} - d_n N \tag{4.2a}$$

$$\frac{dA}{dt} = a_a + b_a A N - d_a A \tag{4.2b}$$

Notice that Equation (4.2a) has the same functional form that the activator of Circuit C in Section 2.4. This is, the promoter of Nanog (N) in this case is auto-activated and competitively inhibited by the component A. The rate of activation of A, on the other hand, is proportional to N and does not saturate.

Again N represents the level of Nanog and now A stands for the joint action of Oct4, Esrrb and Sox2. In this model, the regulated expression

of Nanog is assumed to be cooperative (both for N and A) and subject to saturation. On the other hand, the transcriptional input on A is considered to be linear in N and to depend on A itself —in the form of a positive feedback that accounts for the mutual activation between Oct4 and Sox2—, and saturation is ignored. For the parameters chosen, the system has one single stable state that corresponds to high concentrations of Nanog and Oct4/Esrrb/Sox2. The typical behaviour of this model is shown in Figure 4.7a, which depicts the system's phase space, according to the deterministic equations (Equation (4.2)) in the $N - A$ plane, with a single stable state where both nullclines cross—shown in blue (N) and red (A).

Sporadic transcriptional fluctuations might perturb the stable state and trigger a dynamic response (solid line) that leads to a sudden decrease of Nanog followed by a slow decrease of Oct4/Esrrb/Sox2 levels. These decreases alleviate repression of the Nanog promoter; when Nanog levels are recovered so are the levels of Oct4/Esrrb/Sox2, thus returning the system to the high-Nanog steady state. Figure 4.7b shows the time traces from the stochastic simulations of the circuit (implementation details can be found Appendix A.2) and Figure 4.7a also shows cells sampled from the stationary distribution of the stochastic version by means of grey dots. The stochastic pulses generate, again, the characteristic stationary distribution of Nanog (Figure 4.7c)

This circuit, that captures the essence of pluripotency regulation, shows how excitability might regulate fluctuations that lead to the observed Nanog heterogeneity in pluripotent ES cells. Most importantly, this minimal circuit encodes robust dynamics that are preserved through large variations in the parameter values.

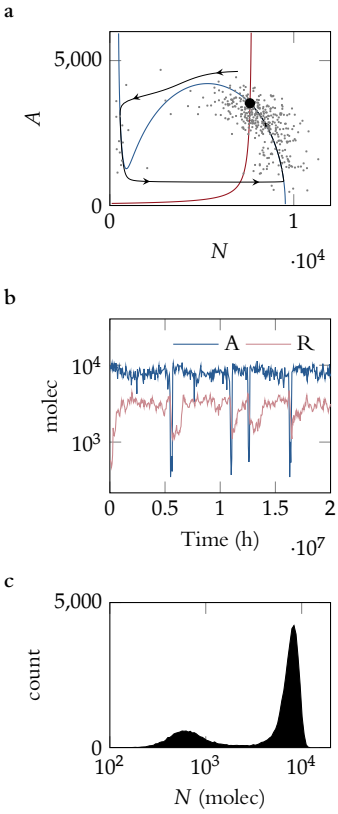

Figure 4.7: **Dynamics of the minimal transcriptional module of pluripotency regulation.** (a) Phase plane with N and A nullclines shown in blue and red, resp. Black line with arrows depicts a deterministic trajectory. Grey dots are sampled from the stationary distribution generated by the stochastic simulation of the model reactions. (b) Sample time traces from the stochastic simulations of the circuit. (c) Marginal stationary distribution of Nanog.

4.2 Correlations and post-translational mechanisms of pluripotency maintenance: the NOC model

So far we have analysed the dynamics of pluripotency regulation at the transcriptional level based on published experimental data of transcriptional interactions. A considerable amount of literature has been published in this regard, and, in fact, transcription regulation has centred the focus of a large number of studies on pluripotency in ES cells.

However, as we have recently noticed (Muñoz Descalzo, Rué, Garcia-Ojalvo, et al., 2012, and see Section 1.1) the degree of pluripotency of ES cell populations is associated with high correlation values amongst, at least, four of its key regulatory factors, viz. Nanog, Oct4, Tcf3 and β-catenin (see Figure 1.8). In other words, our observations suggest that in populations of ES cells the correlation between Oct4 and Nanog protein levels is inversely related to the probability of differentiation and therefore is a measure of pluripotency —so are Nanog/β-catenin, Oct4/β-catenin and Tcf3/β-catenin correlations, as shown in Figure 1.8. These correlations, revealed by quantitative immunofluorescence of single mES cells, are increased when cells are cultured in robust pluripotency (2i+L) conditions when compared to standard conditions (S+L, see Figure 1.8).

A salient feature of the joint distribution of Oct4 and Nanog levels in single cells is a sharp boundary below which no cells can be found in self renewing conditions (ibid.). This boundary represents a lower limit to the levels of Oct4 that a cell can have for a given amount of Nanog, and its tilted layout indicates that this limit varies approximately linearly with the latter. The mathematical models of transcriptional control of pluripotency analysed in the Sections above cannot reproduce such a linear constraint, nor can, to our knowledge, other published models of transcriptional pluripotency regulation.

This raises the possibility that post-transcriptional and post-translational mechanisms –and not only purely transcriptional ones—, which have been shown to be associated with ES cells (Buckley et al., 2012; Faunes et al., 2013) contribute to the regulation of pluripotency. In this regard, we propose and analyse below, through single-cell microscopy and theoretical modelling, a mechanism of post-translational regulation of pluripotency.

In order to test the relevance of post-translational regulation in pluripotency, we first consider a minimal network module involving only Oct4 and Nanog (Figure 4.8), on the assumption that the correlations between these two proteins result from the formation of a complex —which we abbreviate here as O:N— that is known to exist (Berg, Snoek, et al., 2010; Fidalgo et al., 2012; J. Wang et al., 2006; L. Zhang et al., 2007). In this minimal model (NOC model), we assume that Oct4 and Nanog exist in one of two forms: either free, or bound together in a complex. We do not exclude the possibility that the free forms of Nanog and Oct4 interact with other proteins to exert additional functions (see below).

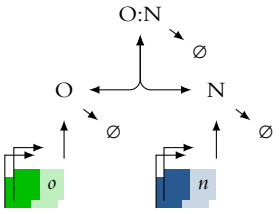

Figure 4.8: **Schema of the NOC circuit.** The NOC core protein network involves the production of free Nanog (N) and Oct4 (O) proteins and the formation of a binary complex Oct4:Nanog (O:N).

This model aims at describing the stochastic dynamics of Oct4 and Nanog expression and translation without relying on any specific transcriptional regulation —as opposed to what we have already analysed in previous Sections of this Thesis.

The post-translational dynamics of the model is approximately characterised by the simple set of rate equations for free Nanog (N), free Oct4 (O) and the Nanog:Oct4 complex ($O{:}N$):

$$\frac{dN}{dt} = k_{-}O{:}N - (k_{+}O + d_N)\,N + \quad _N(t), \tag{4.3a}$$

$$\frac{dO}{dt} = k_{-}O{:}N - (k_{+}N + d_O)\,O + \quad _O(t), \tag{4.3b}$$

$$\frac{dO{:}N}{dt} = k_{+}N \cdot O - (k_{-} + d_{ON})\,O{:}N. \tag{4.3c}$$

The $O{:}N$ complex association and dissociation processes are represented by mass action kinetics with rates k_+ and k_-, respectively. These terms assume, for simplicity, monomeric participation of both proteins in the complex, although the behaviour of the model is qualitatively the same with multimeric interaction (see below). Protein and complex degradation is taken into account through linear decay of Nanog, Oct4 and the complex with rates d_N, d_O and d_{ON}, respectively. Finally, input from the transcriptional layer —newly synthesised proteins— is accounted for by the time-dependent, non-negative terms $_N$ and $_O$. In its stochastic version, the model surmises that in mES cells cultured in standard conditions Nanog is transcribed in infrequent bursts and this entails a long tailed distribution of transcripts, as observed experimentally —see Figure 4.9 and Miyanari and Torres-Padilla (2012); Navarro et al. (2012).

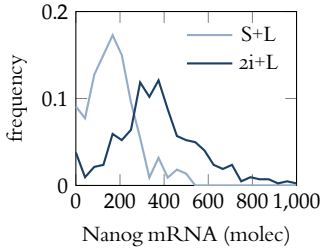

Figure 4.9: **Distribution of Nanog mRNA transcripts in pluripotent ES cells.** Experimental distribution of mRNA molecules obtained by fluorescent in situ hybridisation for Nanog mRNA (RNA-FISH) in mouse ES cells cultured under standard self-renewal (clear blue) and robust pluripotency (dark blue) conditions. Data courtesy of Lars Martin Jakt.

A key assumption of this model concerns the different stabilities for the free proteins and protein complexes. Specifically, free Nanog is as-

sumed to decay much faster than Oct4 and the complex ($d_N \quad d_O, d_{ON}$), as suggested from the protein half-life measurements showing that Nanog is less stable than Oct4 (Muñoz Descalzo, Rué, Faunes, et al., 2013).

The total amounts of Nanog and Oct4 are recovered as: $N \quad = N+O{:}N$ and $O \quad = O+O{:}N$. Intuitively, unstable free Nanog implies that the bulk of Nanog proteins will be in a complex with Oct4 and thus, loosely speaking, the system requires minimally one molecule of Oct4 for each molecule of stabilised Nanog.

As a matter of fact, in the absence of transcriptional inputs ($_N = \quad _O =$ 0) the system has one global attractor at zero towards which all points in the space relax. It is in this relaxation process where the different decay rates play a role in the generation of correlations and the sharp boundary observed in the Nanog-Oct4 distribution. A large degradation rate of free Nanog implies that during relaxation all points in space will rapidly collapse to the plane $N = 0$ —this can be regarded as a slow manifold in the limit of infinitely fast N degradation. The lack of free Nanog imposes a sharp boundary at $N \quad = O$ such that its lower region becomes rapidly devoid of points.

To illustrate this relaxation process, Figure 4.10a depicts some trajectories in the state space formed by free Nanog-free Oct4-O:N complex —the N-O-$O{:}N$ space—, corresponding to the response of the continuous model given above (Equation (4.3)) to finite pulses of Nanog and Oct4 production of different amplitudes. All these trajectories have an initial transient where free Nanog rapidly disappears in two different ways:

i if there is free Oct4 available, Nanog binds to it; otherwise

ii it quickly degrades.

As a result, all trajectories collapse onto the zero-N plane and from there Oct4 and O:N slowly decay to zero. If we translate the free variables of the system (N, O, $O{:}N$) to the measurable variables $N \quad$ and $O \quad$, as shown in Figure 4.10b, all trajectories of points below the line $N \quad = O$ collapse to it. Thus, in the absence of protein production, Equation (4.3) establishes a preferential line (or plane in the 3D space) that defines, in principle, a boundary as the one observed in the experiments.

In order to account for fluctuations in the protein production and, especially in the expression of the Nanog and Oct4 genes, we have implemented a discrete stochastic version of the model (details on its derivation

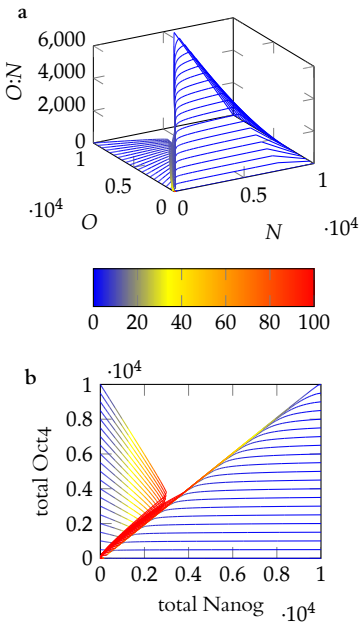

Figure 4.10: **Relaxation dynamics of the NOC model.** (a) Sample trajectories simulated with the deterministic NOC model are depicted in the N-O-$O{:}N$ state space, to show the effect of fast Nanog decay, with all orbits rapidly falling to $N = 0$. Each trajectory has been given initial pulses of Nanog and Oct4 production of different amplitudes. For each orbit colour indicates integration time from the initial condition. (b) The same trajectories in (a) represented in the N -O plane, with the exact same colour scheme.

can be found in Appendix A.2). The transcriptional layer, for which no regulation is assumed, introduces noise in the network of protein interactions —a discrete version of the functions $_N$ and $_O$ in Equation (4.3). The expression of Nanog and Oct4 is modelled with two alleles each, which can be in the active or inactive states, thus allowing for different transcriptional patterns which range from bursty monoallelic expression for Nanog in S+L conditions to its biallelic expression in 2i+L, depending on the values of the state-switching rates. The full set of reactions is shown in Table A.9.

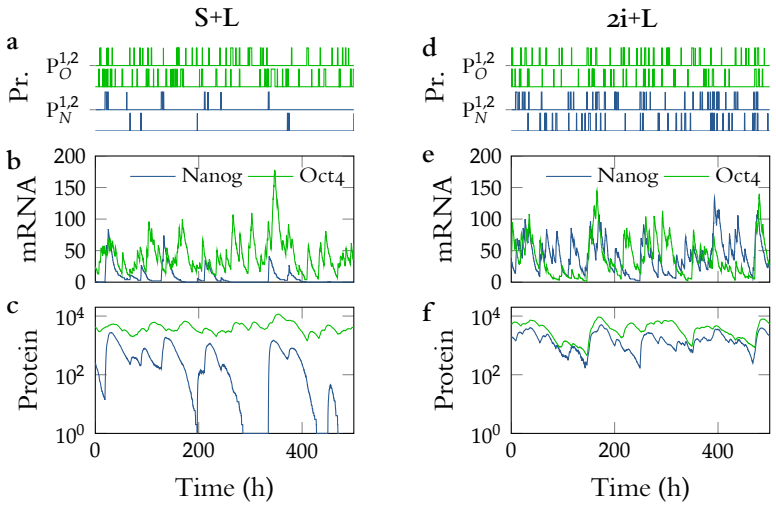

Figure 4.11: **Dynamics of the NOC model in S+L and 2i+L conditions.** Time traces, for cells grown in S+L or 2i+L conditions, of Nanog (blue) and Oct4 (green) promoter activity (panels a an d, resp.), mRNA (b and e) and protein (c and f) levels.

In the framework of the discrete model outlined here, the different culture conditions are represented by the Nanog expression level. In particular, following recent experimental results (Miyanari and Torres-Padilla, 2012; Navarro et al., 2012, and see Figure 4.9), standard self-renewing S+L conditions are implemented by monoallelic transcriptional bursts of Nanog expression, Figure 4.11a, while 2i+L conditions are modelled by increasing of the frequency of the transcription events (and thus its biallelic character, Table A.10 and Figure 4.11d). Figure 4.11 shows the dynamics of Nanog and Oct4 promoter activity (a,d) as well as mRNA (b,e) and protein levels for S+L and 2i+L conditions (d,f). This simple model is able to reproduce Nanog mRNA and protein distribution obtained experimentally in S+L and 2i+L conditions —cf. Figure 4.9 and Figure 4.12c; and also Figure 1.8f and Figure 4.12e— and, to a certain extent, that of Oct4 levels (compare Figure 1.8g with Figure 4.12f). According to the model, the joint effect of continuous Oct4 expression with bursty Nanog transcription and the higher stability of O and $O{:}N$ allows the system to explore the space above the boundary in S+L conditions (Figure 4.12a). In 2i+L conditions, on the other hand, continuous

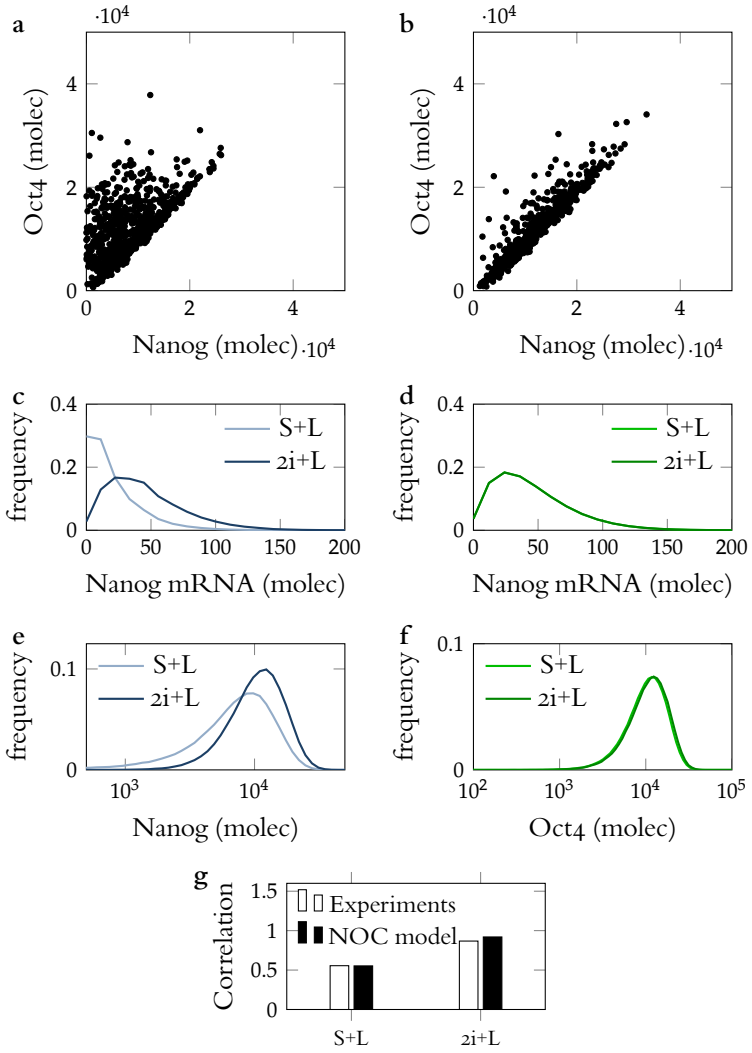

Figure 4.12: **Stochastic dynamics of the NOC model.** Simulation results in standard S+L (a) and 2i+L (b) conditions. Dots represent protein levels of cells sampled from the stationary model distribution of the NOC model. (c) Simulated Nanog mRNA molecule distributions in cells grown in S+L (light blue line) and 2i+L (dark blue line). (d) Simulated Oct4 mRNA molecule distributions in cells grown in S+L (light blue line) and 2i+L (dark blue line). (e) Simulated Nanog protein distributions in cells grown in S+L (light blue line) and 2i+L (dark blue line). (f) Simulated Oct4 protein distributions in cells grown in S+L (light blue line) and 2i+L (dark blue line). (g) Pearson correlation coefficient between Oct4 and Nanog from the experimental data, and the simulations, in cells cultured in S+L and 2i+L.

expression of Nanog leads to uniform distribution of Nanog proteins (dark curve in Figure 4.12e), which are highly correlated to levels of Oct4 (Figure 4.12b and also panel g).

To further assess the role of bursty Nanog expression in S+L conditions, we investigated the dynamics of Nanog expression using the NOC model. In this model, burstiness can be controlled without affecting the average levels of Nanog mRNA. This is achieved by modifying the mean residence-times in the active and inactive states of the Nanog promoters without modifying the ratio of time that the promoters are in the active state, i.e., a timescale parameter ϕ controls the activation and inactivation transition rates $k^{\phi}_{+1} = \phi \cdot k_{+1}$ and $k^{\phi}_{-1} = \phi \cdot k_{-1}$. The upper and middle panels in Figure 4.13 show that the parameter ϕ does not affect the ratio of promoter activity nor the average number of mRNA molecules.

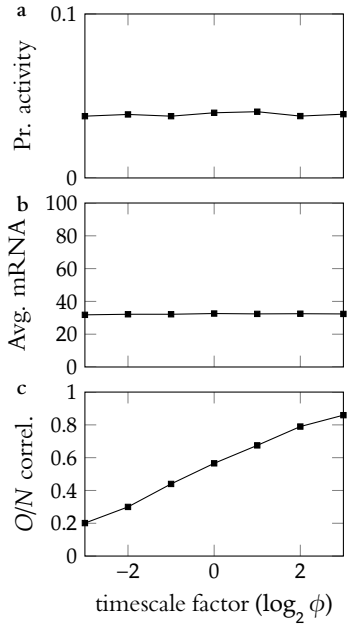

Figure 4.13: **Influence of promoter dynamics in protein correlations.** (a) Nanog promoter dynamics is a key factor controlling the final *O/N* correlation. Variation of the promoter timescale ϕ does not affect the ratio of promoter activity (top) nor the average number of mRNA molecules (centre) but modifies the O/N correlation (bottom panel). (b-d) Effect of the promoter timescale in the $O - N$ distribution for 3 different cases.

This parameter, however, does have a clear impact on the *O/N* correlation, as can be observed in the bottom panel in Figure 4.13: the slower the transition rates, the lowest the correlation achieved, as the system is allowed to remain without Nanog production enough time to become depleted. Conversely, the faster (and less bursty) the Nanog promoter is, the higher *O/N* correlation is achieved, as in the case of 2i+L conditions. Thus, in addition to the absolute levels of Nanog, the dynamics (the time scale) of its expression has a direct impact on the *O/N* correlation.

The minimal NOC model assumes direct interaction between one Nanog molecule and one Oct4 molecule. However, this interaction can be easily extended to multimeric species, which have been reported (Carey et al., 2011; Mullin et al., 2008; Saxe et al., 2009). If, for instance, the interaction is accomplished through a complex involving a Nanog *n*-mer and an Oct4 *m*-mer, the correlations will not change drastically provided the homomeric complexes are as unstable as their simple protein counterparts. What is expected to change is the ratio Oct4 to

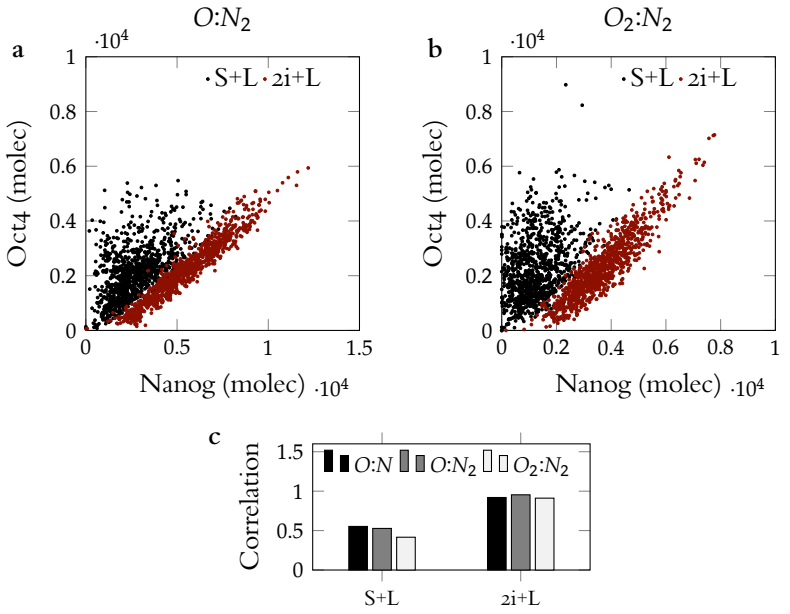

Figure 4.14: **Generalisation of the NOC model to multimeric interaction.** Simulated scatter plots showing Nanog and Oct4 levels under standard S+L (black dots) and 2i+L (red dots) conditions, considering that a Nanog dimer binds one Oct4 molecule (a) or an Oct4 dimer (b). Each dot represents one single cell. (c) Pearson correlation coefficient between Oct4 and Nanog (*O/N*) in cells cultured in S+L and 2i+L for the multimeric interactions shown in panels a and b.

Nanog, which will tend to the rate *m/n*. To illustrate this generalisation of the NOC model, we simulated the cases in which either both Nanog and Oct4 dimerise prior to interact (Figure 4.14a), or Nanog forms a dimer and then is complexed with Oct4 (Figure 4.14b). In both cases, experimental correlations can be reproduced in S+L as well as in 2i+L conditions (Figure 4.14c).

While this minimalistic model can account for the correlations be-

tween Oct4 and Nanog in the pluripotent state, it cannot explain some important observations. In particular, it does not include β-catenin, whose levels have been shown to play a significant role in the regulation of the pluripotency network (Faunes et al., 2013; Lyashenko et al., 2011; Wray et al., 2011). According to the model, absence of Nanog should lead to elevated levels of free Oct4, which experimentally has been shown to promote differentiation, and yet Nanog mutant cells remain pluripotent (Chambers, Silva, et al., 2007). This suggests that additional elements and interactions need to be incorporated into the model.

4.3 *A protein interaction network underlies naive pluripotency: the TBON model.*

Molecular analyses have revealed a dual role for β-catenin in the maintenance of pluripotency: alleviating the repressive activity of Tcf3 on Nanog (Martello et al., 2012; Wray et al., 2011) and binding free Oct4 through the formation of a complex (Abu-Remaileh et al., 2010; Faunes et al., 2013; Kelly et al., 2011). In order to add these interactions to the NOC model, we first analysed the experimental distributions of β-catenin and Tcf3 in relation to those of Oct4 and Nanog at the level of single cells (see Figure 1.8). Under standard growth conditions (S+L) there is no clear correlation between the levels of β-catenin and Nanog (β/N), and the correlation between β-catenin and Oct4 (β/O) is weak (Figure 1.8c and e, white bars). In contrast, when cells are cultured in 2i, the β/N and β/O correlations increase (Figure 1.8d and e, black bars). As in the NOC model, we consider that these correlations might be caused in part by the existence of interactions of β-catenin with Nanog and Oct4, in the form of protein complexes that have been experimentally detected (Abu-Remaileh et al., 2010; Berg, Snoek, et al., 2010; Faunes et al., 2013; Fidalgo et al., 2012; Kelly et al., 2011; J. Wang et al., 2006; L. Zhang et al., 2007).

The expanded model, TBON —after Tcf3, β-catenin, Oct4 and Nanog, which is schematically depicted in Figure 4.15, considers the existence of four protein complexes, O:N, β:N, β:O and β:T, which compete for their individual components, and attempts to reproduce the observed total quantities of protein species, which in this case are[1]

$$N = N + O{:}N + {:}N, \tag{4.4a}$$

$$O = O + O{:}N + {:}O, \tag{4.4b}$$

$$= + {:}N + {:}O + {:}, \tag{4.4c}$$

and

$$= + {:}. \tag{4.4d}$$

As in the NOC model, two alleles (promoters) are considered for Nanog, Oct4 and Tcf3, whereas for β-catenin, the model assumes a constant supply of proteins to the system. Despite it is considered a highly stable protein, in the model the available amount of β-catenin varies from cell to cell, as experimentally observed (Faunes et al. (2013) and Figure 1.8h),

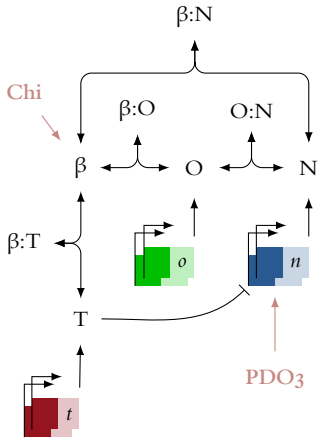

Figure 4.15: **Schema of the TBON model.** See text for details.

[1] It should be noted that in two other terms have to be considered: $P^1_{N,2}$ and $P^2_{N,2}$. These account for the Tcf3 molecules that might be bound to each of the two alleles of the Nanog promoter in its active state. See Appendix A.2 for details

which allows us to introduce a dynamic intercellular heterogeneity in β-catenin levels (see Appendix A.2 for details).

The TBON model (Figure 4.15) concentrates on how the proteins regulate and interact with one another, the only transcriptional regulation included being the inhibition of Nanog by Tcf3 (inhibitory arrow in this Figure). To model the influence of 2i+L conditions on pluripotency, we rely on the separate effects of its two constituent inhibitors: PD03 and Chiron. While PD03 is thought to increase the frequency of Nanog expression (Lanner et al., 2010; Miyanari and Torres-Padilla, 2012), Chiron is considered to stabilise β-catenin and thus neutralise the repressing activity of Tcf3 on the pluripotency network (Martello et al., 2012; Wray et al., 2011). Following these observations, we represented the influence of 2i+L conditions as a double effect (red arrows in Figure 4.15).

A pivotal factor in this model, similarly to the case of the NOC model described above, is the difference in the stability of free Nanog with respect to the complex *O:N* complex. Indeed, the model assumes a hierarchy in the half-lives of the elements of the network, and also in the stabilities of the complexes, that can be arranged from less to more stable as: Nanog Oct4 < O:N < T, β:T < β:N, β:O and β. Although there is currently no data on the particular half-lives of the different protein flavours —free or complexed—, the assumptions made by the model lead to pooled protein half-lives similar to those observed experimentally (see below).

Stochastic simulations of the TBON model (details developed in Appendix A.2) reproduce the experimental protein distributions and correlations amongst all the elements of the extended network in both standard S+L and enhanced pluripotency culture conditions (Figure 1.8c,d,f-i and compare with simulation results in Figure 4.16a-g).

The key element of the competition is the association and dissociation dynamics of the different complexes, together with the differential degradation kinetics of the different species. In this respect, we have assumed a particular grading of protein stabilities, according to their state (free or complexed). To the best of our knowledge, there is no straightforward method to estimate the stabilities of proteins belonging to different complex pools. Conversely, in vivo measurement of the half-life of the total protein —disregarding whether it belongs to a complex or not— is an easy procedure that only requires treating cells with drugs that halt the translational machinery and measuring protein levels at future time

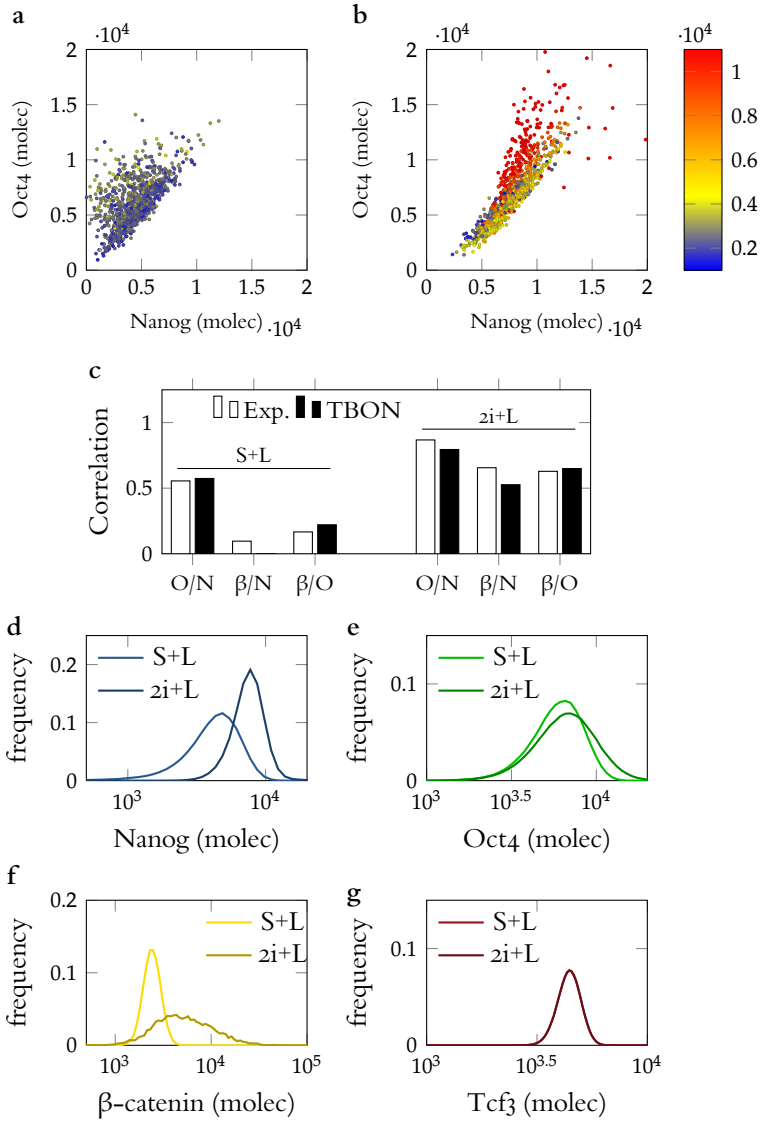

Figure 4.16: **Simulation results of the TBON model.** Simulated scatter plots or Nanog, Oct4 and β-catenin (colour coded) levels in standard S+L (a) and 2i+L (b) conditions; compare with the experimental data shown in Figure 1.8c and d. Each dot is a cell sampled from the model simulations. (c) Pearson correlation coefficient between Oct4 and Nanog (O/N), β-catenin and Nanog (β/N) and β-catenin and Oct4 (β/O) from the experimental data (white bars) and from the TBON simulations (black bars), in cells cultured in S+L and 2i+L. (d-g) Simulated protein distributions of Nanog (d), Oct4 (e), β-catenin (f) and Tcf3 (g) in mES cells grown in S+L (light colour lines) and in 2i+L conditions (dark colour lines). Compare with Figure 1.8f-i.

points either with QIF or Western Blot (WB) (Figure 4.17a-d and i). Using our model, we can simulate this situation by stopping the translation events and letting the post-transcriptional layer relax through an elaborated process that involves both protein degradation and rearrangement of protein pools. Panels e-h of Figure 4.17 show average protein levels from simulations. The results obtained are in close agreement with experimental data (compare panels e-h to a-d in Figure 4.17 and see panel i). These results indicate that the assumptions made in the particular protein species half-lives lead to pooled half-lives that are compatible to those experimentally measured.

Most significantly, this model, and its underlying set of interactions,

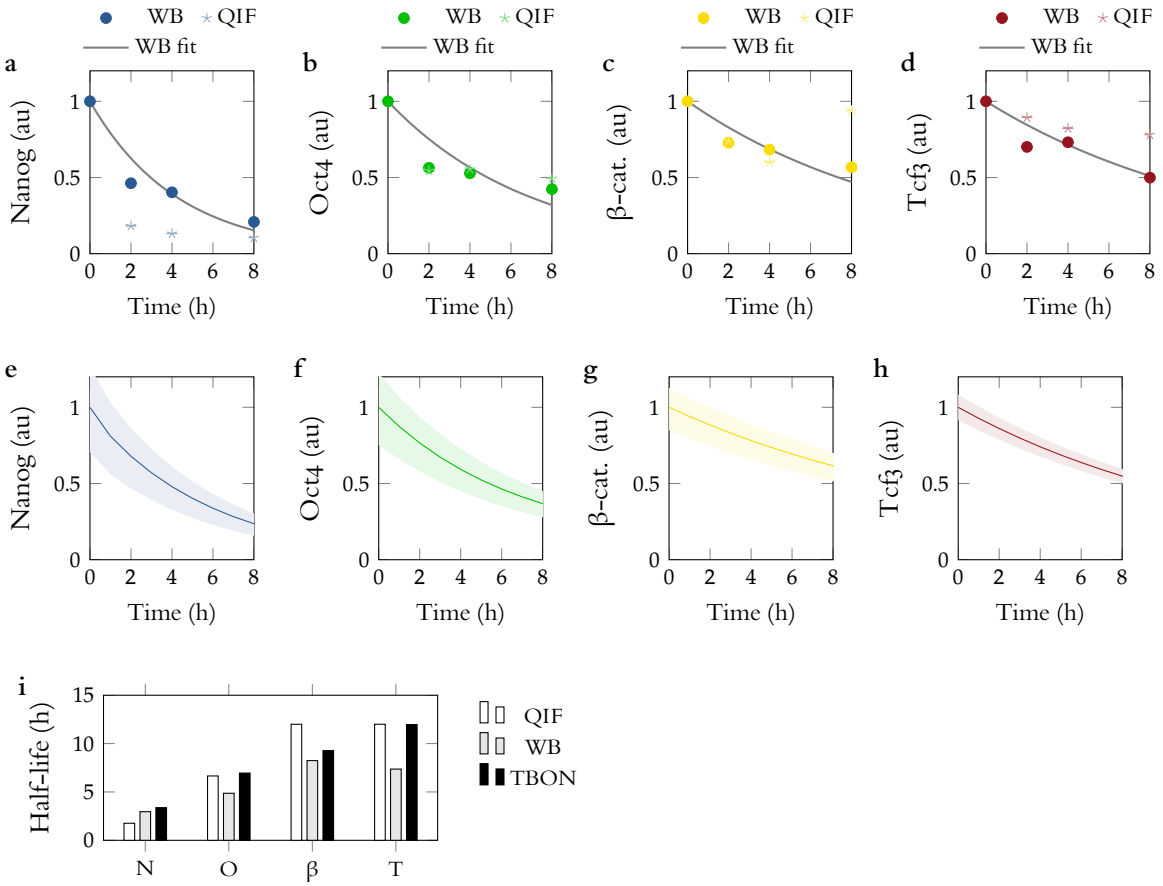

Figure 4.17: **Protein half-lives and the TBON model.** (a-d) Average protein levels normalised to unity after treatment with CHX as measured by WB and QIF. Grey curves fit a linear decay process to WB data. (e-h) Simulated protein levels normalised to unity after inhibition of translation. Stripes indicate inter-quartilic range of the distributions. (i) Protein half-lives for WB and QIF data (a-d) and simulations (e-h), obtained by fitting a linear decay processes.

provides a framework to understand the dynamics of pluripotency, which ultimately relies on the relative levels of Tcf3, β-catenin, Oct4 and Nanog, found as either part of a complex or as free molecules in individual cells. A competition between the complexes for their common components leads to a balanced equilibrium that is dependent on the culture conditions. In this sense, the model provides predictions not only of the distributions of the total amount of protein, but also of its pools of sub-species and complexes, as illustrated in panels h and i of Figure 1.8. In these panels we represent the different pools of each of the proteins by means of the pie charts. The area of these charts is proportional to the average total amount of each protein in a given condition, which allows a comparison of the total amounts of proteins between different conditions. For instance, comparison of Figure 4.16h with i indicates that total Nanog levels in S+L are smaller that in 2i+L. In addition, each circle is split into sectors representing the different pools of protein complexes and free species, with area proportional the corresponding amount of protein. A comparison between panels h and i in Figure 4.16 shows that

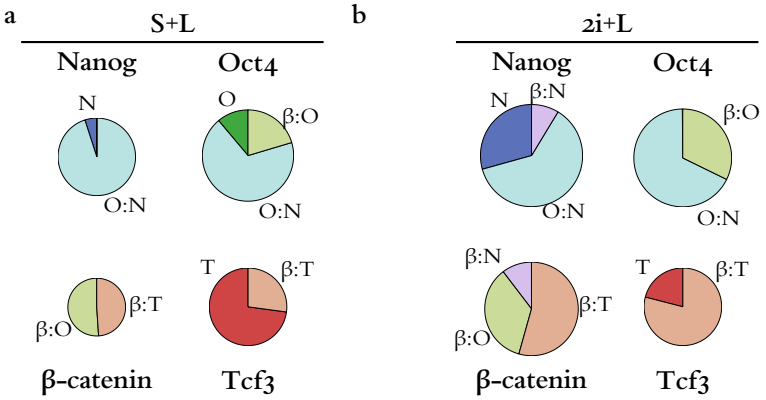

Figure 4.18: **Average distribution of protein species/complexes in the TBON model.** Pie charts showing the relative pools of Nanog, Oct4, β-catenin and Tcf3 found as free molecules (dark colours) or as part of a complex (pale colours), as determined by the TBON model in cells grown in S+L (a) or 2i+L (b). The area of each circle is proportional to the total amount of each respective molecule and so are the circle sectors to the corresponding free proteins and complexes.

the β:O, β:N and β:T complexes should be found more readily in 2i conditions than in S+L, whereas more free Oct4 and Tcf3 will exist in S+L than in 2i, as it is supported experimentally (Faunes et al., 2013). Indeed, Figure 4.19 shows that in 2i+L pluripotency conditions there are much less cells with high amounts of free Oct4.

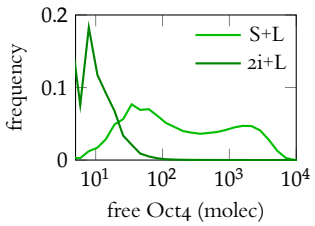

Figure 4.19: **Distribution of free Oct4 in S+L and 2i+L conditions.** Simulated free Oct4 protein distributions in ES cells grown in S+L (S+L, light green) and in 2i+L conditions (dark green).

In the light of these results, we postulate that the main function of the complexes is to stabilise the pluripotent state through the attenuation of the activity of Oct4 (Figure 4.19), represented by free Oct4, which has been shown to lead to differentiation when in excess (Nishiyama, A. A. Sharov, et al., 2013; Nishiyama, Xin, et al., 2009; Niwa, Miyazaki, and A. G. Smith, 2000). In order to investigate this possibility, we have performed a bioinformatic reanalysis of the targets of Oct4 and Nanog, which we detail in Appendix B. The results of this analysis support this contention, and show that although Oct4 and Nanog bind to many common sites in the genome, their effects on gene regulation are not the same (Figure B.1 and Muñoz Descalzo, Rué, Faunes, et al., 2013). Jointly regulated genes, which might be associated with high and correlated (balanced) levels of Nanog and Oct4, are mainly associated with pluripotency. On the other hand the unique targets of Nanog, and most clearly Oct4, include a number of transcription factors and signalling molecules associated with the emergence of germ layers and body patterning. This analysis indirectly supports the idea that the function of the protein complexes is to reduce the amount of Oct4 available to trigger differentiation: the O:N complex reinforces the activity of the pluripotency network, while the β:O buffers any excess of free Oct4 which might lead to the activation of differentiation programs.

The results and model presented here create a framework to analyse ES cell mutants, gain some insights into the molecular basis for their phenotypes, and thereby on the mechanisms underlying pluripotency in wild type conditions. Analysis of mutant cell types is theoretically achieved by adapting the parameter values of the model reactions and adding or removing reactions to the system based on known evidence and experimental observations. Changes introduced in the model to account for these perturbations are summarised in Appendix A.2, based on experimental observations.

Nanog is a key element of the ES cell regulatory network but, while it is absolutely required for the establishment of pluripotency (Silva and A. Smith, 2008), its removal in an established ES cell population only affects the frequency of differentiation of the mutant cells (Muñoz Descalzo, Rué, Faunes, et al., 2013). The model predicts that in the absence of Nanog there will be only a partial coverage of free Oct4 by β-catenin (Figure 4.20a). That is, the mutant will have a mild rise of free Oct4 levels and only the β:O complex will maintain,, albeit unstably, pluripotency in these cells (compare Figure 4.18a and Figure 4.20a). Nanog mutant cells, as predicted by the model, exhibit a moderate increase in the levels of total Oct4 and β-catenin (ibid.) as well as in their β/O correlation Figure 4.20b which might indicate that the frequency of differentiation is sensitive to the levels of β-catenin.

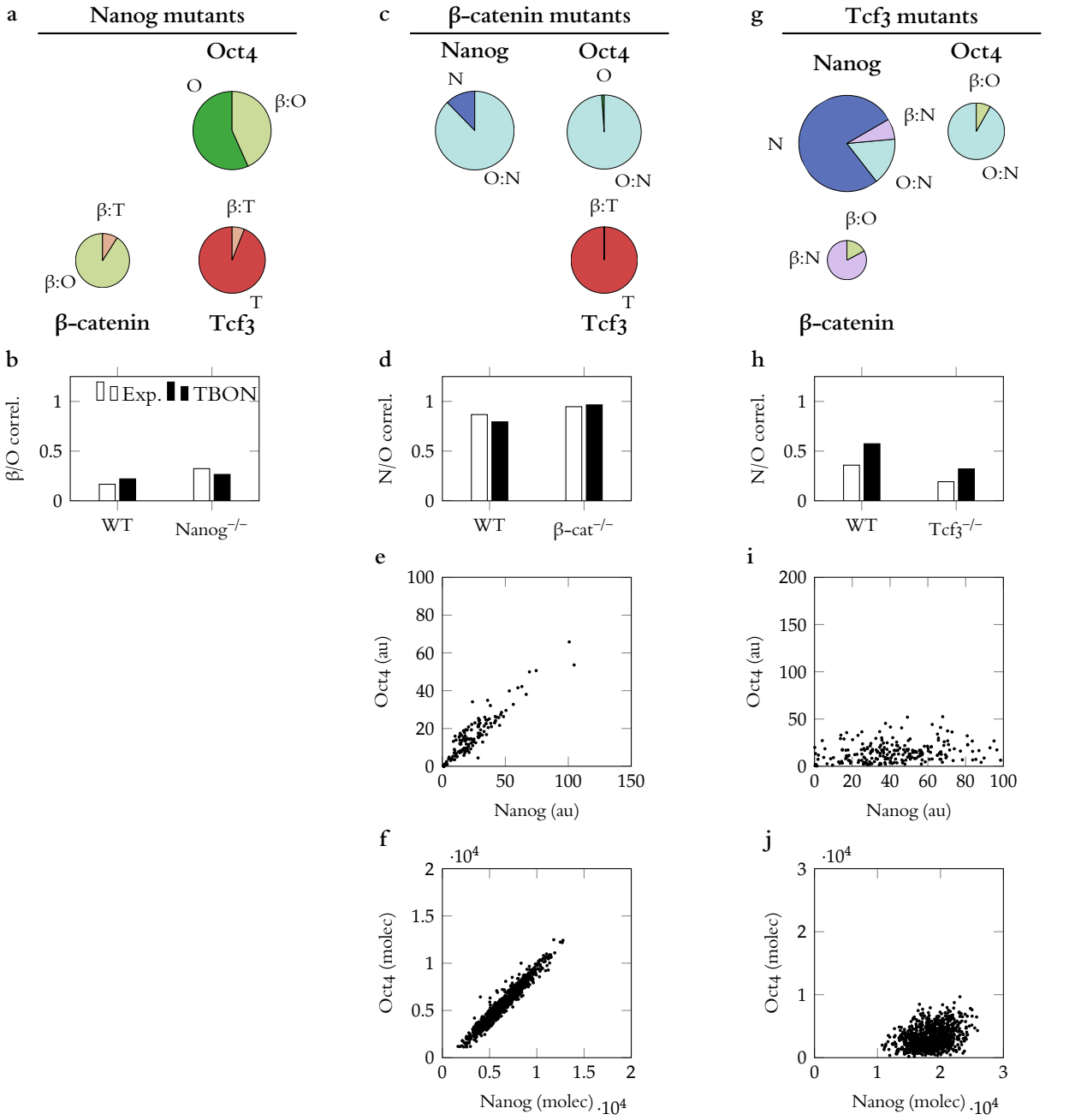

Figure 4.20: **Analysis of mutant ES cells.** (a,b) Nanog mutant cells. Protein pools predicted by the model (a) give rise to a moderate increase in the β/O correlation similar to the experimentally observed (b). (c,f) β-catenin mutant cells in 2i conditions. Loss of β-catenin induces reorganisation of protein pools (c) around O:N and retains the high O/N correlation (d). Experimental (e) and simulated (f) O-N distributions shown. (g,j) Tcf3 mutant cells. Loss of Tcf3 frees the Nanog promoter and boosts Nanog levels (g) far beyond the diagonal boundary (i) as also observed experimentally (j). This increase results in a reduction of the O/N correlation (h).

The model is also able to reproduce and provide a molecular understanding of the phenotype of β-catenin-mutant ES cells, which can only be unstably maintained in 2i+L (Faunes et al., 2013). Although the levels of total Oct4 and Nanog are low in the absence of β-catenin (Figure 4.20c, and compare with Figure 4.18b), the O/N correlation is not different from the wild type (Figure 4.20d), which is in agreement with the notion that a high O/N correlation is a central feature of pluripotency (Muñoz Descalzo, Rué, Garcia-Ojalvo, et al., 2012). The model reproduces this situation (cf. panels e and f in Figure 4.20d) but indicates that the absence of β-catenin leads to an increase in free Tcf3 levels

(Figure 4.20c). Free Tcf3 will inhibit Nanog expression, lowering the amount of the O:N complex and thus contributing to the instability of the mutant cells.

The ability of the model to reproduce and explain mutants with weakened pluripotency is important, but there are also mutants, in particular Tcf3 mutants, that have been suggested to exhibit increased pluripotency (Guo et al., 2011; Pereira, Yi, and Merrill, 2006; Yi, Pereira, and Merrill, 2008). It is thought that the pluripotency of Tcf3 mutants is due to increased levels of Nanog, but we also observe a decrease in the levels of Oct4 and β-catenin, and a collapse of the O/N correlation, as cells with low O/N ratio cross the lower boundary (Figure 4.20i). Introducing these experimental observations in the TBON model (Appendix A.2), we can reproduce the measured decrease in O/N correlation, suggesting that the excess Nanog quenches any excess Oct4 (Figure 4.20, and see Muñoz Descalzo, Rué, Garcia-Ojalvo, et al., 2012). The absence of free Oct4 molecules in the Tcf3 mutant cells could explain the delayed differentiation of Tcf3 mutant cells (Pereira, Yi, and Merrill, 2006; Yi, Pereira, and Merrill, 2008).

We have also examined the response of the TBON circuit to perturbations in Oct4. Complete loss of Oct4 results in the loss of pluripotency (Niwa, S. Masui, et al., 2002; Niwa, Miyazaki, and A. G. Smith, 2000). Indeed, no Oct4 mutant ES cells exist that can stably maintain pluripotency. Here we use mutant cells that express Oct4 via a TET repressible promoter, which can be switched off by addition of Doxycycline in the medium. Modelling of this cell type (Appendix A.2), which requires the sudden stop of Oct4 transcription and leads to a dynamic and irreversible loss of Oct4, suggests that loss of pluripotency is due to a breakdown in the balance of protein complexes. Panels a and b of Figure 4.21 compare experimental and simulated levels of Nanog at 0h,24h and 48h after cell treatment with doxycycline, and Figure 4.21c depicts the evolution of protein complexes predicted by the model. Nanog, in the absence of the possibility of binding Oct4 —which would allow it to escape fast degradation—, can only bind β-catenin, and this has a clear incidence in its levels. Following a 48 h treatment, Oct4 vanishes and this results in a new steady state with a larger pool of free Nanog subject to fast degradation (Figure 4.21c).

Finally, we have investigated the effect of Oct4 overexpression (OE) using ES cells with an additional copy of Oct4 controlled by a Doxycycline. Experimental Oct4 overexpression leads to increased levels of

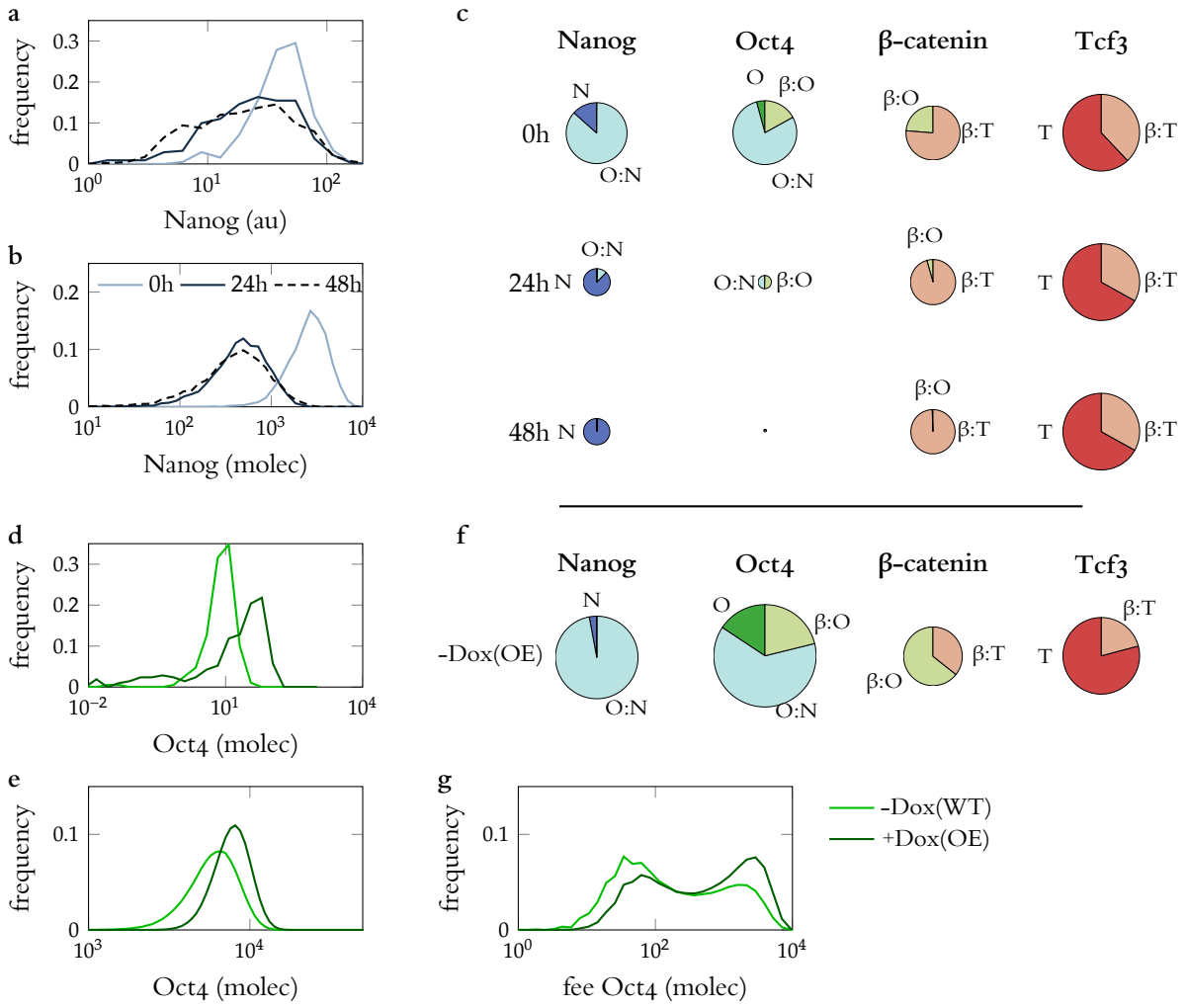

Figure 4.21: **Oct4 perturbations in pluripotency.** (a-c) Loss of Oct4. Experimental (a) and simulated (b) distributions show a decrease in Nanog levels, that stabilise after 48 h, and a reorganisation of protein pools (c) in ES cells after inhibition of Oct4 transcription. Overexpression of Oct4 (d-g), leads to elevated levels of Oct4 protein (d). The model predicts this raise (e) and the elevation of levels of free Oct4 that might explain loss of pluripotency.

Oct4 protein (Figure 4.21d) and to a higher rate of ES differentiation —i.e., lower pluripotency. Incorporation if this third promoter into the model reproduces the increase in Oct4 levels (Figure 4.21e). Indeed, this newly synthesised Oct4 proteins cannot be mopped up by the network, as shown in the composition of protein pools (Figure 4.21f), and thus remain in its free form (Figure 4.21g). Thus, according to our postulate, the high levels of free Oct4 will lead to the expression of its unique targets and therefore to differentiation, as observed experimentally (Nishiyama, Xin, et al., 2009; Niwa, Miyazaki, and A. G. Smith, 2000).

This model allows extensive and quantitative testing of pluripotency in mouse ES cells under different genetic and culture conditions. Other results rendered by the model are the separate effects of Chiron and PD03 inhibitors in some mutant cells; the recovery of pluripotency in Oct4

overexpressing cells by means of β-catenin induction with Chiron; or the effect of transcriptional auto-regulation by Nanog. (Muñoz Descalzo, Rué, Faunes, et al., 2013).

Altogether, the results put forward in this Section suggest that a noise-driven competitive protein interaction network involving Oct4, Nanog, Tcf3 and β-catenin provide a key element of the pluripotency regulatory system.

4.4 Conclusions and discussion

We have investigated the transcriptional and post-translational regulation of pluripotency in Embryonic Stem cells, and proposed several models that account for the experimentally observed phenomena.

At the transcriptional level, we have extended the previous model of pluripotency regulation by Kalmar et al. (2009) in order to incorporate the last findings regarding molecular interactions of the pluripotency factors. Inclusion of this knowledge has led us to a dynamic system (Section 4.1) that can replicate the structure of an ES cell population. That is to say, the model allows the identification of three subpopulations of cells according to its expression state: High (High Nanog (HN)) and Low (Low Nanog (LN)) Nanog populations, which have been subject to extensive studies (Chambers, Silva, et al., 2007; Kalmar et al., 2009); and a third one,Middle Nanog (MN), with intermediate levels of Nanog and considered unstable, that might have been overlooked in the past (Luo et al., 2013). The model indicates that whereas HN and LN are stable states (in the deterministic regime), MN is a transient state, as presumed by the current experimental evidence (ibid.). These results might prove helpful in understanding the internal states of ES cells belonging to each of the three described subpopulations.

The analysis of the models proposed in Section 4.1 indicates that genetic regulation of fluctuations in individual cells is a robust mechanism to control the highly reproducible distributions of Nanog levels in populations of ES cells. In a more general framework of understanding, these results reinforce the notion that, although decisions are made at the single cell level (and thus not reproducible), they result in highly consistent and robust patterns at the macroscopic (population) level (Garcia-Ojalvo and Martinez Arias, 2012).

In Sections 4.2 and 4.3 we propose a novel mechanism of regulation

of pluripotency in mouse ES cells that involves the formation of protein complexes between key factors that alter their stabilities. We have proposed a minimal model of post-translational regulation between Oct4 and Nanog (Section 4.2) that, despite its simplicity, can account for the observed correlations between these two factors under different pluripotent conditions (Muñoz Descalzo, Rué, Garcia-Ojalvo, et al., 2012). By experimentally measuring the levels of four key pluripotency factors in individual cells, we have expanded the complex formation module to a whole competitive network of protein complexes involving Oct4, Nanog, Tcf3 and β-catenin. The new (and larger) model accounts for the major characteristics of pluripotency in mouse ES cells. At the centre of this network is the transcriptional activity of Oct4, whose levels are crucial for pluripotency and differentiation (Nishiyama, Xin, et al., 2009; Niwa, S. Masui, et al., 2002; Niwa, Miyazaki, and A. G. Smith, 2000). We suggest that the Oct4 levels are dynamically buffered through the interactions of four protein complexes (O:N, β:N, β:O, and β:T) that compete for their common elements. The result of this competition is that the levels of active Oct4 are restricted to a range that promotes pluripotency though fluctuations in those levels lead to cells with higher differentiation-promoting potential. A prediction of our model that the lower the levels of Oct4, the higher the degree of pluripotency finds support on recent experiments that show that cells with low stable levels of Oct4 are unable to exit pluripotency (Radzisheuskaya et al., 2013)

The plethora of recent studies is making increasingly clear that pluripotency does not rely on a single linear program of transcriptional activity, but rather on the convergence of several molecular activities that balance self-renewal and differentiation. Our results suggest that a noise-driven competitive protein interaction network involving Oct4, Nanog, Tcf3 and β-catenin provide a key element of this regulatory system.

III

Relaxation and noise in cellular signalling networks

5

Noise, relaxation dynamics and information routing in a cellular signalling network

The networks of genes and proteins that regulate cellular behaviour are subject to non-negligible amounts of random fluctuations, or noise, arising from the intrinsic stochasticity of cellular components. These regulatory networks are, in addition, embedded in temporally variable environments, as a result for instance of circadian and ultradian rhythms affecting the organism to which the cell belongs. The question of how regulatory networks, and in particular signalling networks, respond to this dynamic driving is still open due to an incomplete knowledge of the underlying complex pattern of interactions.

In this Chapter we focus on a Boolean network (BN) model recently proposed by Helikar et al. (2008) that describes the signal transduction network of a typical human fibroblast. The model was created by careful inspection of a large body of experimental literature, resulting in the network schematised in Figure 1.10, which contains 9 inputs that feed signals to 130 internal nodes, linked with each other through a web of 542 interactions.

Below we describe the main structural characteristics of the network and, in Section 5.1, we examine the relaxation dynamics of the network both in the absence and the presence of external fluctuations. We also investigate its response to external cyclic signals in Section 5.2. We then turn our attention to the role of external noise in the information propagation process. In particular, we describe the sensing capability of the network to background chatter (Section 5.3); the effect of this chatter to the network response to periodic stimuli (Section 5.4); and how this network response emerges through robust reorganisation of internal information transmission paths (Sections 5.5 and 5.6).

Mathematically, the BN used here (termed *original* network hereafter) consists of 139 elements, or nodes, connected by 542 links. Nodes represent chemical species, which are assumed to be either active (1) or inactive (0), and edges represent their interactions.

From a topological point of view, this network is very different from classical random Boolean network (RBN) networks (or Kauffman networks, Kauffman, 1969). The average number of inputs per node (in-degree) and the average number of nodes affected by one given node (out-degree) are both 3.9. Neither the in-degree nor the out-degree distributions are uniform throughout the network, contrary to RBN networks (Figure 5.1). For instance, while there are 21 nodes with a single input, the Src protein is affected by 11 different nodes (its in-degree is equal to 11) and affects 28 downstream nodes (its out-degree is equal to 28). In fact, Src is the protein with highest out-degree and closeness centrality. Other interesting nodes are Rac, which integrates information from 13 species (maximum in-degree) and Erk, which is the node that participates the most in the shortest paths among all pairs of species (highest betweenness centrality).

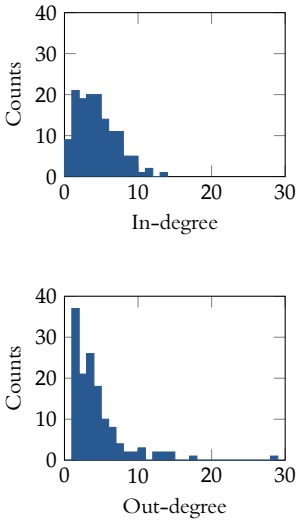

Figure 5.1: **In- and out- degree distribution of the signalling network.**

5.1 *Relaxation dynamics of a noisy Boolean cell signalling network*

The network implements standard Boolean dynamics. That is to say, the states of all species are updated synchronously at each iteration according to a set of node-specific deterministic logic rules. We consider here two types of nodes, determined by their definition: input nodes and internal nodes. On the one hand, the state of input nodes is independent from the state of other nodes of the network and determined externally. These nodes represent signals outside the cell (i.e., the external environment), and may evolve deterministically (for instance, being maintained to a fixed value of activity or oscillating at a specific periodicity) or stochastically (see below). On the other hand, every internal node i has an associated logic rule, f_i, which determines the new state $x_i(t)$ at time t from the states of its k_i incident nodes, $x_{i_1}, \dots x_{i_{k_i}}$ (note that these may include the node i itself), at time $t-1$ (see Figure 5.2). Therefore, for given input nodes and initial conditions of the internal nodes of the network, the states of all nodes evolve in a deterministic and reproducible fashion.

In this manner, the dynamical Boolean system evolves in time. The nodes in the network change their activity state according to the logic rules acting upon them, their current state, and the states of the input signals. Thus, any combination of constant input signals transforms

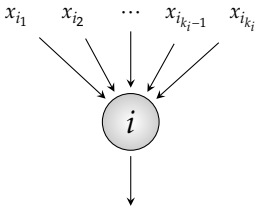

Figure 5.2: **Dynamics of internal nodes.** Each internal node i updates its state at time t synchronously according to the states of its input nodes, $i_1, \dots, i_{k_i}$ at time $t-1$ and its logic rule $f(\cdot)$.

the system into a distinct autonomous system. A typical evolution for a combination of constant inputs is shown in Figure 5.3a. In this example, input nodes are fixed to 000111101 —i.e., following the order of the inputs in Figure 1.10, Stress, IL1/TNF-α, ECM, and α_q are inactive, while Calcium pump, EGF, α_s, α_i, and $\alpha_{12/13}$ are active. The system relaxes, during a transient period, towards a stationary configuration in which some nodes are inactive (black in the figure), some are constantly active (white), and a few cyclically alternate between active and inactive states in a four-step period.

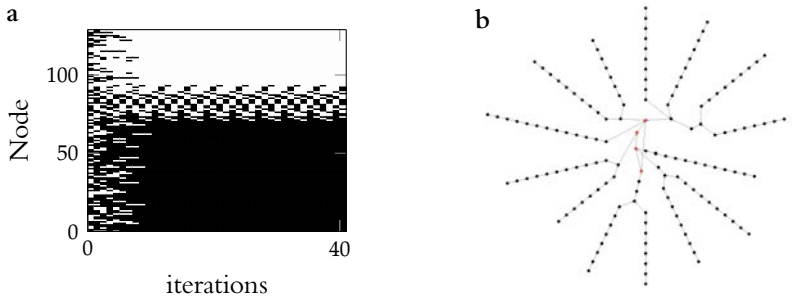

Figure 5.3: **Dynamics of the network.** (a) Evolution of the network nodes for random initial conditions and a fixed combination of input states (000111101) White (black) regions correspond to active (inactive) states. After some initial transient, the network reaches a limit cycle attractor where the activity of some of nodes oscillate periodically, while that of the others is fixed to either the active or inactive state. (b) Example of a cyclic attractor with four ever-repeating states and 14 different transient orbits, with lengths between 9 and 13 iterations,that are attracted to it.

In the random Boolean network literature this is known as the ordered regime. These types of dynamics, with periodic attractors, are an immediate consequence of having a finite number of possible states —2^N for a system with N nodes— and deterministic evolution rules. Figure 5.3b depicts one such periodic attractor. In this Figure, circles represent network states and arrows (shown only as edges) indicate the deterministic transitions between these. Red circles belong to the final attractor and thus are connected in a circular topology. Black circles, on the contrary, indicate network states belonging to transient orbits that flow into the attractor.

Since we are interested in the dynamics of this system, we start by considering the distribution of transient lengths obtained for fixed inputs. This network accepts 2^9 different combinations of input signals (the 9 inputs are fixed to 0 or 1) which will necessarily produce periodic or fixed orbits defining the corresponding relaxation (transitory) duration. To calculate the transients we search first the stable fixed points and periodic orbits reached by the system at long times. Once we identify the attractor for a given specific simulation condition, we define the transient length as the number of iterations needed to reach the attractor.

A full exploration of the phase space of this system is computationally unfeasible. Hence, we partially explore this space by sampling 3000

random initial conditions for every combination of constant input signals. This represents a total number of 1.536×10^6 realisations of the dynamics. From these realisations we have identified 264 206 attractors with very different topological structures. This result shows that the structure of the phase space is far from trivial and totally depends on the activity values of the input signals. Thus, changing these inputs, the shape of the deterministic attractors varies. As an example of this dependence on the input signals, Figure 5.4 (page 103) shows the partial exploration of the network state space for four different combinations of inputs. From this Figure, it is clear that there is a large variation in the number of attractors found from each combination of input signals. Whereas for some input combinations there are only 10 different attractors in the 3000 realisations (panel a), for others the number of attractors reaches 2168 (panel d).

Figure 5.5a shows the histogram of transient durations measured from this set of simulations. The distribution exhibits both a sharp peak of transient lengths centred at around 10 iteration time steps —average and median transient lengths are 22.8 and 16 respectively— and a fat tail with approximately 25 % of the simulations having transients larger than 25 iterations. Each individual input condition leads qualitatively to the same kind of distribution, when only its initial conditions are varied. Thus, a network as complex as the one represented in Figure 1.10 produces, in general, a well defined temporal sequence of responses to static input conditions. To some extent, this behaviour underlies the reliability and adaptability of the system to changes in the external inputs: a well-defined response guarantees that the system will adapt its dynamics always at the same pace. A dispersion of some orders of magnitude in the response duration would produce a variety of evolutions for different conditions, which would difficult the interaction with other cellular processes. On the other hand, we speculate that a well-defined response allows for a predictable and simple behaviour that may be easily coupled with other cellular mechanisms.

The distribution of periods for the limit cycle attractors is not as sharp as that of the transient lengths but shows certain dispersion, with some forbidden periodic orbit lengths (these correspond to flat regions in the complementary cumulative histogram shown in Figure 5.5b). However, here we do not concentrate on the network attractors because they are not robust in the presence of noise, as we discuss below. We will see next that the relaxation length, on the other hand, does persist in the presence of fluctuations.

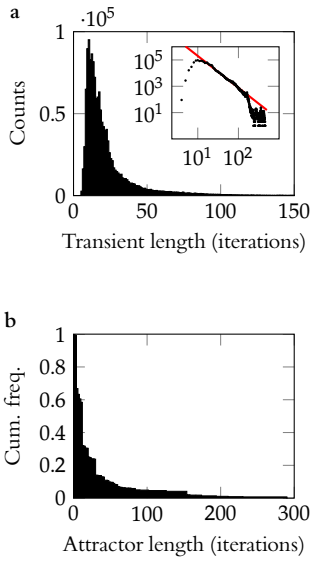

Figure 5.5: **Relaxation statistics of the network.** (a) Histogram of the transient lengths for the 1.536×10^6 different random initial conditions (see main text). Transient lengths were computed by first finding the attractors and then identifying the iteration at which the attractor is reached. The inset shows the histogram in doubly logarithmic axes, which reveals that the tail of the histogram follows a power law (red dashed line) with scaling exponent 2.4. (b) Complementary cumulative histogram of the attractor lengths for the same initial conditions as in (a).

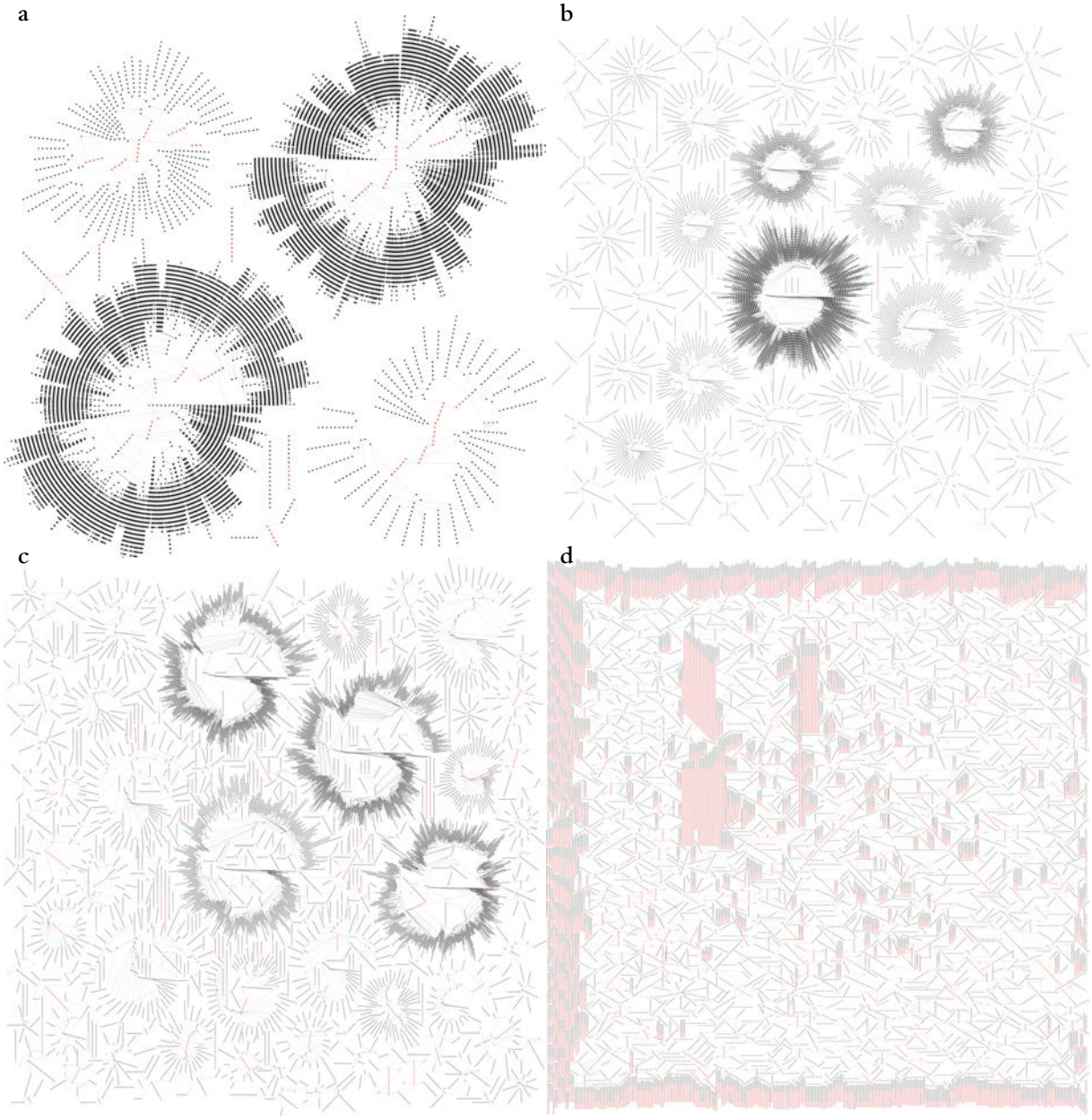

Figure 5.4: **The topology of the state space varies for different inputs.** Partial exploration of the fibroblast network phase space for four combinations of constant inputs. Each dot represents a state of the system and is linked by an edge to the next state according to an iteration of the network dynamics. Each panel shows the states found relaxing the network for 3000 random initial conditions. (a) 10 attractors found, two of which are very large limit cycles (inputs: 101011101). (b) 93 heterogeneous distribution of attractor sizes (inputs: 000000000). (c) 320 heterogeneous distribution of attractor sizes with some large limit cycles (inputs: 100110000). (d) 2168 small attractors creating a disconnected phase space (inputs: 001100011).

Each realisation of the dynamics of the network model can be considered as the dynamics of a single cell. To study the response of a population of cells we average the activity of each node, i, at every iteration, for the $R = 3000$ realisations of the dynamics described above:

$$X_i(t) = \frac{1}{R} \sum_{\text{cell}=1}^{R} x_{i,\text{cell}}(t). \tag{5.1}$$

The average activity, $X_i(t)$, takes a real value between 0 and 1, which

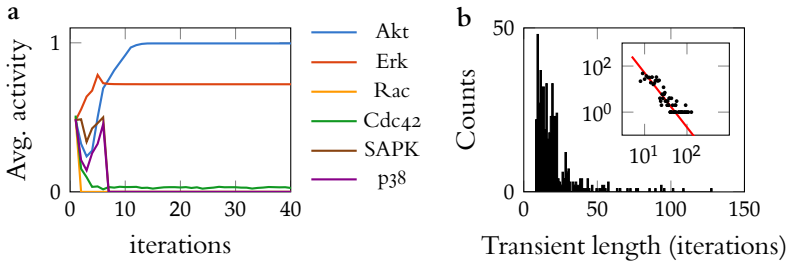

Figure 5.6: **Population dynamics of the system for constant inputs.** (a) Population average for output nodes obtained from 3000 realisations of the network dynamics with different initial conditions, and the same fixed combination of input states as in panel (a). Different combinations of fixed inputs lead to different population averages of the outputs. (b) Histogram of the transient lengths of the population dynamics for the 512 possible combinations of input states. Transients for the population dynamics are computed as the maximum transient lengths of the average activities of single nodes. These node-specific transients were calculated by first establishing the range of values of the stationary state (within a certain tolerance) and then finding the iteration at which the node's average activity value enters this range and never leaves it again. The population relaxation durations are close to the median transient lengths of individual realisations of the network dynamics (cf. panel b and Figure 5.5a). The inset depicts the histogram in doubly logarithmic scale. The plot indicates that this histogram has again a fat tail, although in this case the number of data points is too small to reveal a clear-cut power law.

quantifies the probability of the corresponding node to be active at time t. The question is now whether the average activity behaves in the same manner as the single-cell dynamics that we have examined above, and in particular if it exhibits a well-defined transient duration. It should be noted here that we are not coupling networks, we simply deal with statistics of the network over different initial conditions for fixed sets of inputs. Figure 5.6a shows the temporal evolution of the average activity of the six output nodes of the network, as defined in Figure 1.10, for the same input conditions shown in Figure 5.3. In this particular case, some of the output nodes exhibit a zero activity, while some others reach an approximately constant non-zero activity. The distribution of transient lengths at the population level —which we can measure as the maximum time required for every node to reach its final value—, shown in Figure 5.6b, is similar to the distribution for single cells (cf. Figure 5.5b).

The environment in which cells are embedded changes irregularly in time. As a result the cell receives noisy, randomly fluctuating signals, which introduce new and usually fast time scales, affecting its dynamics. To introduce this unpredictable evolution of the cell environment into the model, here we allow random fluctuations in the activity of the input nodes. We define a probability q for an input to be active, and at each iteration of the network dynamics we draw the state of this input node from a Bernoulli distribution with probability of success equal to q —i.e., $\mathrm{P}\ [x_i(t) = 1] = q$ for input node i. The parameter q therefore determines the average level of background chatter in the network, which is lowest for $q = 0$ and highest for $q = 1$. Note that q additionally controls the degree of variability in the input sequence, which is maximal at $q = 0.5$, and disappears for both $q = 0$ and $q = 1$ (see Figure 5.7).

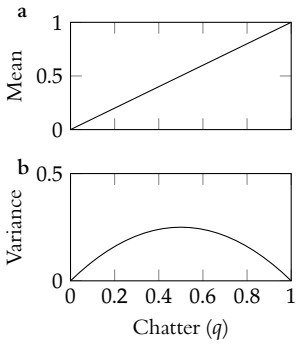

Figure 5.7: **Chatter — input noise** Chatter level at each input node is determined by q, the probability of being in the active state and thus its average value (a). The variance is given by $q(1-q)$ and reaches its maximum at $q = 0.5$.

This chatter model, together with the ensemble representation introduced in Equation (5.1), allows for a realistic description of stochastic fluctuations, as it dilutes the effect of flipping input-node states at the macroscopic level (Bornholdt, 2008), provided the number R of cell realisations is large enough. With this description, the chatter levels correspond to the population average activities of the input nodes.

The random sequences of activity states obtained for inputs with a constant chatter level provide the network with an unstructured signal that lacks temporal information. We consider in this Section the case of maximum variability, $q = 0.5$, for all input nodes.

In order to investigate the effect of noise on the response of the system, we have performed simulations with different initial conditions and realisations of the chatter. An example of the evolution obtained for a given realisation is given in Figure 5.8a. This plot shows that the chatter destroys the deterministic attractors obtained for constant inputs, described in the previous Section. At the population level, however, each node reaches an approximately constant average level of activity, as shown in Figure 5.8b, which plots the population activity of the six output nodes of the network —the intermediate nodes have a qualitatively similar behaviour. A comparison between Figure 5.5a, Figure 5.6b and Figure 5.8b indicates that the transients do not change qualitatively in the presence of chatter. As in the case of constant inputs, we can measure the population relaxation duration as the number of iterations needed for every node to reach its stationary average level. For the particular case of Figure 5.8b the estimated transient length is 11, which falls within the distribution of deterministic transients shown in Figure 5.6b.

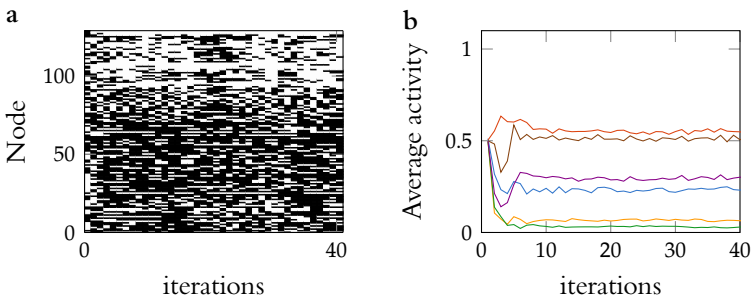

Figure 5.8: **Effects of chatter acting upon the input nodes.** (a) Dynamic evolution of the network nodes for a given realisation of chatter ($q = 0.5$) affecting all input nodes. In this case the evolution of the network is erratic and does not reach stable or periodic attractors [compare the evolution shown here with that presented in Figure 5.5a]. (b) Population average for the output nodes [colour coding as in Figure 5.6a] obtained from the 3000 realisations of the network dynamics for a fixed value of chatter $q = 0.5$ for all input states.

This robustness of the relaxation of the network dynamics in the presence of noise is a desirable property of the system. However, it is also important to see how flexible the network is in dealing with different time scales, because in certain situations, as in the presence of circadian rhythms, the system is coupled to signals that may vary in time with a

well-defined spectral structure. To that end, we now study the response of the network to temporally modulated signals.

5.2 *Network frequency response to cyclic environments*

In order to determine how well the network reacts to a variety of external time scales, we now consider an input condition in which a specific input node is subject to a periodic alternation of its activity level, while the remaining inputs receive a noisy chatter background. As we have done in the previous Section, we simulate different noise realisations for the same inputs and random initial conditions, and calculate the population average of the output nodes —which can be related to specific cellular responses. We quantify the frequency response of these nodes by calculating the power spectral density (PSD) of the corresponding average signals, using the Welch periodogram method. In Figure 5.9a we show the PSD of the average activity of a specific output node (Akt) for increasing stimulation frequency, ν_0, being applied to a particular input node (Stress), when the rest of inputs are subject to a fixed chatter level $q = 0.5$. This result shows that the periodic input signal is transduced by the output nodes to some extent: the peak of the spectrum of the output node's activity (darker region) corresponds in all cases to the frequency applied to the input signal. The strength of the response at the input frequency, however, depends on the value of that frequency, indicating that the network does not respond equally well to all frequencies. To show that, we plot in Figure 5.9b,c the values of the power spectral density of the average activity of all six output nodes at the input frequency ν_0, when this frequency varies. The Figure shows two example cases (panels b and c) corresponding to the periodic modulation of two different input nodes, and reveals a variety of frequency responses including low-pass, high-pass and band-pass filtering, together with other more complex behaviours. In all cases, however, the frequency cutoffs are located in the range ν_0 0.05–0.20, which correspond precisely to periods in the range 5–20 that are close to the peak of the relaxation durations that we have observed in previous Sections.

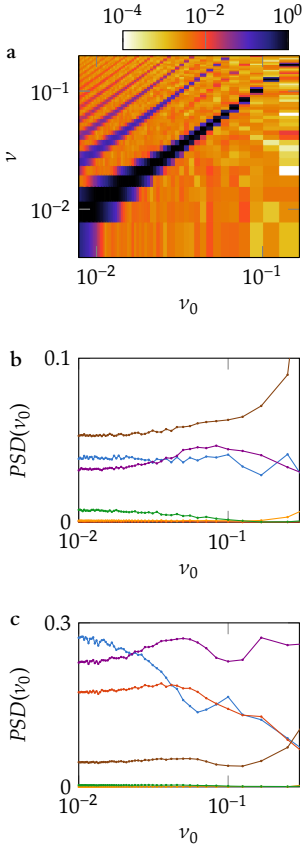

Figure 5.9: **Response of the network to periodic driving.** (a) Power spectrum of the average activity of an output node (Akt) versus stimulation frequency, ν_0, being applied to a given input node (Stress), while the other input nodes are subject to a fixed chatter level ($q = 0.5$). (b,c) Frequency filtering of the network. Power spectral density at the input frequency ν_0, normalised to the power of the input signal, as a function of ν_0, for one input node and fixed chatter level ($q = 0.5$) for the rest of inputs. The periodic driving is applied to (b) Stress and (c) EGF input nodes.

It is interesting to note here the influence of chatter on the network's response to external signals. As we have seen, the noisy signal introduced by chatter destroys the structure of the deterministic attractors of the system. By contrast, it maintains the characteristic relaxation times of the network, and helps transducing structured inputs, in the form of periodic signals, into their corresponding outputs (see Domedel-Puig

et al. (2011) and next Section). To some extent, the dynamics in the periodically driven case become a sequence of relaxations between the two states involved in the driving signal. The synergistic effect between noisy and structured inputs found here is an example of integration of signals by cells. Given that cells operate in a noisy environment while being required to respond to temporally structured inputs, we can expect the results obtained here to be relevant in realistic settings.

In this Section and the previous one we have analysed the dynamic response of the fibroblast network to perturbations (its relaxation dynamics) and to periodic forcing (oscillating signals) in the presence of chatter, or background noise. The focus of these two Sections has been put mainly on the dynamics and time-scales of the network. By contrast, the remaining Sections of this Chapter are committed to the analysis of the effects of varying levels of external noise on information propagation through the network.

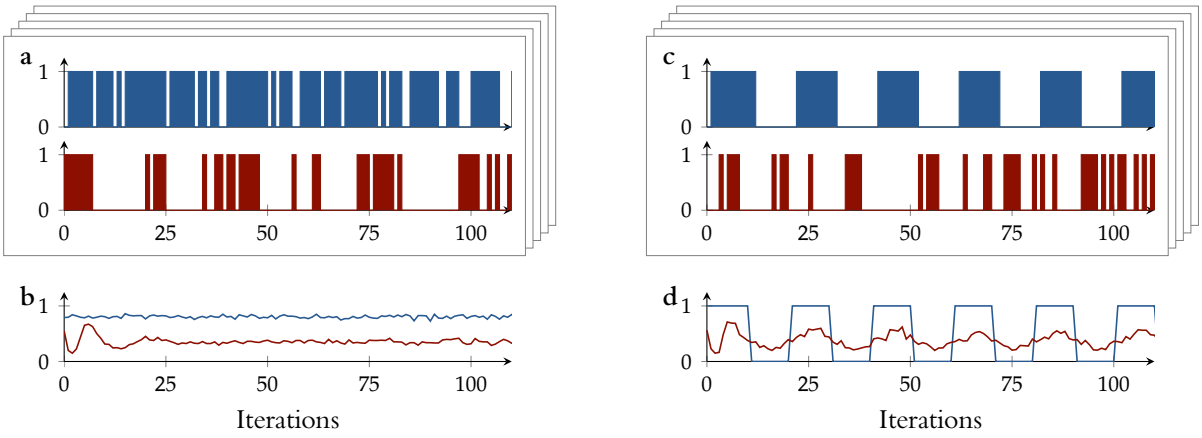

Figure 5.10: **Population dynamics of the cell signalling network.** (a) Single-cell response of the network to a constant chatter level. Here we show a realisation of the network dynamics in which the states of all input nodes randomly switch between 1 and 0 at a chatter level of $q = 0.8$. For clarity, we show the evolution of only one input node (Stress, in blue) and one output node (p38, in red). (b) The population average signal for the two nodes, calculated averaging data from R=201 realisations, is shown at the bottom. (c) Single realisation showing the response of the network to periodic stimulation of one input node. We show one network realisation in which input node stress oscillates (period $_0 = 20$), and the other input nodes are set to chatter level $q = 0.8$. (d) Recovered population average signal shows that output nodes also oscillate.

Figure 5.10a depicts a realisation of the network dynamics obtained with unstructured sequences for all inputs set at chatter level $q = 0.8$. For clarity, only the input Stress is shown, in blue and the output p38 in red. At the population level, both nodes show constant activity after a short transient (see Figure 5.10b) and, in this case, averages over cell realisations are equivalent to temporal averages.

In contrast to the unstructured case, oscillating (structured) inputs whose states turn on and off periodically introduce temporal information to the network that propagates through the network. Panels c and d of Figure 5.10 illustrate this situation. In this case, the input Stress oscillates and the rest of inputs fluctuate with a fixed level of chatter, $q = 0.8$. The average population activity of the output node p38 also oscillates at the periodicity of the input, therefore recovering the temporal information

supplied by the external stimulation. Below we analyse and describe the conditions in which these structured and unstructured signals are transmitted, and the role played by background chatter during this process.

5.3 *The logic of the original network allows sensing the background chatter level*

To study the contribution of chatter to the network dynamics, we again consider the response of the network to unstructured inputs. These have been implemented with a constant chatter level q for all the input nodes, as in the previous Section. Under these conditions, the population activities of the output nodes reach a steady state that corresponds to its temporal average

$$\langle X_i(t)\rangle = \lim_{\infty} \frac{1}{R} \sum_{\text{cells}=1}^{R} \sum_{t=0} x_{i,\text{cell}}(t)$$

$$= \lim_{\infty} \frac{1}{} \sum_{t=0} X_i(t),$$

where $X_i(t)$ is the ensemble or population average defined in Equation (5.1). Figure 5.11a shows the temporal average of the population signal for all the output nodes, for increasing levels of chatter. The average population activity increases monotonously for three of the outputs (Akt, Erk and Rac) as the chatter level increases. In particular, we observe that the average activity of Erk is approximately proportional to the chatter level. On the other hand, the average population activities of the other three outputs (Cdc42, SAPK and p38), depend non-monotonically on the chatter, becoming maximal for an intermediate value of q. Thus, the original network responds to background chatter in a non-trivial manner.

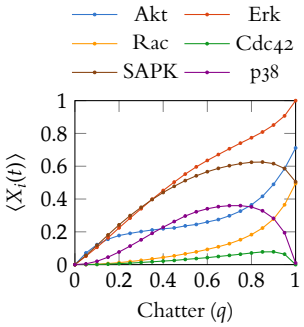

Figure 5.11: **Temporal average of the population activity of the output nodes for increasing chatter levels.** In these simulations, the activity of all input nodes is unstructured (i.e., the probability of the input nodes being active is set to a constant value q). The average activity of some output nodes increases monotonically, while it peaks at intermediate chatter levels for other outputs.

We now ask to what extent the effects of constant chatter described above can be attributed to the specific connectivity architecture of the fibroblast network being used here. To address this issue, we generate and study the behaviour of random versions of the network. We have taken two distinct approaches to randomise the original network, keeping in both cases the topology of the network constant, while varying the rules of the nodes. In the first case (altered logic (AL), see Figure 5.12), the network is randomised by sampling a completely new random rule for each node, with the only constraint that the number of ones in the truth table is the same as in the original table. The algorithm used to

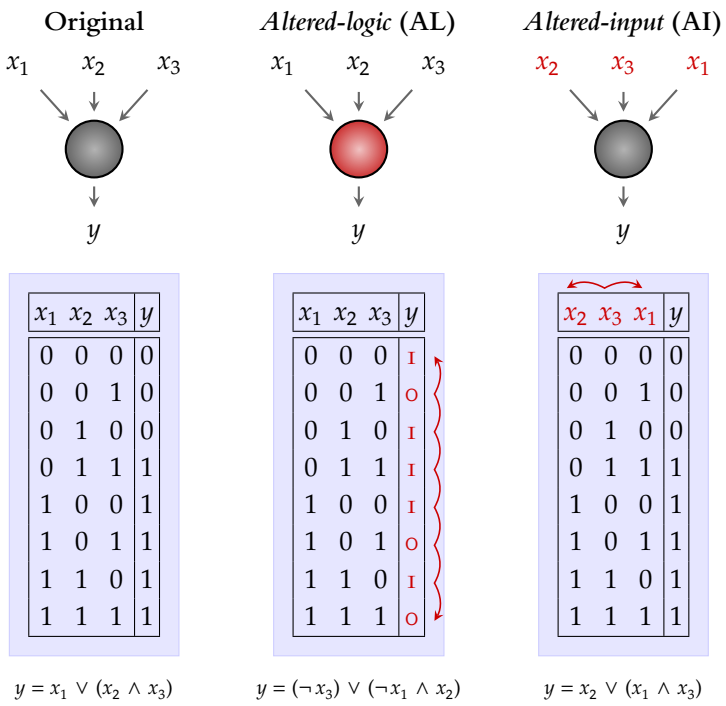

Figure 5.12: **Network randomisation.** The update rule of one of the nodes of the original network is shown as an example (left column). The first randomisation method (altered-logic, middle column) consists on shuffling the output of the update rule, in such a way that the logic of the logic gate changes but the number of active output states is maintained. The second randomisation method (altered-input, right column) consists instead on shuffling the inputs of the network node while maintaining the logic of the update rule (the logic is given below the truth table in each case).

generate this kind of network is the following: for every internal node, i, of the network, we take its original truth table, consisting of 2^{k_i} entries, and count the number of entries that lead to an active state of node i, say $n(i)$. For the new network, we reassign the $n(i)$ entries that lead to the active state for this node. Taking into account that there are $\binom{2^{k_i}}{n(i)}$ different possible logic rules for every node i, we see that the number of possible networks produced in this way is extremely large. These networks are structurally equivalent to the original BN in the sense that the links forming them are maintained. However, the distribution of logic rules is completely different. In particular, the original network contains a high proportion of canalising functions (Helikar et al., 2008; Kauffman et al., 2004). A node rule f_i is said to be canalising for a given input if its output is completely determined by at least one specific value of that input, independently of the values of the remaining inputs. As an example, consider the canalising rule $f(x_1, x_2, x_3) = x_1 \vee (x_2 \wedge x_3)$. Note that, with this logic, whenever x_1 takes the value 1, the output is 1, independently of the truth values of x_2 and x_3. In this case, f is canalising for x_1, and x_1 is a canalising input of f. A total of 96 nodes in the fibroblast network are canalising for at least one of their inputs in the original network. In addition, 168 from the 542 interactions are canalising. For instance, Src protein acts in a canalising manner for 8 of its immediately downstream

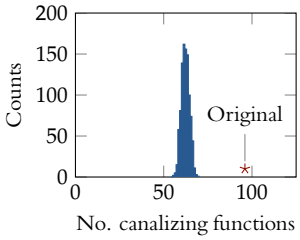

Figure 5.13: **Topological comparison of the networks.** Distribution of the number of canalising functions for the AL networks, as defined in Figure 5.12. The arrow marks the number of canalising functions of both the original and the AI networks (see text below).

proteins while PI4K, Mekk1 and PI3K have 4 canalising inputs. There are no paths from inputs to specific outputs built fully from canalising functions, except the direct link going from ECM to Rac. By contrast, the distribution of canalising functions for the family of AL networks is very narrow and centred on 63 canalising functions (see Figure 5.13).

We built a second type of randomised network by shuffling the inputs of the logic rules of each inner node (altered input (AI) networks, see Figure 5.12). In this case, the procedure for generating the network is the following: for each internal node, i, with k_i inputs, we randomly draw a permutation, $: 1, \dots, k_i \quad 1, \dots, k_i$, defining the new logic rule, f_i, as:

$$f_i(y_1, \dots, y_{k_i}) = f_i \; y_{(1)}, \dots, y_{(k_i)} \tag{5.3}$$

Even though the roles of the input species for every node are changed, this building procedure ensures that the type of logic rule for every node is the same as in the original network. Therefore, the number of canalising functions of the network is preserved. The type of networks produced in this way, while having different functional connectivity than the original network, are topologically very similar to it. Again, taking into account that there are $k_i!$ different logic rules for every node, we see that the number of possible networks produced following this algorithm is very large.

We have simulated and analysed 100 random networks of each type. The results obtained from multiple realisations of the randomised AL version of the network indicate that its response is, in general, not sensitive to the chatter level. Figure 5.14 shows the results for five sampled

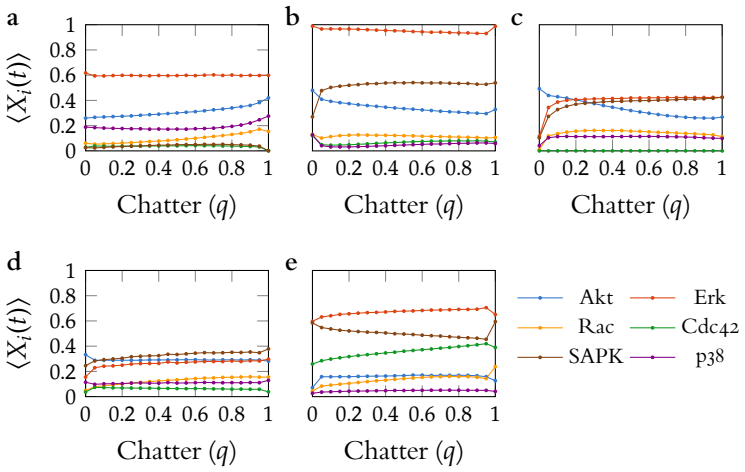

Figure 5.14: **Temporal average of the population activity vs chatter in AL networks.** Examples of five of the 100 randomised altered-logic networks in which the logic rules of the nodes are changed randomly: the network activity is basically insensitive to positive chatter levels. The population activity was estimated from $R = 201$ cell realisations for each network. After dismissing a transitory period of 160 iterations (Rué, Pons, et al., 2010), the temporal average of the population was calculated.

AL networks, in all these cases the population average activities are ap-

proximately constant for varying levels of chatter. Therefore, the responsiveness to chatter is not guaranteed solely by the network topology, but seems to require a particular type of logic rules governing the dynamics of the nodes. To check whether this is indeed the case, we used the

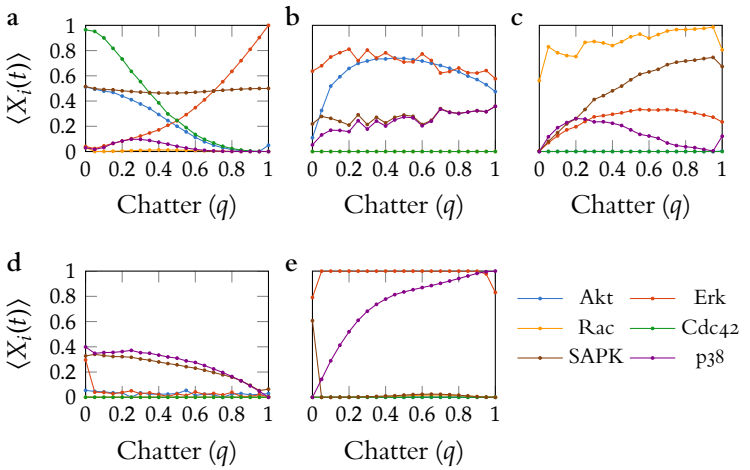

Figure 5.15: **Temporal average of the population activity vs chatter in AI networks.** Five realisations out of 100 of the altered-input network models, in which the logic rules are maintained but the inputs are shuffled randomly. The population activity was estimated from $R = 201$ cell realisations for each network. After dismissing a transitory period of 160 iterations (Rué, Pons, et al., 2010), the temporal average of the population was calculated.

second family of randomised networks, the AI networks. This randomisation is less severe than the previous one, since it maintains the type of logic rules in the network. We observe that networks of this AI family are sensitive to chatter levels in a similar way to the original, experimentally-based network (Figure 5.15).

Figure 5.16 shows the distributions of the number of output nodes that respond to the unstructured inputs as the chatter level varies for AL and AI network families. A node is considered responsive when the dynamic range of the average activity $\langle X_i \rangle$ (difference between maximum and minimum values for varying q, as in Figures 5.14 and 5.15) surpasses a given threshold, taken here equal to 0.25 in the units of the y-axis of Figures 5.14 and 5.15, without loss of generality. Taken together, these results reveal that the biologically realistic network studied here responds in a non-trivial manner to a constant level of background chatter, and it is the distribution of logic rules of the network nodes, which is far from random, that determines this responsiveness.

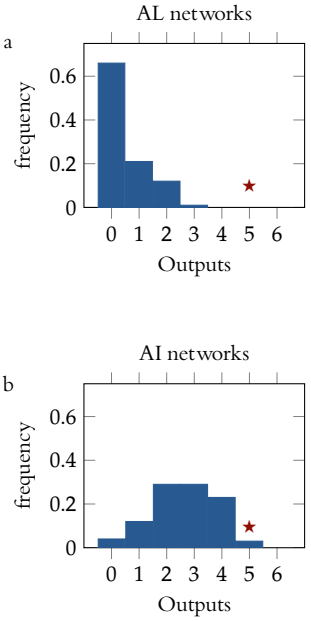

Figure 5.16: **Quantifying the responsiveness of randomised networks to unstructured inputs.** Distribution of the number of output nodes that respond to the unstructured inputs of Figures 5.14 and 5.15, for 100 realisations of each of the two types of network randomisations described in Figure 5.12: AL (a), and AI (b). Red stars indicates the number of responsive inputs of the original biological network.

5.4 *Background chatter enhances the network response to periodic stimulation*

We now study the ability of the network to process and transmit structured information under different levels of background chatter. To this purpose, we examine the response of the network to the periodic stimulation of a specific input node, maintaining the rest of inputs at a constant

chatter level (see examples of realisations and population activity for this type of input in Figure 5.10c,d). Contrary to what is observed for unstructured inputs (Figure 5.10a,b), the output signals obtained in these settings do have temporal structure. This is illustrated in Figure 5.17a, which shows the population average, $X_i(t)$, for all the output nodes upon periodic stimulation of the input node Stress. Some outputs (Erk and Cdc42) do not show a significant response to the periodic modulation of Stress, while others (Akt, Rac, SAPK, and p38) do oscillate at the period of this input. We have studied in detail the periodic stimulation of every input node of the fibroblast network, and find that they all show the same qualitative phenomenology (Domedel-Puig et al., 2011), with the exception of the α-ligands. These ligands differ from the rest of inputs in that they represent generic pathways. For example, α_s ligands correspond to the signals that use the α_s subunit of the G protein (which include epinephrin, glucagon, TSH, and more), while α_i ligands (like acetylcholine, serotonin and angiotensin) use the α_i subunit, etc. They possibly fail to respond because their logic has somehow been altered during the generalisation process. For clarity of presentation, we focus hereafter on the periodic perturbation of Stress only.

We quantify the amount of periodicity in the network's response using, again, the power spectral density (PSD) of each output node (see Section 5.2). Figure 5.17b shows the value of this quantity at the frequency of the input, $\mathrm{PSD}_i(\nu_0)$, as a function of increasing q levels. This plot shows that the response at the input frequency changes in a non-trivial manner as a function of the chatter level (specially for Akt, Rac, SAPK, and p38). The different outputs reproduce the input periodicity in distinct ways, in some cases displaying their maximum response at intermediate q values. The stress-activated protein kinase (SAPK), for instance, seems to respond better to periodic stimulation by input Stress at chatter levels close to $q = 0.3$. Another stress-activated protein, p38, is also an interesting example that presents two ranges of high response for chatters around $q = 0.4$ and 0.9. This behaviour implies that the most responsive output to a given input varies as the chatter level changes. In the particular example shown in Figure 5.17b, the most responsive output under increasing chatter values follows the sequence SAPK, Akt, p38 and Rac. In this sense, the network acts as a system capable of selecting its dominant output depending on the degree of background activity.

We now examine the extent to which the chatter-dependent ability of the network to select its response depends on its connectivity. To

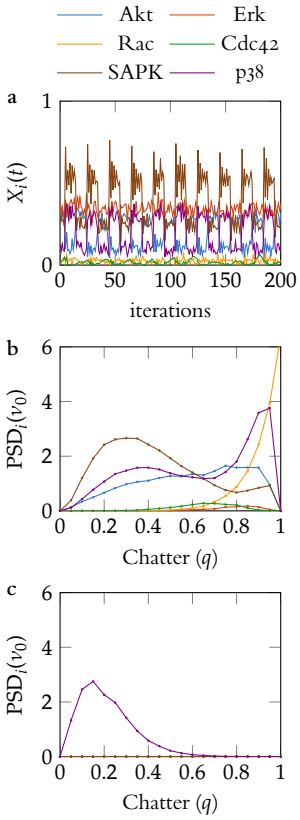

Figure 5.17: **Network response to periodic stimulation of the input node Stress in the presence of chatter.** (a) Temporal evolution of the population average of the output nodes (R=201 cells) for a chatter level of $q = 0.3$ in the original network. Some species (Akt, SAPK and p38) oscillate at the frequency of the input signal. Power spectral density (PSD) at the input frequency obtained from the population average of each output node versus chatter level, q, in the original network (b) and an AI randomised network (c).

address this question, we now perform the same numerical experiment for the two randomisations of the network described previously (and shown in Figure 5.12). As in the case of constant inputs discussed above, AL networks in which the logic at all nodes is randomised (right column in Figure 5.12) are again insensitive to chatter, and in fact they do not respond to the periodic input at all. In contrast to the case of constant inputs, where AI networks did respond somewhat to the levels of chatter (Domedel-Puig et al., 2011), we observe here that this second family of randomised networks is barely able to sense the periodic input, and are thus unable to show a sensitivity of the output to background chatter. A particular example of the response of such weakly randomised network to a periodic modulation of the Stress input is shown in Figure 5.17c. Only p38 responds at all in this case.

Figure 5.18 shows the distribution of output nodes that respond to a modulation of the stress input for the family of AL networks (panel a) and AI networks. From this Figure, it is clear that randomised networks are barely responsive to periodic modulation of its inputs.

This suggests that the topological structure of the network and the distribution of logic rules of the nodes are not sufficient for a successful information processing, but the original logic rules for each node are needed.

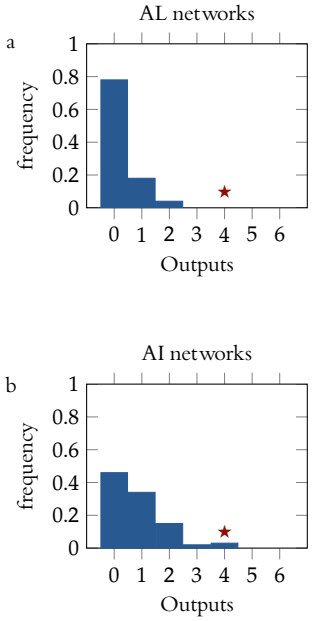

Figure 5.18: **Quantifying the responsiveness of randomised networks to structured inputs.** Distribution of the number of output nodes that respond to a modulation of the stress input, for 100 realisations of each of the two types of network randomisations described in Figure 5.12. A node is considered responsive when the dynamic range of the power spectral density at the stimulation frequency (difference between maximum and minimum values) for varying q surpasses a given threshold, taken here equal to 3 % of the maximum PSD of the input, without loss of generality. The arrows indicate the number of responsive nodes of the original biological network.

In the BN field there is an open discussion with the use of a synchronous updating scheme for the simulation of biologically relevant models, as it has been shown to produce artefacts. Indeed asynchronous updating has been shown to destroy some of the attractors found in networks previously simulated with a synchronous-updating framework (Greil and Drossel, 2005; Klemm and Bornholdt, 2005a). The phenomena described above in this Section have been performed using this synchronous updating approach. Thus, it is fair to ask whether this is an artefact of the simulation scheme or not. To verify that the results obtained are not a consequence of the simulation method, we have implemented different asynchronous updating schemes

In a first asynchronous method, the updating order of the nodes is randomised at each iteration, and all nodes of the network are updated sequentially following this order. In that way, one iteration in this scheme corresponds to updating once the whole network, as in the synchronous scheme. However, this is an extreme change with respect to the synchronous method, in the sense that we switch from a parallel

updating order (synchronous update) to the opposite scenario in which the asynchronous updating sequence is specific to each iteration. To examine an intermediate situation, we have implemented a second asynchronous update method in which every iteration is composed of two steps. During each of these steps, a different partition of nodes is updated synchronously. Nodes are assigned to the early-update partition according to a given probability p, which we set to 0.05. Both asynchronous update schemes described here share the fact that they select the nodes randomly at each iteration. This fact strongly suppresses the ability for the structured ordering coming from the periodic input to systematically percolate spuriously through the network. However, the results (see Figure 5.19) show that the non-trivial influence of chatter on the response to a periodic input persists in the presence of asynchronous updating. The transmission of the input signal in this case can thus be attributed only to the robustness of the dynamical behaviour with respect to these causal destructive updates.

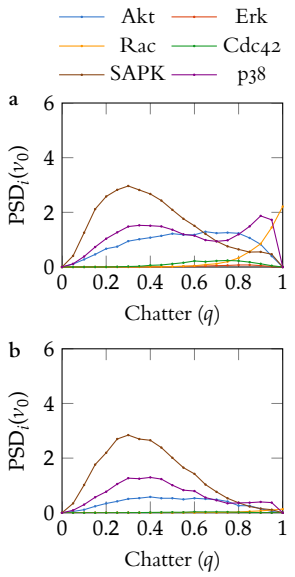

Figure 5.19: **Robustness to asynchronous updating.** Network response to periodic stimulation of input node stress ($\nu_0 = 0.05$) in the presence of chatter, quantified by the power spectral density of the output nodes, for three different asynchronous updating methods (see text). The power spectral densities are calculated for time series of 1600 iterations, after dismissing a transitory period of 160 iterations in the data, and for a population of size R=201 for each q. (a) Results for the two-step partial randomisation update scheme, with a probability of being updated in the first step equal to 0.05. (b) Results for the complete randomisation update scheme.

Heterogeneity introduced by chatter influences the information-transmission capabilities of the network. Our dynamic definition of chatter, however, also implies a dynamic component that cannot be disgreggated from this heterogeneity. Therefore, it is reasonable to ask whether this dynamic nature of the fluctuations is relevant in the information-transmission process or not. Indeed, the temporal variation associated with chatter prevents the system from becoming trapped in (deterministic) attractors in which some nodes might become permanently unresponsive to the input signal (as we have discussed in Section 5.1). Thus the response of the network to chatter goes beyond a mere sensitivity to different average levels of network activity in different nodes. To verify this, we have run simulations in which the chatter is static (fixed to its initial value, *quenched* chatter), and calculate the number of cells (i.e., realisations) that are responsive to the input, comparing it with the case of temporally variable chatter. To compare both situations, we consider the number of individual cells that respond to a given periodic stimulus. We quantify this with the maximum cross-correlation of the average activity between the stimulated input and each output. The cross-correlation between to given nodes i and is defined as:

$$c_i(\ ) = \frac{1}{-} \sum_{t=} X_i(t) - \langle X_i \rangle \quad X(t - \ ) - \langle X \rangle \qquad (5.4)$$

where $\langle\cdot\rangle$ represents the standard time average defined in Equation (5.2). The cross-correlation function $c_i(\ )$ measures how similar the population signals X_i and X are when the second signal presents a temporal delay of with respect to the first one. Although it is known that correlation does not imply causality, if two linked nodes are correlated, it is often reasonable to assume that the source node transmits information to the target node of the link. For each edge of the network, a simple measure of information transmission is, hence, the maximum value of the cross-correlation with respect to the delay ,

$$_i = \max_{0,\ldots,\ _0} |c_i(\ )| \tag{5.5}$$

where $_0$ is the period length and the absolute value in the formula accounts for phase inversions. A cell is thus considered responsive if the maximum value of the cross-correlation function between the input signal and a certain output response is larger than a given threshold (chosen here equal to 6.25 % of the maximum correlation, without loss of generality). Our results, displayed in Figure 5.20, show that the fraction of responsive cells for a given input is substantially smaller in the case of quenched chatter (panel b) than for dynamic chatter (panel a), for all chatter levels. Thus temporal variability of the chatter plays a relevant role in the behaviour reported.

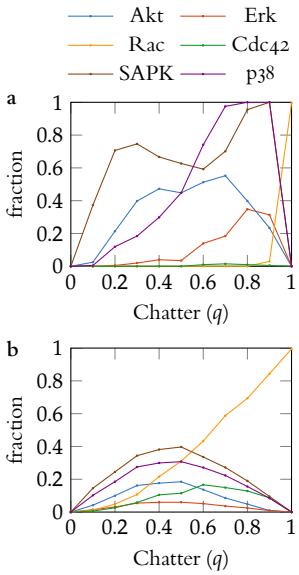

Figure 5.20: **Role of temporal variation of the chatter.** Fraction of responsive cells (realisations) versus chatter level q for (a) 201 realisations of temporally varying chatter, as considered in the main text, and (b) 2001 realisations of quenched chatter. A cell is considered responsive if the maximum value of the cross-correlation function between input (stress) and output surpasses 6.25 % of the perfect score.

Next we study how the information is transmitted from the stimulated input to the dominant output, and which nodes participate in this transmission. To that end, we calculate the PSD at the stimulation period, ν_0^{-1}, for all network nodes. The network nodes that present a clear peak centred at the stimulation frequency in the power spectrum are considered to be responsive. We also compute the maximum cross-correlation, in absolute value, between the average signals of all pairs of interacting nodes i and (Equations (5.4) and (5.5)). This measure is dependent on the input stimulus and the global chatter level. Figure 5.21 (page 116) shows this information in the case of a periodic modulation of the input Stress, and for two different values of the chatter level, $q = 0.3$ and $q = 0.9$, which correspond to the conditions for which SAPK and p38, respectively, are dominant outputs (Figure 5.17b). A common feature of both panels in Figure 5.21 is that there are some internal nodes that reproduce the periodic input signal (i.e., they have high values of the power spectral density at the input frequency), and which are usually connected to each other by high cross-correlation values. However, there are also

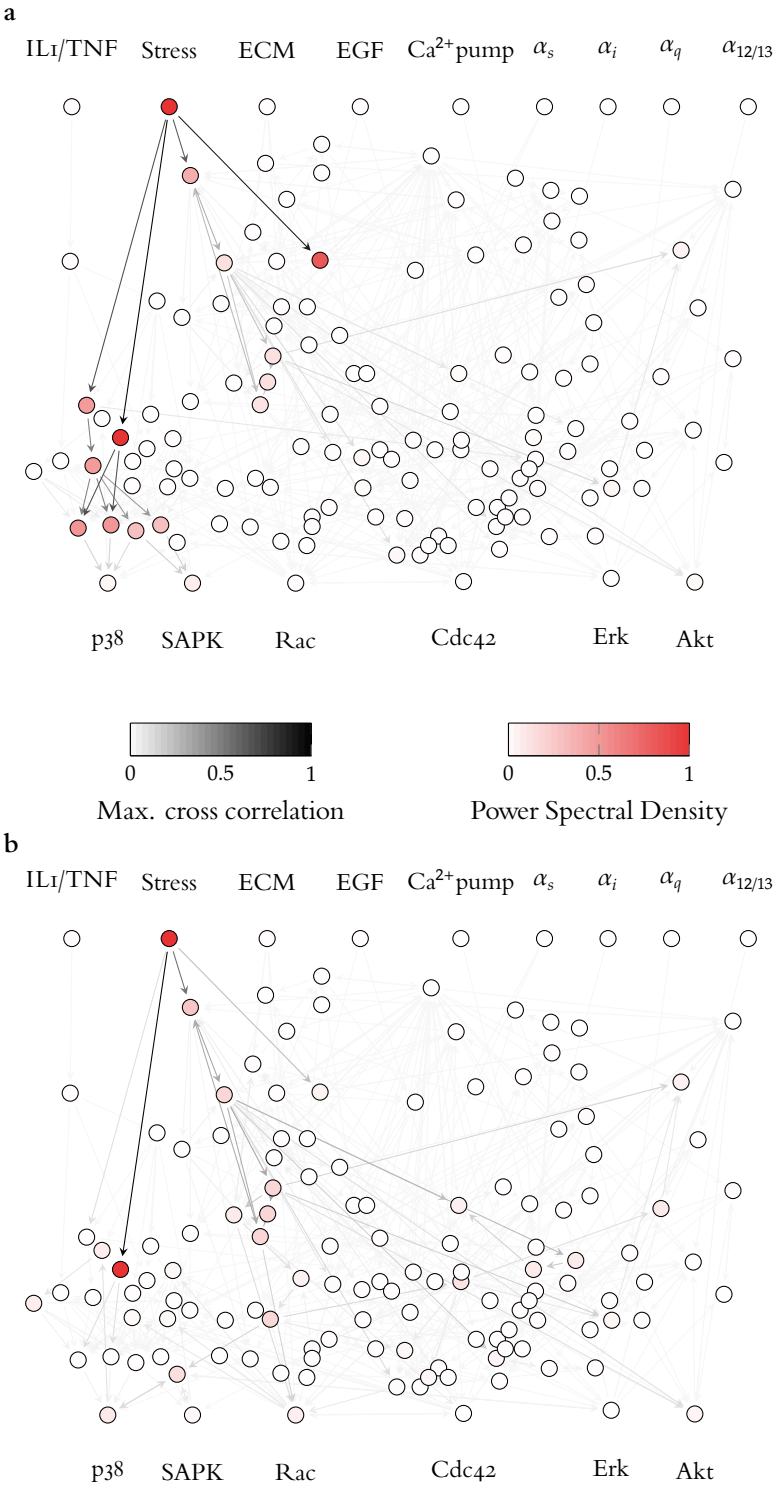

Figure 5.21: **Chatter level determines the set of nodes that respond to the input modulation.** Power spectral density at the input frequency (colour coding of the circles) and maximum cross-correlation (in absolute value) between pairs of nodes (grey coding of the edges) for a periodic modulation of the Stress input and two values of the background chatter level affecting the rest of inputs: (a) $q = 0.3$, (b) $q = 0.9$.

important differences between the two chatter levels. For instance, when $q = 0.3$ (Figure 5.21a), most of the nodes that transmit the signal from the stimulated node to the dominant output node (in this case, SAPK) are not so active for $q = 0.9$ (Figure 5.21b), and a different set of nodes transmit the information from the Stress input to p38, Rac and Akt, which now become dominant outputs. Together, these results show that chatter is able to select which output responds dominantly to a given input by determining the set of internal nodes that are most affected by the input. These nodes in turn signal downstream until a given output node is reached.

5.5 *Chatter level determines the information path*

Figure 5.21 shows that the chatter level sets which groups of internal interconnected nodes convey information from the inputs to the outputs. These groups of nodes and the links between them constitute preferred paths of information transmission. In order to characterise which of these paths are dominant in transmitting information, we resorted to optimisation algorithms of graph theory. We built a directed weighted graph using a measure based on this assignment. The idea is to determine the paths going from each input node to each output node in such a way that each edge forming this path has a high enough maximum cross-correlation. Hence, every path going from a node i to a node is formed by a set of edges, $(i_1, i_2), (i_2, i_3), \ldots, (i_{-1}, i)$, in such a way that $i_1 = i$ and $i =$. The criterion we chose was to assign a weight equal to the inverse of maximum cross correlation, $_i = {}_i^{-1}$, to each edge $(i,)$, and to look for paths that minimise the sum of these weights. To do so, we defined a correlation score or cost function, , for each path, as the inverse of the sum of weights along it:

$$((i_1, i_2), (i_2, i_3), \ldots, (i_{-1}, i)) = \left(\sum_{k=1}^{-1} {}_{i_k, i_k+1} \right)^{-1} \qquad (5.6)$$

To obtain the paths with maximum correlation score, we used a fast implementation of the Kth shortest path algorithm (Martins and Pascoal, 2003), which identifies the first K paths with minimum sum of weights, and thus the maximum path score (Kth-shortest path algorithm with $K = 10$ shortest paths). This approach is well suited for our problem because it penalises large paths, in which case more terms are added to the path score, and paths where at least one of the edges has a low

maximum cross-correlation value, $_{i_k,i_{k+1}}$. We set the number of top shortest paths to be found, K, to a high enough value, such as $K = 10$, which ensures that all paths with high scores are identified. After finding the highest score paths for all q values, a threshold-based filtering is set to find those paths with scores higher than a certain level, $_c$.

Each of the paths found using this method was assigned a score equal to the inverse of the sum of weights. Thus, the higher the score of a path, the higher the correlations of its constituent interactions. Those paths with highest score in terms of sums of these weights are what we define as dominant paths. In Figure 5.22, we show the score of the best paths found going from input Stress to p38 (panel a). In panel b we show the value of the score as a function of the chatter level q, for each of the paths found (for results corresponding to other input-output combinations see Domedel-Puig et al., 2011). The detected paths (Figure 5.22a) are

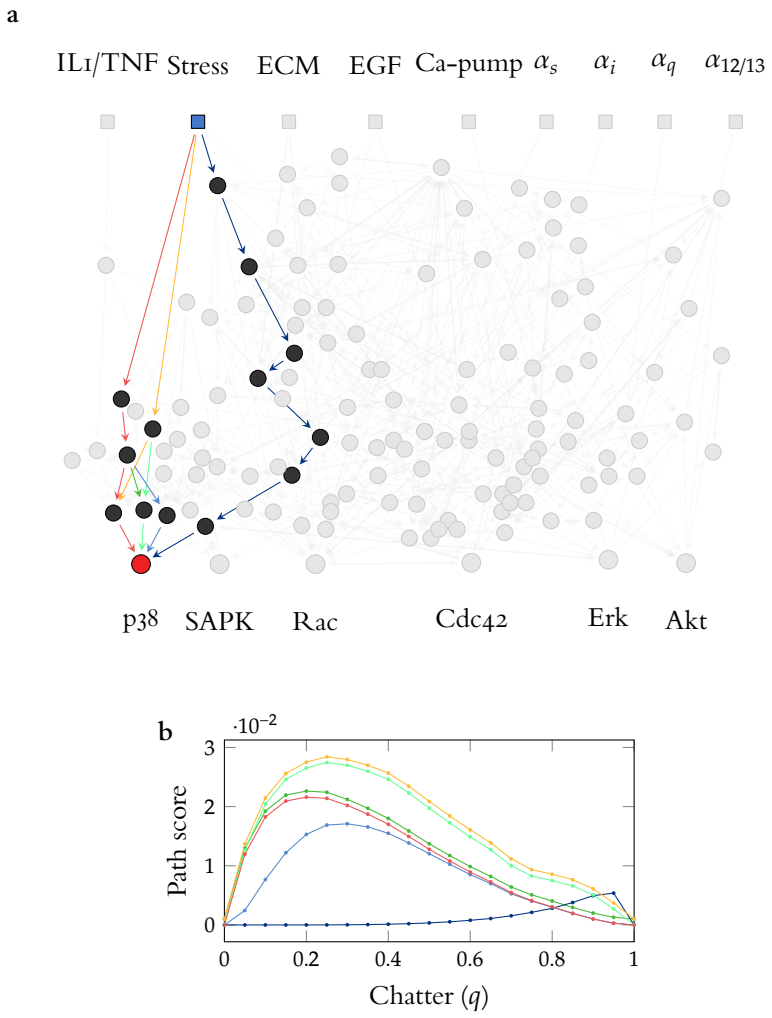

Figure 5.22: **Paths of information flow for varying levels of chatter.** (a) Detected dominant information transmission paths (see text). (b) Path scores of the dominant paths versus chatter level for a periodic modulation of the stress input, when considering p38 as output.

short, as they usually involve between 3 and 7 intermediate species. For

low chatter levels, a group of paths emerges involving the MKK3 and MKK6 activation of p38. This group of paths is responsible for the first peak in the power spectral density of p38 shown in Figure 5.17b. At high levels of chatter, these paths fade out, and the oscillatory behaviour of p38 (second peak in Figure 5.17b) becomes then due to inhibition by the MAP phosphatases (MKP), which in turn are activated through the adenyl cyclase (AC-cAMP) pathway.

5.6 *The network presents a balance between robustness and responsiveness*

From the results of the previous Section it is clear that different dominant paths emerge as a consequence of varying chatter levels. The remaining question is what happens to those internal nodes of the network not involved in the aforementioned paths when chatter level varies. To address whether they change their processing capacity, we analyse the sensitivity to chatter variation of the power spectral density at the input frequency, $\mathrm{PSD}_i(\nu_0)$, and of the maximum cross-correlation (in absolute value) between edges, $_i$.

As in many dynamic systems, a measure of how sensitive the response of the system is to a certain parameter can be introduced for our model. In particular, we can define measures of sensitivity to chatter level for both nodes and edges of the network. We define the sensitivity of nodes to chatter as the variations of power spectral density at the input frequency with respect to variations in the chatter. In mathematical terms, the sensitivity of node i can be defined as $S_i(q) = \frac{}{q}\,\mathrm{PSD}_i(\nu_o)|_q$ Likewise, we define the sensitivity of edges to chatter variations as $S_i\,(q) = \frac{}{q}\;{}_i\,(q)|_q$ Both derivatives are estimated by centred differences with a differentiation step of $q = 0.05$.

We call these two magnitudes, respectively, node sensitivity [$S_i(q)$ for node i] and edge sensitivity [$S_i\,(q)$ for the interaction pair $(i,\ )$] (see Appendix A.3 for additional details).

Figure 5.23 summarises the results obtained in the case when the network is driven by an oscillating stress signal (see Appendix A.3 for other inputs). Both nodes and edges are coloured according to the maximum (in absolute value) sensitivity for all chatter levels. Blue (red) colour indicates a positive (negative) variation in the direction of increasing chatter. Colour intensity indicates the magnitude of this maximum variation. In this Figure, it can be observed that only a small number of nodes and

edges of the network have a significant variation of power and correlation when varying chatter levels. Note that most of the sensitive nodes and edges are involved in one or more dominant paths at a given chatter level (see Figure 5.21 and Figure 5.22). Thus, while species belonging to paths involved in information transmission are sensitive to variations in the chatter level,the rest of the network nodes seem to be robust against these variations.

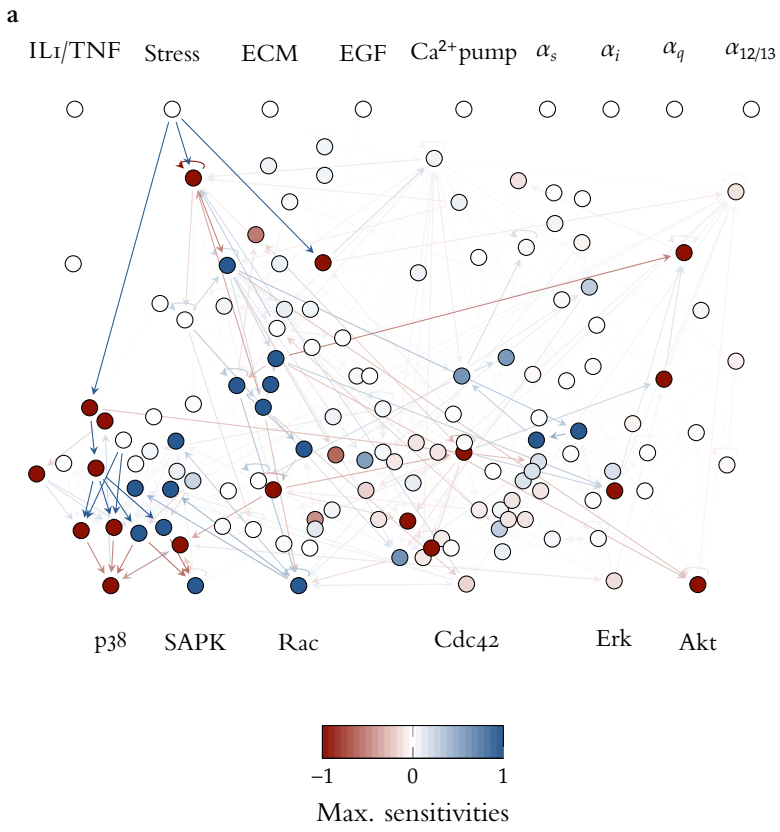

Figure 5.23: **Maximum sensitivity of nodes and edges to chatter variation for periodic input stress.** Red (blue) colour scales for negative (positive) values of the sensitivity. Only a small number of nodes and edges present a high maximum (in absolute value) sensitivity. Most of these nodes and edges are part of at least one of the paths previously identified as dominant (see Figure 5.21 and Figure 5.22). The major part of nodes and edges not involved in information paths are robust to chatter, since they are insensitive to it.

5.7 *Conclusions and discussion*

Signalling is a key cellular process. Only with a complex and efficient flow of information across their biochemical machinery are cells able to react to changing environments and to fulfil the basic operations needed to maintain life. This highly complex behaviour cannot be built under the basis of isolated specialised processes, but must rely on a certain level of global coherence. Even though a large amount of studies in the past have been devoted to examine specific mechanisms in the signalling system of the cell, little attention has been paid until the last decade to the integration of the different parts of the signalling machinery.

Here we are interested in this global aspect of the signalling mechanisms. In order to handle the complexity of the signalling machinery, we considered a reduced version of the system by decreasing the number of details while still capturing the main properties of the process. Boolean networks provide us with a mathematical framework that, even though being highly simple and schematic, keeps the essence of the information processing underling signalling activity. Within that context, we have used a recently developed Boolean network (Helikar et al., 2008) that models the signalling network in human fibroblast cells, and that has been compiled from an extensive experimental literature analysis. We have used this network to study the global response of the system to external inputs of different nature. Each type of input considered introduces information into the system at different scales. We have computed the response of a single cell —understood in our model as a single realisation of the dynamics— and of cell populations — understood as the average of multiple realisations sharing equivalent input conditions.

It is well known that random Boolean networks evolve, in the absence of noise, towards steady or periodic attractors. We observed this type of behaviour in our (non-random, experimentally motivated) *original* network model when the input states of the network are fixed. The network evolves from different initial conditions (which may correspond to different cellular states) towards specific attractors by relaxing its dynamics during a transient period. We computed the duration of these transients and observed that their distribution peaks at around 10 to 20 iteration units. Thus, this signalling network reacts to different kinds of fixed inputs within the same time scale. This characteristic time of the dynamical response of the network is approximately constant, which means that the network goes from any region in phase space to its corresponding attractor in a narrow range of iterations. This behaviour is a manifestation of the responsiveness and reliability of the network. When considered as a population by averaging the evolution of all the cellular ensemble, the system behaves similarly.

Purely constant inputs are, nevertheless, not realistic biologically. Cells work in an environment that changes in time due to biochemical noise, variations of control conditions, or even coupling with other cells in multicellular organisms. We have addressed this situation by including a random signal with rapid variability, the chatter, to the network inputs. Noise introduces a fast time scale into the system that is transduced down the network, and which might interfere with its information processing

abilities, specially taking into account its high connectivity. This does not occur, however. The only effect of the input chatter is the elimination of the deterministic attractors of the network, while the transient lengths are kept essentially the same as in the case of fixed inputs.

In this scenario, where the presence of noise does not destroy the system's responsiveness, we ask how the network processes a well-defined external signal at different frequencies. Our results show that, at the population level, the network is able to process input signals with frequencies in the range of the network's relaxation frequency. This fact has clear biological implications, as it establishes that a particular biological system, such a transduction network, has a limited dynamic range for which it is able to process input signals. A recent experimental study has revealed that the NF-κB pathway responds in a non-trivial manner to the frequency of periodic stimulation by TNF-α (Ashall et al., 2009). Even though the components of that pathway are not included in our network (the input signal is) and the pathway considered is small and intrinsically dynamical, the modelling results presented here could provide a framework for understanding how the frequency response of signalling networks is related with their relaxation properties. Also, this phenomenon partially explains why the network reaches a stationary state when only noise is present in the inputs, since we observe that high frequencies are always filtered out. Furthermore, in some cases noise is required for the network to transduce the oscillating input at the allowed range of frequencies.

The range of non-trivial frequency responses reported here —low-pass, band-pass, and high-pass— implies that the signalling network being considered does not behave simply as a linear cascade, since in that case, the information would be simply propagated with a certain delay, of the order of the cascade length. The average shortest path length from inputs to outputs in the network of Figure 1.10 is 3.87, which is much shorter than the median transient lengths reported here. Thus, the shortest path lengths alone cannot account for the relaxation. One could also argue that, because the initial conditions are random, the relaxation duration is governed by the maximum shortest path length. However, in this network the maximum shortest path has 12 links, and thus cannot account for the non-negligible amount of transients spanning tens of iterations. Taken together, these facts suggest that the information in this signalling network is not simply percolated through a linear cascade, but reverberates inside the network for a certain time, due to long feedbacks.

Indeed, the network used here is not directional. For instance, it contains 50 auto-loops which roughly correspond to 10 % of all the links, 26 direct feedback loops (A B A, 10 % of links), and 55 % more two-step feedback loops (A B C A) than a corresponding random network.

Our results show that signalling networks that integrate multiple dynamic inputs exhibit a well-defined relaxation behaviour that is robust to the presence of noise —in the form of background chatter—, while being flexible enough to provide the cell with a rich toolbox of frequency responses to periodic input signals. Indeed, natural cells live in environments whose composition affects the way in which they function. An example is the interstitial fluid surrounding cells in higher organisms, which affects processes as important as embryogenesis, tissue morphogenesis, remodelling and cancer progression (Rutkowski and Swartz, 2007). While being subject to purely external sources of variation, cells also contribute to modifying their surroundings by secreting multiple signalling molecules (Karagiannis, Pavlou, and Diamandis, 2010). In a given physiological situation only a small subset of those signals will carry information relevant to the cell (Hsueh et al., 2009), with the rest constituting a source of background noise, or chatter, that is bound to affect the cell's response to the relevant inputs. The level of chatter will depend on numerous variables, including cell type, tissue, developmental stage, health status, etc. The information-carrying signals, on the other hand, are frequently dynamical, since oscillations in cell physiology are ubiquitous (Stark, Chan, and George, 2007), and in some cases markedly periodic, driven for instance by regular biological rhythms such as those generated by circadian clocks (Keller et al., 2009). In this Chapter we have also examined how the information transmission capabilities of a human cell subject to a periodic input signal depend on the amount of background activity.

At the experimental level, context-dependent signalling is beginning to be unravelled. For example, the Alliance for Cell Signalling (AfCS) recently compared the effect of 22 individually applied inputs (cytokines, GPCR ligands, TLR ligands, and tyrosine-kinase receptor ligands) upon 42 cell outputs (cytokine production, protein phosphorylation, calcium and cAMP levels), to the effect of all possible pairwise combinations of those inputs (Natarajan et al., 2006). According to the results of that study, only a small number of ligands are able to control cellular outputs independently from the other inputs. In contrast, most inputs act as

modulators of signal transduction, providing the cell with the ability to perform context-dependent signalling. Our results fit well with these findings, as we see that the level of background activity of the input nodes determines the capability of the cell to respond to other inputs (in particular, to follow both unstructured and structured signals). In their work, Natarajan et al. (2006) coin the term "interaction agent" to refer to the network circuits that couple different signalling pathways. They claim that such circuits would be silent in single ligand experiments and become active upon multiple input signalling, causing the non-additive effects observed for certain pairs of inputs. In our theoretical study we effectively observe different areas of the network being used at specific chatter levels, thus supporting the existence of these circuits.

Our results show specifically that a detailed signalling network, carefully compiled from published experimental data (Helikar et al., 2008), responds in a non-trivial manner to background chatter, both intrinsically and in the presence of a periodic modulation of one of the inputs.

This work presented here extends the findings of Helikar et al. and colleagues, who created the network in the first place and studied its stationary response to different input levels for increasing intensities of noise. They concluded the network divides biological stimuli into categories, since it reduces the full range of possible external inputs to a limited number of cell responses, in a manner that is robust to noise.

We divert from the work of Helikar et al. (ibid.) in that we focus on the dynamics of the system. Having explored the relaxation time and frequency response of the network in the previous Chapter, we now have shown that chatter is able to enhance the response of certain outputs to a given input when tuned to optimal levels. Given that chatter controls the amount of stochasticity acting upon the network, this is a situation reminiscent of stochastic resonance, a phenomenon in some physical and neurosensory systems by which the detection of a weak signal is enhanced by noise (Wiesenfeld and Moss, 1995). Recent studies have shown that signalling networks prioritise dynamic range over signal strength (Janes, Reinhardt, and Yaffe, 2008). This entails a linear relationship between the input signal and the output response of the network, which ensures that the reaction of the network to an oscillatory input will also be oscillatory with the same main frequency, for a large wide range of input amplitudes. Our results corroborate this finding, and extend it by assigning a relevant role to the background chatter coming from other input nodes, which enhances the frequency response.

Our results also show that varying chatter levels allow the network to select which output nodes respond preferentially to a given input. Indeed, output switching is achieved via a mechanism that places few requirements on the temporal structure of contextual, non-specific signals. Thus, we conjecture that cells could use environmental noise —to which they are unavoidably subject— to choose among alternative information routes, and eventually among different cellular responses. Randomised versions of the original network in which the topology of the connections is maintained —and only the integration rules at the nodes are altered— fail to reproduce this property, indicating that the chatter-driven selectivity reported here is fine-tuned to the specific architecture and logic of the experimentally-supported network.

Concurrent with the ability of background chatter to select the dominant output for a given input, chatter also selects the network path through which information is transmitted. The nodes belonging to these preferred paths can be expected to form the classifier hyperspace proposed by Oda and Kitano (2006) in their study of the Toll-like receptor signalling pathway. These preferred paths are sensitive to chatter and allow transitions between different information processing scenarios that underlie different output responses. In that way, a given signalling network can have multiple working states that are selected by the background chatter. The rest of the network nodes not belonging to the preferred paths, on the other hand, remain insensitive to chatter. In that way, we can conjecture that signalling networks have a built-in balance between responsiveness and robustness within their coupling architecture, and this balance is modulated by background chatter. The results presented here indicate that background activity levels are key for determining the response of the cell to a given input, by allowing the emergence of novel system-level properties such as information routing, output switching, and context-dependent signalling.

6

Integration of cellular signals in chattering environments

In the context of cellular signalling, it is reasonable to think that environmental noise might be used by the cell in a constructive manner to detect and amplify information-carrying signals. Indeed, theoretical evidence in this direction has been presented in Chapter 5 above. However, the implication of this type of mechanism on signal integration processes is still an open question. In fact, despite the effort made during decades, signal integration is far from being fully understood. One of the reasons behind this problem is that a detailed theoretical description of a cell is proving difficult, because cells are complex objects. The chemistry of cells operates at multiple levels, spanning for instance signalling interactions, genetic expression, and metabolism, and each of these levels of operation is mediated by many chemical species that interact in a strongly non-linear manner, forming highly complex dynamical webs.

In this Chapter, we discuss work that moves away from classical signalling studies that focus on only one ligand, and examine instead the effects of multiple ligands upon the same cell, in the presence of external noise. In particular, we consider whether the ability of cells to integrate signals from one or more sources of information is altered when other input signals provide a background chatter. To study this problem, we consider the same dynamic description of the signalling network described in Chapter 5 above. (Helikar et al., 2008). Using this model we explore systematically, by means of numerical simulations, the effect of background noise in the integration of input signals.

Altogether, the results presented in Chapter 5 reveal that the dynamical response of the Boolean signalling network to external modulation is, at the same time, versatile and robust to noise. The rich dynamical be-

haviour observed in this model, consistent with the diversity of responses observed in cellular organisms, is the result of the complex information-processing ability of the network. The capacity to combine information from different input sources to produce a coherent output response is possible only if the different signalling pathways can interact within the network. In fact, Natarajan et al. (2006) introduced the term "interaction agent" to refer to the network circuits that couple different signalling pathways. In this Chapter, we are interested in understanding how noise affects pathway interaction and thus signal integration. In Section 6.1 we describe the effect of background chatter on the network processing of a single external signal. In Section 6.2 we extend the study to two the integration of two simultaneous input signals, and in Section 6.1 we compare the network integration function to that of simple two-input logic functions. This comparison allows us to uncover the effect of global background chatter on the network integration function of the external signals.

6.1 Integration of a signal with background chatter

The ability of chatter to select the information route going from a periodically driven input node to the output nodes, described in Chapter 5 above, stresses that signalling networks are not a passive objects that merely transmit information, but that they also process this information, modifying it, on its way from the emitting (input) node to the receiver (output) node. This phenomenology leads us to ask if the background chatter level q (defined in Chapter 5 above) modifies the activity of the network not only when one node is perturbed periodically, but also when this node is maintained at a specific (different) chatter level q_1. In other words, we ask if the routing ability of chatter revealed in the dynamical context of one periodic input also exists in a more static situation, when that node is maintained at a specific chatter level q_1, different from the fixed background chatter level q that affects all other input nodes.

In that direction, we systematically studied the average activity of the output nodes when changing the specific chatter level, q_1, of a given input node, under different values of the global background chatter, q. Specifically, we scanned the model in two dimensions for each input i: for q_i ranging from 0 to 1 —in 0.10 steps—, and a common q equal for the remaining 8 inputs, also ranging from 0 to 1. We compute the average activity of the output nodes at steady state over 2000 different realisa-

tions. These different runs involve different realisations of the stochastic input variables as well as different random initial conditions. Following the approach presented in Chapter 5 on the transient behaviour of the network in the presence of chatter, all temporal averages are computed on a time window of 200 iterations and after disregarding the initial 100.

The quantitative behaviour of the network outputs is specific for each of the considered inputs. Furthermore, in most cases the dependence of the activity of output nodes on q_i changes when the common background chatter q is varied. This can be seen for instance in Figure 6.1, which for the sake of simplicity shows only the case when the chatter of the input node IL1/TNF-α, q_1, is modified. These results confirm the fact that the processing of one signal depends on the rest of input signals.

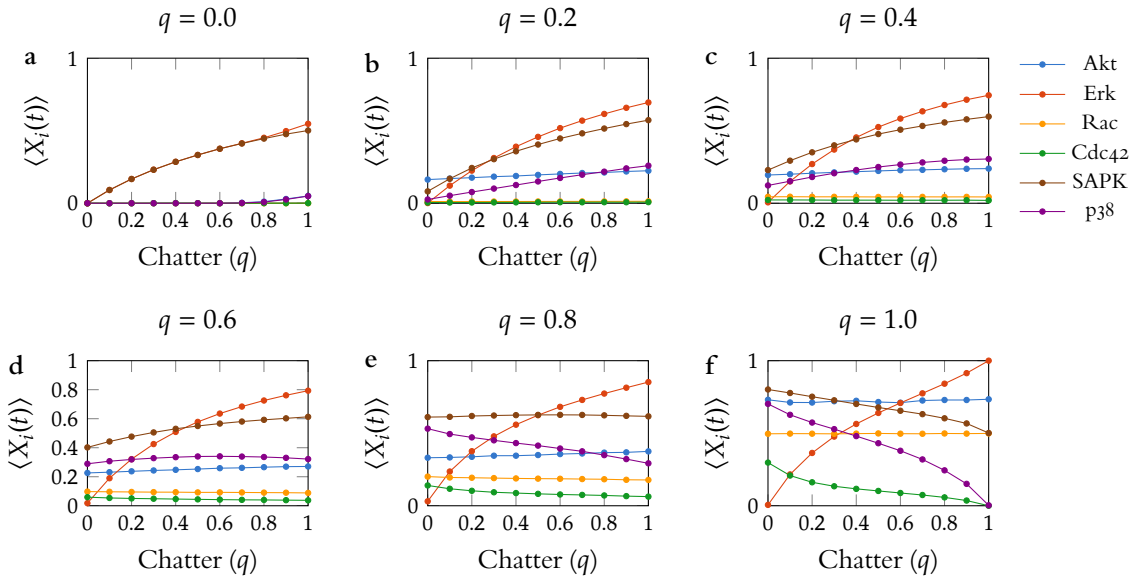

Figure 6.1: **Average activity of the output nodes changes with chatter level.** Average activity of the output nodes versus the chatter level, q_1, of the input node IL1/TNF-α for different values of the background chatter $q = 0.0, 0.3, 0.5, 0.7, 1.0$. Averages are computed over 2000 cell realisations lasting 300 time iterations each, after dismissing a transient of 100 iterations.

Different output nodes in Figure 6.1 display different behaviours: for instance, the monotonic increase of Erk with q_1 is basically unchanged when increasing the global background activity q. The opposite situation is shown for SAPK and p38, which switch their monotonically increasing dependence with q_1 at lower q to a monotonically decreasing dependence at high q. These two examples highlight the dramatic influence of the background signalling upon the processing of one specific ligand. In this case, a low activity of the background signals would enhance the activity of SAPK and p38 when IL1/TNF-α increases, but would inhibit their output activity with IL1/TNF-α when the background chatter is larger. The average activities of Akt, Rac and Cdc42 are independent of the degree of activity of IL1/TNF-α, but grow when the global back-

ground chatter level increases. Thus, these output nodes are not sensitive to IL1/TNF-α but they are activated by the background chatter. Qualitatively similar changes with q are observed for other inputs.

6.2 Integration of two signals under background chatter

The results of Figure 6.1 support the idea that the integration of one single input cannot be studied in isolation from the activity of the other inputs received by the cell. These results are consistent with the dependence of the average activity of some output nodes on background chatter level, and with the effect on the integration of periodic signals shown in the previous Chapter. Continuing in this direction, we now examine the effect of background chatter when two different input nodes are activated with two different and specific chatter levels, q_1 and q_2, respectively. To that end, we perform a comprehensive set of numerical simulations combining all the possible pairs of inputs, with the values for the q common to the remaining seven inputs ranging from 0 to 1, and averaging over 2000 different initial conditions and realisations of the random variables. Four examples of the outcome of these simulations are presented in Figure 6.2 (page 131), which shows the average activity (in colour code) of the different output nodes when integrating the levels of various pairs of input nodes, for increasing background chatter levels, as indicated. The output activity surfaces obtained in this way, or integration maps, change as the background chatter level increases. In particular, in the case of IL1/TNF-α and EGF as inputs (Figure 6.2a-f), both inputs have to be ON at low background chatter to activate p38, whereas at high background chatter EGF has to be ON and IL1/TNF-α has to be OFF for p38 to be turned on. In other cases the network goes from being sensitive only to input q_1 (Figure 6.2g,s) to integrate the two inputs in different ways (Figure 6.2l,x), as the background chatter increases. The network can also transition from being completely insensitive to both inputs (Figure 6.2m) to develop an integration capability in the presence of noise (Figure 6.2n-r). Other combinations of input pairs and outputs exhibit non-trivial changes in their integration maps as chatter varies.

6.3 The network integration logic changes with chatter

The transitions observed in Figure 6.2 are relevant for cellular behaviour. Their existence suggests that the net effect of the network is to act as a switch mechanism which selects the type of q_1/q_2 integration as a func-

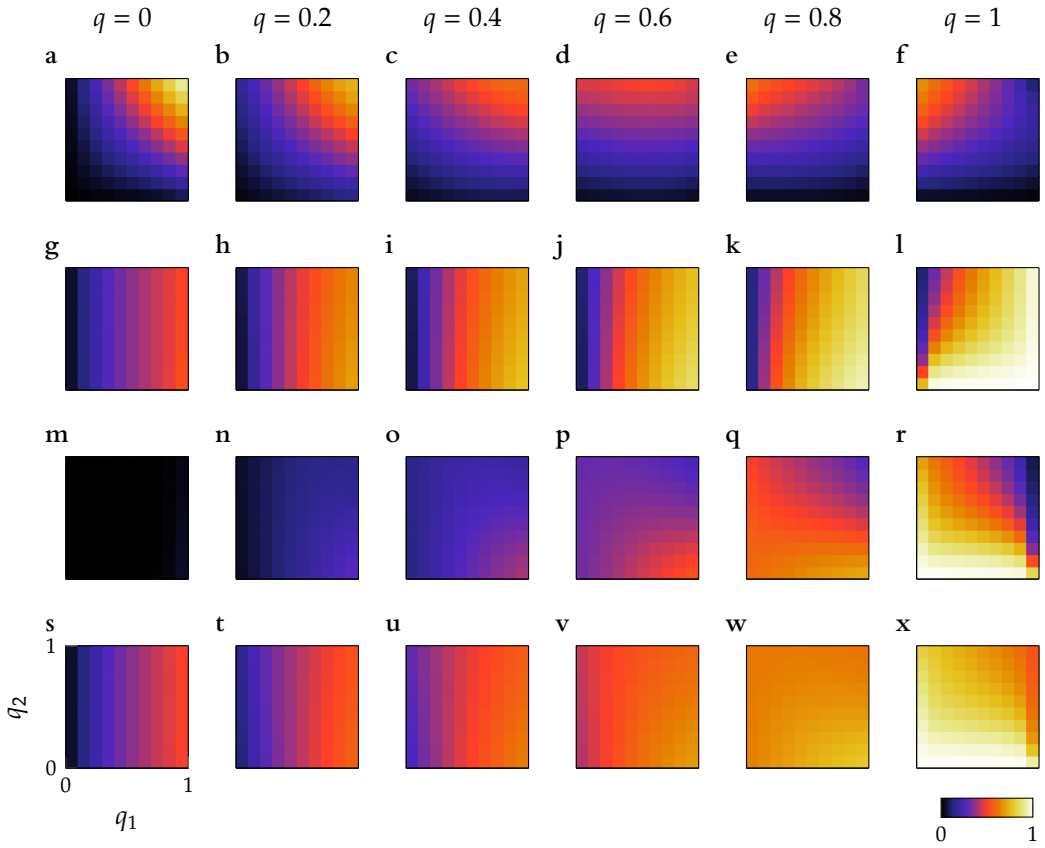

Figure 6.2: **Two-input integration maps.** (a) Integration maps for four combinations of two inputs and an output for varying background chatter q. Averages are computed, for each pair of q_1, q_2 values, over 2000 cell realisations of duration 300 iterations, and after dismissing an initial transient of 100 iterations. The first input, measured by q_1, is in all cases IL1/TNF-α. The second input (q_2) and the output are, respectively: (a-f) EGF and p38, (g-l) ECM and Erk, (m-r) ECM and p38, and (s-x) ECM and SAPK.

tion of the different background chatter levels. This behaviour may be interpreted in terms of simple logic gates with two inputs. Figure 6.3 (page 132) shows the integration maps corresponding to all possible two-input logic gates: two of these (invariant ON and invariant OFF) are independent of the inputs, four of them are sensitive to only one of the inputs ('x', 'NOT x', 'y' and 'NOT y'), and the rest (ten gates) integrate the information from both inputs in order to produce the output. These integration maps can be represented analytically, assuming that the gates are supplied with two independent and binary random processes, x_1 and x_2, with average values q_1 and q_2. The result of this calculation is given in Table 6.1. If we compare the surfaces shown, for instance, in Figure 6.2a,f to the aforementioned 16 basic gates, we observe that those two limit cases bear some resemblance to the 'x AND y' and 'y NIMP x' gates. Hence, despite the complexity of the network under study, we observe that some output nodes behave as functions of simple two-input logic gates. Importantly, the level of background chatter selects the effective logic gate that will operate at any given condition.

To compare the input-output transfer function of the network at dif-

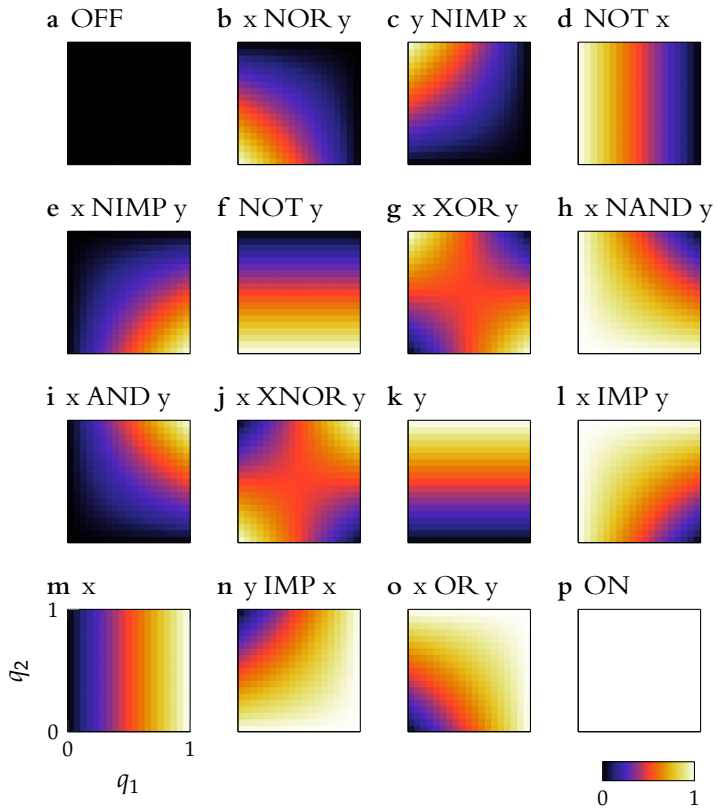

Figure 6.3: **Integration maps for the 16 different rules with two inputs.** The names of the logic gates are based on Mayo et al. (2006). Colour coding is the same as in Figure 6.2.

Rule	Expr. ($_i$)	Prob. (p)	Name
0	0	0	OFF
1	$\neg(x \vee y)$	$1 - q_1 - q_2 + q_1q_2$	x NOR y
2	$\neg x \wedge y$	$q_2 - q_1q_2$	y NIMP x
3	$\neg x$	$1 - q_1$	NOT x
4	$x \wedge \neg y$	$q_1 - q_1q_2$	x NIMP y
5	$\neg y$	$1 - q_2$	NOT y
6	$x \wedge \neg y \vee \neg x \wedge y$	$q_1 + q_2 - 2q_1q_2$	x XOR y
7	$\neg(x \wedge y)$	$1 - q_1q_2$	x NAND y
	$x \wedge y$	q_1q_2	x AND y
	$\neg(x \wedge \neg y \vee \neg x \wedge y)$	$1 - q_1 - q_2 + 2q_1q_2$	x XNOR y
10	y	q_2	y
11	$\neg x \vee y$	$1 - q_1 + q_1q_2$	x IMP y
12	x	q_1	x
13	$x \vee \neg y$	$1 - q_2 + q_1q_2$	y IMP x
14	$x \vee y$	$q_1 + q_2 - q_1q_2$	x OR y
15	1	1	ON

Table 6.1: **Definition of the 16 integration rules with two inputs shown in Figure 6.3.**

ferent background chatter levels, $f(q_1, q_2; q)$, with the simple integration patterns given by the two-input logic functions, (q_1, q_2) [see Table 6.1 in page 132], we use the mean absolute difference (MAD) measure, $\mu_i(q)$, defined as:

$$\mu_i(q) = \quad f(q_1, q_2; q) - \ _i(q_1, q_2) \ dq_1 dq_2, \tag{6.1}$$

where i runs over the 16 integration maps of all possible two-input logic gates.

We consider all the combinations of pairs of input nodes, under the full range of background chatter values. A diversity of behaviours is observed: in some cases, the behaviour obtained from the dynamical evolution of the network cannot be attributed to any single logic gate. Instead, it displays a combination of gates at some specific background chatter level —see for instance all cases in Figure 6.2 in which the responses 0 (black) and 1 (white) are absent. In other cases, such as the one corresponding to the top row in Figure 6.2, the situation is simpler and a marked transition between simple logical gates is observed as background chatter is modified. This transition is quantified in terms of $\mu_i(q)$ and shown in Figure 6.4a. This Figure indicates that, at $q = 0$, the integration of the two inputs at the network level closely resembles an AND gate (Rule 8), as reflected by a small MAD to this rule. As the global level of chatter (q) increases, a transition from Rule 8 to 2 takes, and at $q = 1$ the integration of the inputs is essentially analogous to Rule 2 (y NIMP x). Other, somewhat less clear, transitions between simple logic gates are shown in Figure 6.4b-d. Figure 6.4c, for instance, shows a transition from a complete lack of information processing (Rule 0, OFF state), to the network-level integration of the two inputs in the form of a NAND gate (output is active as long as one of the inputs is inactive, Rule 7).

The transitions between simple binary logic gates are, to some extent, a generalisation of the results obtained for the integration between a single input and background chatter. The generalisation to interactions obtained for an increasing number of input signals may define a hierarchy of simple behaviours which somehow classifies widely varying input signal combinations into a set of simple biological responses (Helikar et al., 2008). These results show that the combination of the topology of the network and the specific logic functions defined at each node or, in other words, the full set of protein-protein interactions of the signalling network, are far from random, as we already pointed in Chapter 5. The function and dynamics of the nodes of the original network (biologically

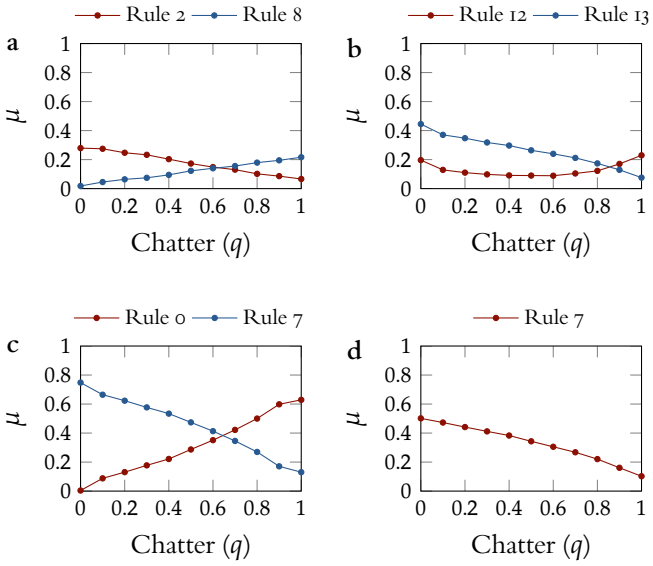

Figure 6.4: **Transition between simple logic gates.** Transitions, as measured by $\mu_i(q)$, for increasing background chatter q, for the four cases of Figure 6.2. Only the logic gates for which the mean absolute difference is smaller than a given threshold (here taken equal to 0.14 without loss of generality) for any value of q are shown. The meanings of the different lines given in the legends follow the notation of Table 6.1.

based, as opposed to randomly built), are organised in such a way that this hierarchy of integration levels, which classify the output response of the cell, end up in a well-defined pattern of robust responses that are biologically useful.

6.4 *Conclusions*

Cells are constantly exposed to fluctuating environmental conditions. External signals are sensed, processed and integrated by cellular signal transduction networks, which translate input signals into specific cellular responses by means of biochemical reactions. It is a well established fact that cells are capable of simultaneously processing multiple sources of information, a task that requires integration of very heterogeneous sorts of signals. A few isolated molecular modules have been uncovered in the last few years that implement such signal integration tasks (Padrick and Rosen, 2010; Prehoda et al., 2000). In vivo, however, these modules with well-established integration rules, are embedded in large signalling networks and affected by many concurrent and possibly randomly fluctuating signals which might influence and bias their function. Given the many examples existing in the physical sciences of the constructive influence of random fluctuations, it is reasonable to think that signalling networks, made up of signal integration modules, might have been tailored by evolution to cope with unavoidable noise to which cells are subject, in particular from its environment. Here we have addressed the

issue of how background noise affects the signal integration capabilities of the signalling network of a typical human cell, namely a fibroblast. Using a Boolean model of this network, we have analysed the integration of one or more signals for different levels of background chatter affecting the rest of inputs of the network. We have observed that

i the response to an input signal depends on the level of background chatter affecting all other input nodes,

ii even though the network is highly complex, the outputs respond in a very simple way to pairs of interacting input signals, and

iii variation of the chatter level causes transitions between these simple pattern responses.

This type of behaviour has been observed experimentally by Hsueh et al. (2009); Janes, Albeck, et al. (2005) and might suggest that the cell adapts its information processing machinery according to the environment.

IV

Transient spatio-temporal population dynamics in bacteria

7

Dynamics of cell death during biofilm formation

Development of multicellular organisms and microbial communities requires the accurate coordination of a large number of complex events: from fast and controlled cell proliferation, specialisation and rearrangement; to precise position-specific programmed cellular death. In these systems, and specially those of simpler organisms as bacteria, it is becoming more clear that cells marshal in a structured manner as a result of basic self-organising principles rather than due to large hard-coded genetic programs. In this Chapter we focus on one such cellular self-organising process: the assembly of *Bacillus subtilis* cells into macroscopic structures called biofilms. In particular, we concentrate on the transient spatio-temporal cell population dynamics that lead to the emergence of mesoscopic structures of massive cellular death. These previously unknown patterns of localised cell death spatially focus mechanical forces during biofilm growth and expansion, and thereby initiate the formation of the characteristic macroscopic wrinkles (see Figure 1.11).

7.1 *A coarse-grained population dynamics model of* Bacillus subtilis *cellular death*

In a recent work we have shown that, in the case of the soil bacterium *Bacillus subtilis*, these complex wrinkle structures arise as a consequence of massive cellular death orchestrated in time and space (experimental results are described in Section 1.3). In order to provide the conceptual understanding to these spatio-temporal patterns of death during biofilm formation, we developed a spatial, coarse-grained model that captures the density dynamics of a colony of bacteria growing on agar (Figure 7.1). Similar to previous theoretical approaches to microbial pattern formation, our model is based on a system of coupled partial differential equations that determine local cell density as a function of space and time. This

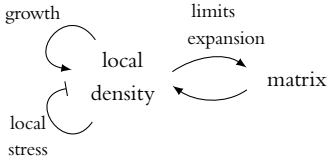

Figure 7.1: **Schematic of the mathematical model of CDP.** In the model local cell density is governed by density-dependent (logistic) growth and death and the ECM maintains density by preventing expansion and, in addition, local stress and cellular waste restrict cellular growth.

phenomenological model makes only the following five assumptions that are grounded on and constrained by experimental observations:

i Local cell density is increased by heterogeneous logistic growth. Heterogeneous cell growth in the model is based on experimental measurements of the spatial distribution of metabolically active cells, as reported by the carboxyfluorescein diacetate (CFDA) marker (Figure 7.2a). This data allowed us to calculate the structure function of the growth rate pattern (Figure 7.2b) and use it as the basis for the spatial growth heterogeneity in the model (Figure 7.2b–d and see below).

ii We assume that cell death occurs when density exceeds the carrying capacity due to global nutrient depletion and local stress such as toxin production.

iii Growth-driven increase in local cell density is counteracted by the tendency of cells to spread out in two-dimensional space.

iv Colony expansion and the resulting cell motion is affected by the presence of extracellular matrix. In particular, the *epsH* matrix gene deletion strain allowed us to quantitatively determine that cell density spreads twofold faster in the absence of the extracellular polysaccharide relative to the wild type strain (Asally et al., 2012).

v Furthermore, we assume that cells in regions of high density express matrix components, as supported by fluorescent reporter expression from the *yqxm* operon promoter (ibid.).

vi Finally, to account for the observed ring of high cell density, we incorporated a similar ring of high density in the starting condition of our simulations.

In addition, we utilised the experimental data to constrain modelling parameters wherever possible (see Table A.14).

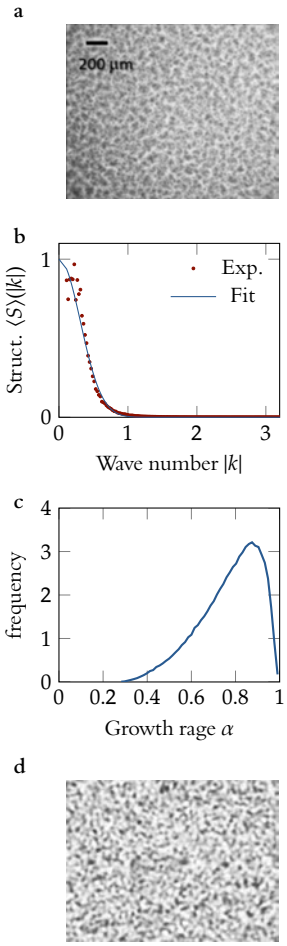

Figure 7.2: **Spatially heterogeneous growth.** (a) Metabolic reporter CFDA image of biofilm colony at 22 h indicates spatial heterogeneities in cellular growth. (b) Analysis of the spatial properties of the colony growth indicate that the heterogeneities have a characteristic wave number profile (red dots), that can be fitted with Equation (7.6). (c) Distribution of amplitudes used in the model, according to Equation (7.7) for $\phi = 0.25$ and $\alpha_0 = 1\,\text{cells}\,\mu\text{m}^{-1}\,\text{h}^{-2}$. (d) Computer-generated growth rate profile with spatial correlation $\lambda_c = 20\,\mu\text{m}$, $a_0 = 1$, $a_1 = 1$, fitted in panel (b).

With these postulates, the model consists of a reaction-diffusion equation that accounts for the dynamics of the density of living cells, ρ, coupled with three additional fields representing the carrying capacity, k, waste products, w, and extracellular matrix, m. The cell density equation reads:

$$\frac{\rho(\ ,t)}{t} = \alpha(\ )\rho(\ ,t)\ \ 1 - \frac{\rho(\ ,t)}{k(\ ,t)}\ \ + \quad D\,(m(\ ,t)) \quad \rho(\ ,t), \qquad (7.1)$$

where the vector $= (x, y)$ indicates the spatial coordinates. Following the first hypothesis above, bacterial growth is considered to follow a logistic law,

$$\alpha(\ )\rho\ \ 1 - \rho/k(\ , t)\ \ ,$$

with a spatially heterogeneous growth rate $\alpha(\ )$, which is kept constant in time (see below) and a dynamic local carrying capacity, $k(\ , t)$. In addition, colony expansion due to cell growth and the resulting cell motion is also considered, with diffusivity depending on the extracellular matrix, $D(m)$. The local carrying capacity, which is the maximum cell density that can be sustained at each point in space, is assumed to behave in a switch-like manner:

$$k(\ , t) = \frac{K(t) + k_1 \cdot (w(\ , t)/k_k)^{n_k}}{1 + (w(\ , t)/k_k)^{n_k}}, \tag{7.2a}$$

where

$$\frac{dK(t)}{dt} = -\frac{\gamma_K}{|\Omega|} \int_\Omega \rho(\ , t) d\ , \qquad K(0) = k_0, \tag{7.2b}$$

and

$$w(\ , t) = \int_0^t \rho(\ ,\ ) d\ . \tag{7.2c}$$

Here $K(t)$ is the maximum global carrying capacity that the agar medium can hold. This quantity decreases at a rate that depends on the average cell density on the whole system. Ω is the region of the 2D space considered for the experiment, and its area size is given by $|\Omega| = \int_\Omega d\ .$ For numerical integration purposes this area is considered as a square (see below). The term $w(\ , t)$ is the local accumulation of waste generated by living cells and its dynamics simply track the local history of cellular density. In the absence of waste, the local carrying capacity reaches its maximum value $k(\ , t) = K(t)$, but as waste accumulates beyond a certain concentration threshold, k_k, it decreases to a lower value, k_1. The steepness of this change is determined by the Hill coefficient n_k. Figure 7.3 depicts the dependence of the local carrying capacity on the waste, w.

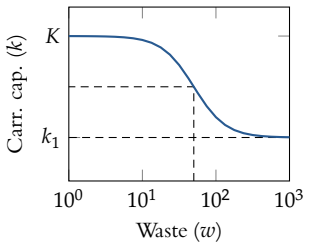

Figure 7.3: **Incidence of waste on carrying capacity.** In the absence of waste (w), the local carrying capacity (k) reaches the maximum value (K). Local accumulation of waste decreases the carrying capacity locally, and thus the maximum density that cells can reach at a given point in space.

Furthermore, as cells grow they start producing extracellular matrix. The model assumes that local production of extracellular matrix is triggered in a switch-like manner, when cell density is above a threshold k_m:

$$m(\ , t) = \lambda_m \int_0^t \frac{\rho(\ ,\ )/k_m{}^{n_m}}{(1 + \rho(\ ,\ )/k_m{}^{n_m})} d\ . \tag{7.3}$$

The presence of the matrix component is assumed to have a direct impact in cell mobility by significantly reducing the effective cell diffusivity, $D(m)$:

$$D(m) = D_0\ ^{-m}. \tag{7.4}$$

Finally, the model incorporates a Sytox cell death reporter, which is activated with high cell densities. The dynamics of the reporter are thus modelled by a sigmoidal activation and a linear degradation:

$$\frac{ds(\ ,t)}{dt} = \alpha_s \frac{\rho(\ ,t)^{n_s}}{k_s^{\ n_s} + \rho(\ ,t)^{n_s}} - \lambda_s s(\ ,t). \tag{7.5}$$

Figure 7.4 shows the Sytox activation curve (in blue) used in the model. The choice of this particular density-dependent Sytox reporter is based on experimental observation, as depicted by the grey crosses in this figure, which correspond to observed (rescaled) pixel values of density (8 h) and Sytox (14 h).

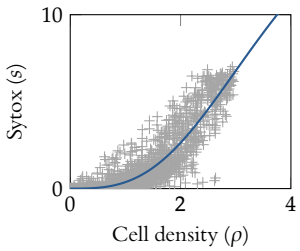

Figure 7.4: **Sytox dependence on cell density.** Nonlinear cell density-dependent activation of the Sytox reporter as modelled by Equation (7.5), together with the experimental data of the bright field at 8 h and the Sytox signal at 14 h (grey crosses).

A CRUCIAL ASSUMPTION of the model is the non-homogeneous profile of growth rates, $\alpha(\ )$, which is motivated by the experimental observation of spatially heterogeneous metabolic activity thorough the entire colony and which might arise due to differences at the microscopic level. In order to incorporate this heterogeneity into the model, we use computer generated heterogeneous growth rate profiles with specific spatial correlation and amplitude distribution. Following Garcia-Ojalvo and Sancho (1999), we generate a random variable $(\ )$ with spatial correlation

$$\langle (\ ),\ (\ )\rangle \propto {}^{-\frac{|-|^2}{2\lambda_c^{\ 2}}}{}^{2}, \tag{7.6}$$

that leads to a structure function that matches the one experimentally observed, as shown in Figure 7.2b, and to a Gaussian distribution of amplitudes. As cells are neither expected to grow infinitely fast nor get completely stalled, the growth rate profile is defined in the finite support $[\phi \cdot \alpha_0, \alpha_0)$ as

$$\alpha(\ ) = \alpha_0 \cdot \ \phi + (1-\phi)\frac{1}{1 + {}^{-a_0 - a_1\ (\ )}}\ . \tag{7.7}$$

In all the simulations we take the following parameter values $\lambda_c = 20\,\mu\text{m}$, $a_0 = 1, a_1 = 1$ and $\phi = 0.25$. Figure 7.2c shows the distribution of amplitudes for α using these parameter values, and Figure 7.2d shows an example of the simulated growth rate profile. All model parameters are shown in Table A.14.

Simulation of colony growth and cell death patterns

To give a sense of the local population dynamics, we first analyse the homogeneous solution of the model. That is to say, we disregard the spatial heterogeneities in growth rate and assume spatially homogeneous initial conditions. This leads to dynamics in 0 spatial dimensions —as in these conditions the diffusion operator cancels out—, that serve us to understand the role of different growth rates in the appearance of the Sytox signal. In this case, the system can be reduced to the set of ordinary differential equations:

$$\frac{d\rho}{dt} = \alpha\rho\left(1 - \frac{\rho}{k}\right), \tag{7.8a}$$

$$k = \frac{K + k_1 \cdot (w/k_k)^{n_k}}{1 + (w/k_k)^{n_k}}, \tag{7.8b}$$

$$\frac{dK}{dt} = -\gamma_K \rho, \tag{7.8c}$$

$$\frac{dw}{dt} = \rho, \tag{7.8d}$$

$$\frac{ds}{dt} = \alpha_s \frac{\rho^{n_s}}{k_s{}^{n_s} + \rho^{n_s}} - \lambda_s s. \tag{7.8e}$$

From these equations it is evident that waste accumulates unlimitedly and that the maximum carrying capacity K fades to zero. Thus, at long times the carrying capacity becomes k_1 and establishes the stationary cell density. However, before reaching this steady level, the cell population suffers a transient overshoot that triggers cell death, as shown in Figure 7.5. This Figure shows simulated time-traces of Sytox, cell density and carrying capacity for four different values of the growth rate ($\alpha = 1, 0.6, 0.4$ and 0.2) according to Equation (7.8), with initial density $\rho(0) = 0.01$ cells μm^{-2}. All other values have been taken from Table A.14. Panels a (Sytox signal) and b (density) of this Figure show that the smaller the growth rate, the longer it takes for cell density and Sytox signal to peak. In addition, the later this peak occurs, the smaller its amplitude, as the maximum carrying capacity that can be reached decreases in time, as Figure 7.5c shows. Thus, small differences in the growth rate lead to heterogeneous distributions of amplitudes and times of appearance of the maximum cell density, cellular death and the Sytox burst.

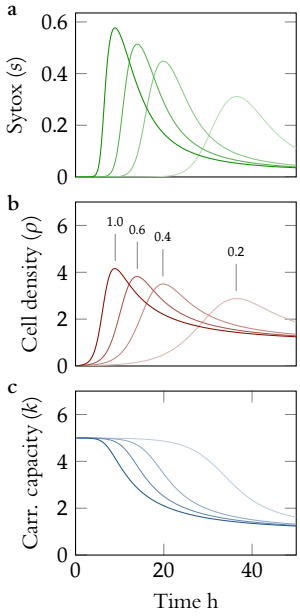

Figure 7.5: **Simulation of the 0-dimensional system.** Simulated time-traces for Sytox signal (a), cell density (b), and carrying capacity (c), for four different values of the growth rate (indicated in panel b).

Next, we numerically analyse the spatially heterogeneous case, replicating the experimental setup (see details on numerical schemes in Appendix A.4). To do so, we emulate the experimentally observed initial

conditions (Figure 7.6), which display a prominent ring of high cellular density due to the coffee ring effect (Yunker et al., 2011). This effect causes cells to accumulate at the edge of the suspension drop deposited on the Petri dish at the beginning of the experiment. In the model, we therefore randomly place spots of cell density in a circular domain with a higher probability to be close to the boundary of the drop, as shown in the diametrical probability density profile depicted in Figure 7.6b.

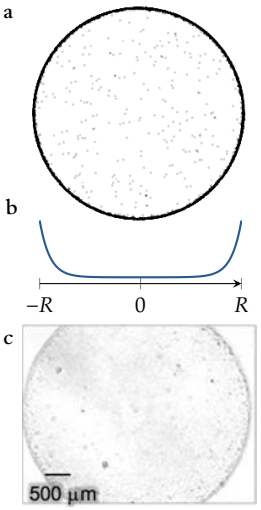

Figure 7.6: **Initial cell density conditions.** Random initial conditions for cell density used in the simulations (panel a) match those observed experimentally (panel c). In the simulation, cell colonies are randomly placed in the space, according to a radial distribution (panel b) that decays exponentially from the ring border to the centre.

Figure 7.7 shows the spatio-temporal dynamics of the model defined by Equations (7.1), (7.3) and (7.7), using the growth rate profile from Figure 7.2d and the initial conditions from Figure 7.6a. The model shows how a simple mechanism of heterogeneous cellular growth together with extracellular matrix production and waste accumulation can account for the observed spatio-temporal CDP (cf. Figures 1.12 and 7.7). According to this model, at the early stages of the colony formation micro-colonies grow and rapidly expand due to the absence of extracellular matrix. This growth and expansion is more pronounced around the initial ring of density. As soon as these areas get populated by cells (9 h in Figure 7.7) cell mobility diminishes due to the appearance of extracellular matrix. Together with the fact that there are areas with low growth rate, which get populated at a slower rate, an initial anti-spot-like pattern of high cell density appears (the interior of the ring is completely filled of cells, with the exception of some empty spots). In high cell density areas, waste

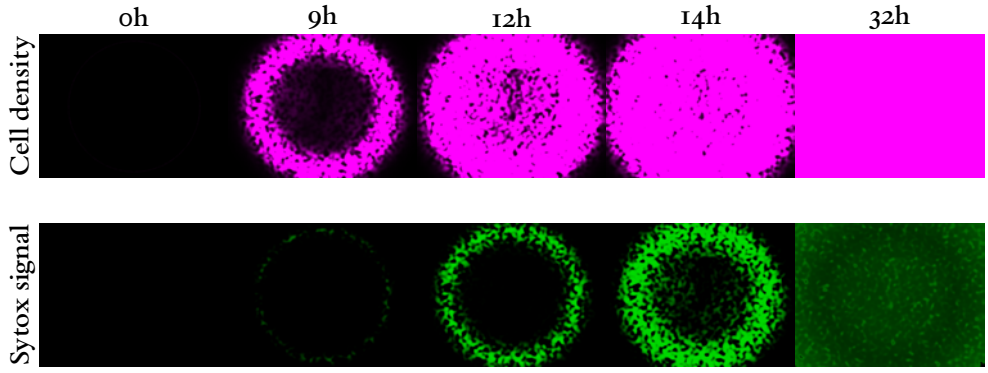

Figure 7.7: **Simulated cell death pattern dynamics.** The mathematical model, with the parameter values from Table A.14 reproduces the observed spatio-temporal dynamics. Compare with Figure 1.12.

accumulates and the local carrying capacity decreases, causing the early burst of Sytox signal (12 h and 14 h in Figure 7.7). As the early burst fades out, the slowly growing spots start reaching their local carrying capacity and the late burst occurs (32 h in Figure 7.7).

A particular property of biofilms is that they can develop even starting from a single-cell colony. We have observed that cell death patterns also occur in colonies grown from individual cells. The model described here

is able to reproduce the experimental results in those conditions as well (Figure 7.8).

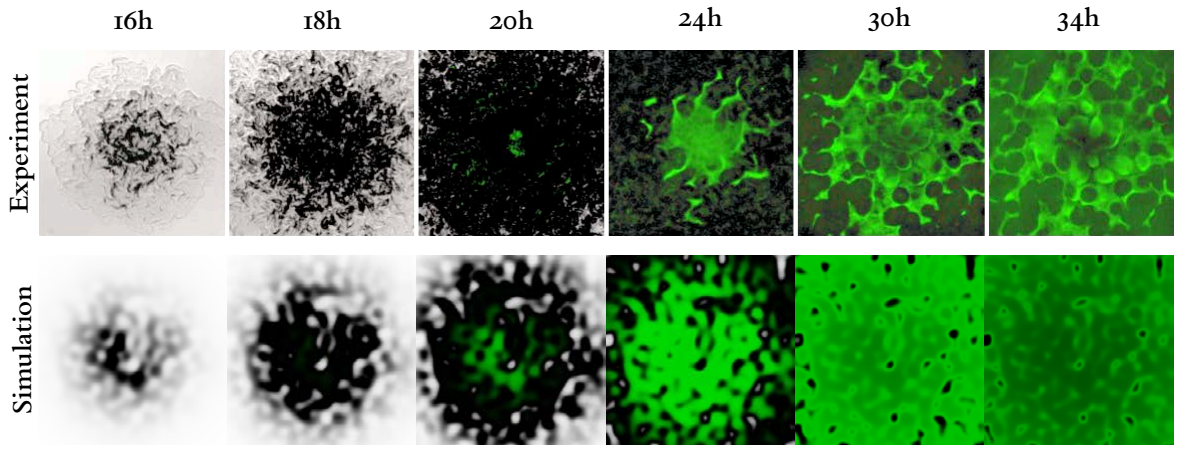

Figure 7.8: **Patterns of cellular death in a colony starting from a single cell.** (a) Experimental film strip of cell density (bright field: grey scale) and death (Sytox: green) during formation of a biofilm from a single cell. (b) Frames obtained from the integration of the model with a single colony as the initial condition. The bright field (in black) is computed from cell density, and Sytox levels are shown in green.

7.2 Identification of dynamic CDP clusters

To analyse the dynamic pattern of developing biofilms, we developed a correlation-based technique to group together areas of the microscope field that varied similarly over time relative to the mean (Asally et al., 2012). In particular, we were interested in spatio-temporal variations in the CDP as observed by Sytox fluorescence. Initial examination of the CDP revealed that certain areas of the CDP would demonstrate fluorescence at different times during the biofilm development, and that many of the fluorescence areas were highly localised. We quantified this phenomenon in both time and space.

We sought to identify deviations from the mean behaviour of the field and areas that deviated in a correlated fashion in time. To that end, for each time-lapse microscopy frame of the Sytox signal, we first subtracted the mean fluorescence intensity at each time point from each pixel at that time point. Pixels with a value greater than the mean would have a positive value, and those with intensity less than the mean would have a negative value. Using this method we were able to see which areas were brighter or dimmer than the mean fluorescence over time. To segment these areas, we employed *k*-means clustering analysis using a Pearson correlation-based metric (1 −) (details can be found in ibid.). From this analysis, we could separate areas of high Sytox intensity earlier in the experiment (early burst) from localised clusters occurring later (late bursts). Notably, we observed that some clusters had high relative Sytox intensity earlier in the biofilm development, corresponding to high initial

cell density, and that other clusters had high relative Sytox later in the time-lapse sequence, showing a clear temporal separation, as shown in Figure 1.13 (and reproduced in Figure 7.9a for clarity).

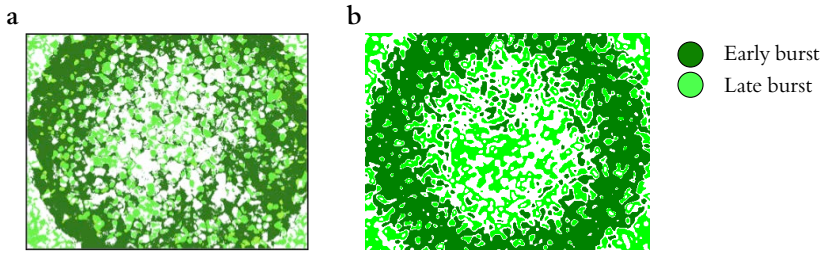

Figure 7.9: **Population dynamics during biofilm formation.** Spatio-temporal patterns of cell death consist of an early- and a late-burst of massive bacterial death. The early-burst appears in the form of isolated spots while the late-burst death occupies approximately the complementary space. (a) Experimental CDPs. (b) Simulation CDPs.

In order to bring insight into the CDP dynamics from the modelling perspective, the same quantitative analysis used in the experimental time-lapse imaging was applied to the simulation results. A *k*-means clustering was applied to the pixel (grid point) values of the Sytox field and the clusters were classified using the same criteria as in the experimental analysis. Figure 7.9b shows the clustering results for the simulations corresponding to wild type conditions.

7.3 Robustness of the model to parameter changes

The CDP can quantified by a spatial heterogeneity coefficient of the late-burst pattern. This coefficient is defined as the square of the perimeter length of the region divided by its area and normalised by 4 , and gives a sense of the complexity of the late burst patches —it is one for a perfect circle and increases as this is stretched anisotropically. In order to assess the robustness of the model to parameter changes, we performed a parameter sensitivity analysis using the spatial heterogeneity as a measure. We integrated the model equations with single parameter perturbations of 25 % their original value (cf. Table A.14, performed the clustering analysis as described above, and computed the spatial heterogeneity of the obtained CDP. Figure 7.10 shows that the spatial heterogeneity coefficient is robust to these parameter perturbations. In this way, the model accounts for the basic spatio-temporal dynamics of the wild type CDP, and does so robustly with respect to parameter variation.

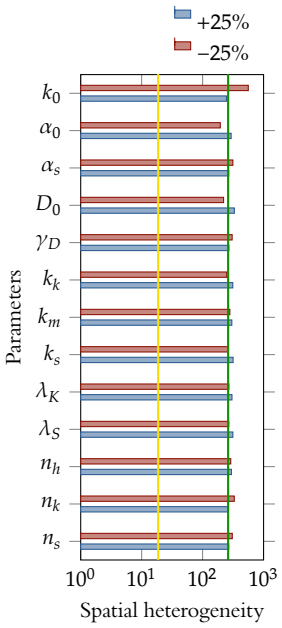

Figure 7.10: **Spatial heterogeneity does not vary considerably when parameter values are perturbed.** Robustness of the model to 25 % variations in the parameter values has been assessed through the spatial heterogeneity coefficient of the spatio-temporal dynamics. The green line marks the average spatial heterogeneity for the default (wt) parameters, whereas the yellow line corresponds to the average spatial heterogeneity of the *epsH* mutant (see below).

7.4 The extracellular matrix as a key component

The robust dynamics described by the model relies on a fast initial expansion of cells that is suddenly terminated by the production of the extracellular scaffold. It is thus reasonable to ask whether strong perturbations in the matrix production can alter the spatio-temporal patterns. In our model, complete disruption of matrix production is achieved by setting the parameter k_m to infinity. Simulations of the model in this scenario, shown in Figure 7.11, already indicates that cell density grows in a homogeneous manner and no spotted CDP is observed. From this picture it is clear that, in the absence of matrix production, death (Sytox) patterns do not appear. Thus, the model predicts that matrix production is decisive for the generation of CDPs, as it precludes cells in high-density regions from invading the low populated ones. To test this prediction experimentally, we utilised a *B. subtilis* strain that contains a deletion of the *epsH* gene, involved in extracellular polysaccharide synthesis and hence in the generation of the ECM. Fluorescence imaging of the *epsH* deletion mutant strain indicated that colonies formed with these cells do not generate

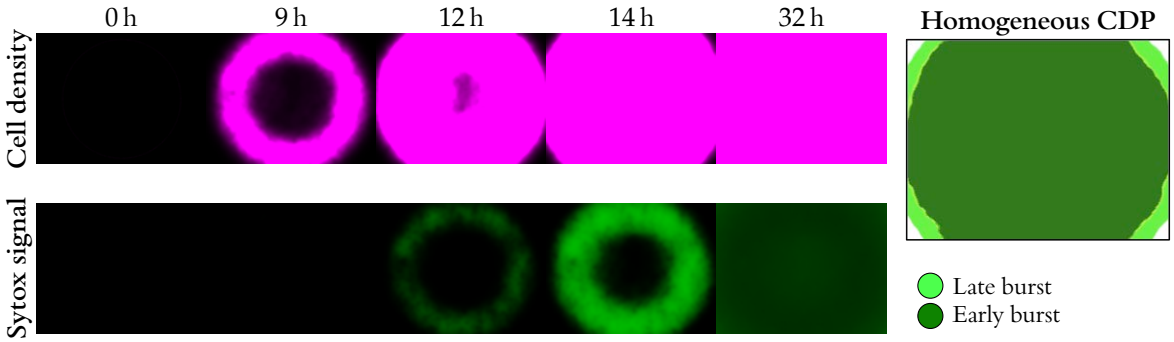

Figure 7.11: **Simulation in the absence of matrix production.** Numerical integration of the mathematical model for the *epsH* mutant ($k_m = \infty$) shows that the colony undergoes a spatially homogeneous growth followed by a uniform wave of cellular death starting at the high density ring (as indicated by Sytox signal). Hence, in this case the CDP consists only of a homogeneous ring of death.

the characteristic spotted pattern of cell death (see Figure 7.12). In accordance with the model prediction, these results show that extracellular matrix production is crucial to generate the CDP. Furthermore, these cells lacking the *epsH* gene were shown to be incompetent to generate wrinkled structures (Figure 1.15).

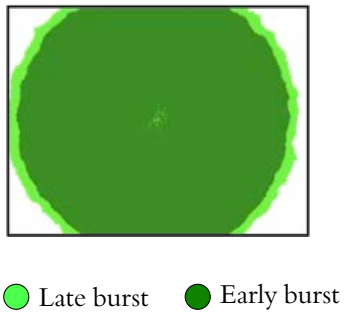

Figure 7.12: **Lack of matrix components leads to an homogeneous CDP** The *epsH* deletion strain generates an homogeneous spatio-temporal patterns of cell death, as shown here in the experimental spatio-temporal clustering of Sytox signal.

We further tested if the model was sensitive to partial changes in the regulation of matrix production. To do so, the main parameter responsible of the regulation matrix production was strongly perturbed. In specific, we varied the value of the matrix accumulation density threshold parameter, k_m, by several orders of magnitude and analysed the spatial heterogeneity of the resulting spatio-temporal dynamics. The obtained results are shown in Figure 7.13 and indicate a progressive loss of the spotted CDPs as matrix production is weakened. From these results we conclude that regulation of matrix production is a crucial aspect in the formation of the spatially localised structures of the observed CDP.

Clearly, the mesoscopic population dynamics model developed in this Chapter accounts for the spatio-temporal CDPs observed in wild type *Bacillus subtilis* cells. Simulations approximating very dilute starting

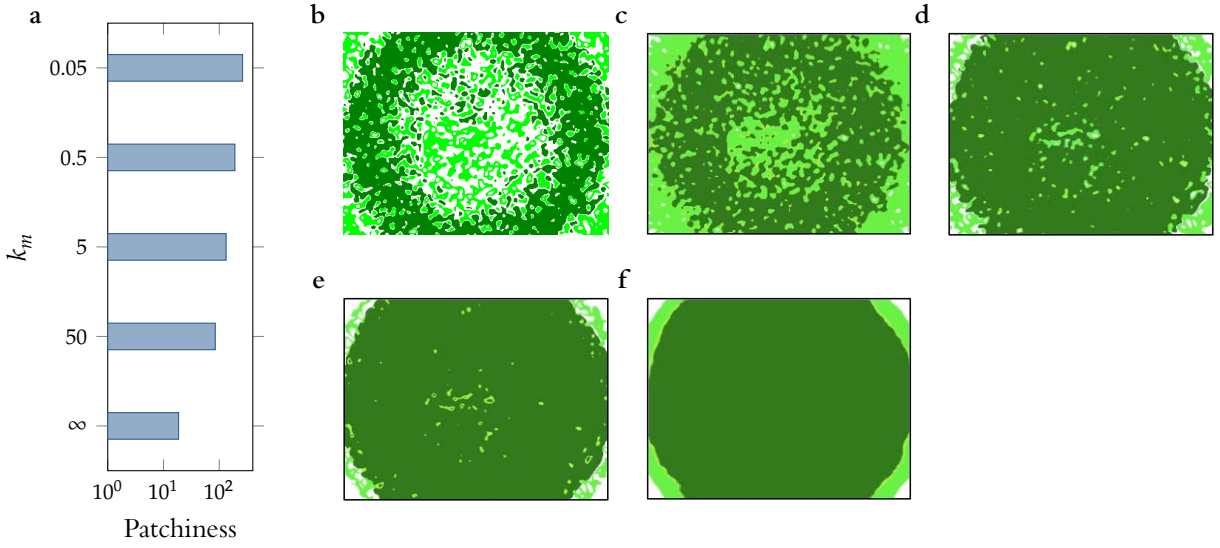

Figure 7.13: **The model is sensitive to large changes in the matrix accumulation density threshold.** The model is sensitive to large perturbations of k_m. The spatial heterogeneity coefficient decreases as this parameter is increased (a), and this is due to the disappearance of the late burst patches in favour of a continuous early burst (b–e) that approaches the CDP of the *epsH* deletion strain (f).

culture conditions (colonies starting with a single cell) are also consistent with experimental observations. In addition, the model predicted the loss of cell death patterns in matrix deficient strains. Therefore, this simple model accounts for the observed cell density and death patterns and provides a conceptual explanation for how the matrix can contribute to the generation of a spatially localised CDP.

7.5 *Conclusions and discussion*

Even simple unicellular organisms such as the soil bacterium *Bacillus subtilis* can form communities with complex three-dimensional structures called biofilms. These biofilms, which are composed by millions of cells glued together by the extracellular matrix, grant the microbial colony with enhanced resistance to extreme conditions, such as exposure to antibiotics or detergents. This increased level of protection is partly due to the macroscopic convoluted wrinkle structures that are generated during biofilm development. Until not long ago, it was poorly understood how these structures could emerge from the mechanical, chemical and genetic interactions at the cellular level.

Our recent experimental and theoretical results (Asally et al., 2012) have brought up evidence that, underlying the formation of wrinkles there is a well-defined spatio-temporal pattern of bacterial death. This mesoscopic, localised, non-homogeneous patterns of cellular death give rise to mechanical forces that can generate macroscopic 3D structures.

Indeed, by experimentally characterising the existing mechanical forces in the biofilm, we noted that the nucleation of wrinkle formation was initiated in areas of premature cell death.

By developing and analysing a simple spatial model of population dynamics which accounts for cellular growth and death, together with the formation of the extracellular matrix, we were able to get deeper insight into the generation of the death patterns. Specifically, we identified the source of the localised pattern formation as a combination between the heterogeneous cellular growth and the dynamics of matrix production. Bacterial growth and death events at a mesoscopic level give raise to spatial inhomogeneities in cell density that are accentuated by the effect the matrix on cellular motility. In fact, according to the model, matrix production is a key element in the generation of the spotted pattern of cellular death and, indeed, its removal leads to homogeneous bacterial growth an death. We then experimentally validated this strong prediction by the model by measuring cellular death in an extracellular matrix deficient strain and verified that this lack of death pattern lead to a flat biofilm with no wrinkles.

These results, grounded on experimental and theoretical analysis, can set a conceptual framework for observations made in other biological systems, including developmental multicellular organisms in which growth and/or cell death concurs with genetic and chemical programs to drive mechanical forces. Well studied systems such as dorsal closure during fruit fly morphogenesis, for instance, might bear some resemblance (with all due caution) to the system studied here, as mechanical forces seem to be directed toward the site of apoptosis.

V

Final remarks and Outlook

8

Conclusions

In this Thesis we have presented theoretical and computational results on some specific transient biological and cellular problems, that can be divided in four groups:

i dynamics of gene-regulated transient stochastic responses via excitable pulses,

ii transcriptional and post-translational mechanisms of pluripotency preservation in Embryonic Stem cells,

iii mechanisms of signal transduction and integration in a cell signalling network in the presence of background fluctuations, and

iv spatio-temporal population dynamics of bacterial communities during biofilm formation.

Although the works described in this dissertation are only theoretical, in some cases these have gone hand in hand with experimentation (carried out by close collaborators). One such work is the research on pluripotency regulation in Embryonic Stem cells. Experimental analysis of individual pluripotent ES cells brought up a previously unknown correlation between key pluripotency factors. Our theoretical work on this data resulted in a novel model of regulation based on post-translational interactions. The proposed model made two predictions: the presence of binary protein complexes in cells cultured in certain pluripotency conditions; and the association of high levels of free Oct4 with differentiation. The former prediction was experimentally validated with biochemistry, and the latter was substantiated by reanalysis of gene expression data.

Another example of fruitful experimental-theoretical collaborative research in this Thesis is the work on biofilms. Experimental investigation of wrinkled biofilms of *Bacillus subtilis* by our collaborators resulted in the discovery of a dynamic pattern of localised cell death that spatially focuses mechanical forces, and thereby initiates wrinkle formation. Based on known facts on bacterial growth, death and matrix production, we put forward a coarse-grained model of population dynamics that reproduced the spatio-temporal patterns observed. Careful computational analysis of the model revealed the central role of matrix production. In this case, the model predicted that, in the absence of extracellular matrix, the patterns of cellular death would vanish. This prediction was also experimentally validated using a matrix deficient bacterial strain. Indeed, these bacteria were not even capable of forming wrinkles, further reinforcing the suggested link between localised cellular death and focusing of mechanical forces as the initiators of wrinkle formation.

These two examples of collaborative research should illustrate the spirit of this dissertation: that mathematics as a language to state biological facts proofs to be a powerful tool to generate and state experimentally testable

hypotheses. That is to say, to rigorously apply the scientific method, as in other fields of science such as physics and chemistry, to identify basic principles or "laws" of biology.

Summary of results and outlook

On the dynamics of gene-regulated transient stochastic responses via excitable pulses

Dynamical behaviour is prevalent in many of gene regulatory processes. In some cases, gene regulatory responses consist of stochastically triggered transient pulses of protein expression. Two such cases are the transient differentiation into competent state by stressed *B. subtilis* bacteria (Süel et al., 2006), and the basal dynamics of the tumour suppressor p53 in mammalian cells (Loewer et al., 2010). The dynamics displayed by these processes can be understood as instances of excitability, the same basic principle ruling neuronal action potentials. We have shown that three different minimal architectures with only two interacting genes are capable of generating regulated transient responses through different excitable mechanisms. The three architectures, which are either known to exist in nature or have been synthetically assembled to exhibit pulsating behaviour, open the question of whether design degeneracy (i.e., different circuit architectures that perform the same task) is expected in natural systems. In this sense, recent studies (Çagatay et al., 2009; Kittisopikul and Süel, 2010) have shown that such degeneracy can be raised when considering the effect of noise. This "second order" distinctive feature provided by the architectures in the presence of random fluctuations needs to be further considered in each particular biological context (Çagatay et al., 2009). In this sense, other recently described excitability mechanisms should also be considered (Meng, Huguet, and Rinzel, 2012; Wieczorek et al., 2011), and genetic circuits implementing them analysed, as these might encode different biological roles.

We have further described how noise in a simple excitable gene circuit can expose other elaborate dynamics, such as regular generation of pulses with polymodal distribution of lengths. We have shown that this effect can appear as a combination of stochastic coherence and another noise-induced effect generated in the vicinity to a Hopf-stabilising bifurcation. The type of oscillations produced in this situation, with uneven periods, is known to occur in the cell cycles of some types of Chinese hamster cells, fission yeast cells and frog eggs. The fact that this behaviour can be found in different biological systems raises the question of whether this is ruled by a common basic principle. We have thus provided a candidate for such a principle, which requires only a very simple and common gene regulatory motif (a combination of activation and inhibition, plus noise). However, further experimental measurements at the single-cell level would be required in order to assess if the proposed mechanism is plausible.

On heterogeneity and correlations in the pluripotency network of Embryonic Stem cells

The route towards differentiation from pluripotent stem cells is usually understood to occur through a transition from a state with high absolute levels of the pluripotency factor Nanog (HN) to a state with low levels (LN). Under this scenario, an excitable mechanism can explain the low frequency fluctuations observed in the activity of Nanog in mouse Embryonic Stem cells (Kalmar et al., 2009). Careful analysis of the structure of an ES cell cultures (Luo et al., 2013) indicates that, in addition to the HN and LN subpopulations, there exists a third

subpopulation with intermediate levels of Nanog (MN). We have investigated the mechanisms of regulation of Nanog at the transcriptional level, and proposed a candidate circuit that can give a mathematical sense to the three observed subpopulations. The gene circuit proposed generates noise-triggered excursions from the HN state (to the MN state), with a small fraction of these being attracted by the LN state. Thus, the model, which precisely reproduces the experimentally reported profile of Nanog expression, predicts that cells transiting from high to middle levels of Nanog would be prone to jump to the low Nanog state and therefore become primed and ready for differentiation. However, in order to experimentally validate this prediction, extensive time-lapse imaging data of individual ES cells should be generated and processed.

By using single-cell quantitative immunofluorescence microscopy (with fixed cells), we identified that high correlations between protein levels of four key pluripotency factors, namely Nanog, Oct4, Tcf3 and β-catenin, are associated with increased pluripotency conditions. In the light of these results, we investigated and modelled how the activity of these four factors is regulated at the protein level. By doing so, we uncovered a key role for post-translational regulation in the maintenance of pluripotency, which complements the already mentioned transcriptional regulatory layer. Specifically, we found that the activity of a network of protein complexes involving the four factors can account for the behaviour of ES cells under different pluripotency conditions (i.e., different genetic and culture conditions). Our results, complemented with a bioinformatic analysis of published genetic data, suggest that the function of the network is to buffer the transcriptional activity of Oct4, which appears to be the main determinant of pluripotency. The protein network suggested explains the mechanisms underlying the gain and loss of function in different mutants, and brings us closer to a full understanding of the molecular basis of pluripotency. Our results open new questions around the biochemical details behind the role of complexes in the stabilisation of the key factors and the interaction between the transcriptional and post-translational layers of pluripotency regulation. These questions will have to be addressed from a multidisciplinary perspective that, again, combines theory and experimentation.

On relaxation and noise in cellular signalling networks

The analysis of isolated small genetic circuits is motivated by the growing evidence that most cellular processes are regulated by highly modular networks of interactions. Yet, most experimentally based protein-protein interaction and cellular signalling networks exhibit a high degree of connectivity among all their parts (Oda, Matsuoka, et al., 2005). There are several biological processes such as information transmission through cellular signal transduction that can only be addressed from a more global perspective. Cells are not passive objects which filter out blindly the modifications of the environment in which they live. On the contrary, they receive information from different environmental sources, process this information, and consequently make decisions. These actions allow them to adapt their behaviour to the information received. We have studied, in this respect, the roles of external background fluctuations in the process of information transmission and integration in a large signalling network. To address this issue in a way that explicitly accounts for the complexity of the system under consideration, we have used a Boolean signalling model. Our results indicate that the level of background chatter shapes the response of the whole signalling network to external information carrying signals, and modulates the activity levels of the network outputs similarly to the known stochastic resonance effect observed in other physical and

biological systems. Furthermore, we have observed that the chatter level effectively determines the paths of information flow within the network by selectively activating different components of the signalling network. In this way, different levels of background noise allow the network to select between different responses, given the same stimulus. We have also analysed the effect of background activity in the integration of pairs of external stimuli and found that the network response can be mapped into simple integration patterns to some degree of approximation. Interestingly, these integration patterns depend on the levels of background chatter the cell is subject to and thus varying levels of background activity allow for transitions in the way some external signals are integrated in the cell.

Together, these results indicate that cellular signal transduction and integration processes are dependent on the levels of environmental fluctuations, suggesting a trade-off between robust and context-dependent decision making in these networks. In this regard, it remains to be seen whether or not these phenomena are present in other models of cell signalling and even in real experimentally controllable cellular signalling networks.

On transient spatio-temporal population dynamics in bacteria

Finally, we have investigated an example of multicellular, colony-level process that might encode a biological function. Specifically, we have looked into the cellular growth and death dynamics taking place in microbial communities of *B. subtilis* during the emergence of sticky, wrinkled structures visible to the naked eye, called biofilms. By experimentally observing biofilms through a fluorescence microscope, we discovered a pattern of localised cell death that spatially focuses mechanical forces, and thereby initiates wrinkle formation.

Based on the experimental findings, we proposed a mesoscopic spatial model of population growth. The model, which considers the interactions between heterogeneous cell growth, local environment and matrix production, could robustly account for the complex transient spatio-temporal dynamics of the observed cellular death. In addition, it predicted that the extracellular matrix is key for the formation of such dynamics and the lack of matrix production components disrupts the complex pattern, a fact that was experimentally verified. These results, together with the experimental evidences that massive death triggers formation of wrinkles, substantiate the idea that this complex system encodes a community-level biological function: the enhancement of the colony resistance under stressful unfavourable conditions. These results set an elementary basis for the control of macroscopic wrinkle structures in biofilms from the microscopic cellular interactions.

VI

Appendix

A

Details of the studied computational models

In this Appendix we describe the mathematical models presented in this Thesis. These details have been kept out of the principal text to improve its readability.

A.1 Mathematical aspects of excitable gene circuits

In Chapter 2 we analyse the dynamics of three genetic circuits, viz. Circuits A, B and C (as named in that Chapter), that regulate transient cellular responses via excitable pulses. Below we derive the model equations from the sets of reactions. The numeric bifurcation analyses have been performed with the XPPAUT software (Ermentrout et al., 2003).

Derivation of Circuit A equations

Circuit A in Section 2.1 implements the core module that triggers transient competence-differentiation under nutritional stress (Süel et al., 2006) in *B. subtilis*. The continuous equations of this model (Equation (2.1)) can be implemented stochastically by the set of reactions shown in Table A.1. In this table, reactions (a) and (b) constitute the basal or leaky transcription of X and Y, while reactions (c) and (d) account for regulated expression of X and Y respectively. The mRNA production rates in these last two reactions are modelled by the activation and inhibition Hill functions, respecively:

$$f(x, k_2, k_x, n) = k_2 \frac{x^n}{x^n + {k_x}^n} \quad \text{and} \quad (x, k_4, k_y, m) = k_4 \frac{1}{1 + {x/k_y}^m},$$

where x and y are the respective concentrations of X and Y (see below). Reactions (e) and (f) correspond to translation of mRNA molecules into proteins. Degradation of mRNA (reactions g and h) is considered both as linear decay processes, whereas protein decay is subject of linear

Table A.1: **Set of reactions of Circuit A.**

Transcription	
$P_X \xrightarrow{k_1} P_X + m_X,$	(a)
$P_X \xrightarrow{f(x,k_2,k_x,n)} P_X + m_X,$	(b)
$P \xrightarrow{k_3} P + m,$	(c)
$P \xrightarrow{(x,k_4,k_y,m)} P + m,$	(d)
Translation	
$m_X \xrightarrow{k_5} m_X + X,$	(e)
$m \xrightarrow{k_6} m + Y,$	(f)
mRNA degradation	
$m_X \xrightarrow{k_7} \varnothing,$	(g)
$m \xrightarrow{k} \varnothing,$	(h)
Protein degradation	
$X \xrightarrow{k} \varnothing,$	(i)
$Y \xrightarrow{k_{10}} \varnothing,$	(j)
Competitive enzymatic degradation	
$X + P \xrightarrow{k_{+11}/} X{:}P,$	(k)
$Y + P \xrightarrow{k_{+12}/} Y{:}P,$	(l)
$X{:}P \xrightarrow{k_{-11}} X + P,$	(m)
$Y{:}P \xrightarrow{k_{-12}} Y + P,$	(n)
$X{:}P \xrightarrow{k_{13}} P,$	(o)
$Y{:}P \xrightarrow{k_{14}} P.$	(p)

or unregulated degradation (reactions i and j) as well as competitive enzymatic degradation by the protease P (reactions k to p), the levels of which are considered constant.

Note that equations with more than one reactant are affected by the system size parameter Ω [1], which is scaled out when considering equations over concentrations ($x = \Omega^{-1}\mathrm{X}$ and $y = \Omega^{-1}\mathrm{Y}$) rather than absolute numbers (Kampen, 2007).

[1] Here, the system size paremeter is $\Omega =$ $\cdot N_A$, where is the cell volume and N_A is Avogadro's number, and thus relates numbers of molecules to concentrations.

In this manner, we can derive a full set of continuous rate equations from this set of reactions. Let m_X, m , x, y, p, $x{:}p$ and $y{:}p$ be the concentrations of the species m_X, m , X, Y, P, X:P and Y:P, respectively. Then the corresponding rate equations are

$$\frac{dm_X}{dt} = k_1\Omega^{-1} + k_2\frac{x^n}{x^n + k_x{}^n}\Omega^{-1} - k_7 m_X, \tag{A.1a}$$

$$\frac{dm}{dt} = k_3\Omega^{-1} + k_4\frac{1}{1 + \ k_y/x^{\ m}}\Omega^{-1} - k_8 m \ , \tag{A.1b}$$

$$\frac{dx}{dt} = k_5 m_X - k_9 x - k_{+11} p \cdot x + k_{-11} x{:}p, \tag{A.1c}$$

$$\frac{dy}{dt} = k_6 m \ - k_{10} y - k_{+12} p \cdot y + k_{-12} y{:}p, \tag{A.1d}$$

$$\frac{dp}{dt} = -k_{+11} p \cdot x - k_{+12} p \cdot y \tag{A.1e}$$

$$+ k_{-11} x{:}p + k_{-12} y{:}p + k_{13} x{:}p + k_{14} y{:}p, \tag{A.1f}$$

$$\frac{dx{:}p}{dt} = k_{+11} p \cdot x - k_{-11} x{:}p - k_{13} y{:}p, \tag{A.1g}$$

$$\frac{dy{:}p}{dt} = k_{+12} p \cdot y + k_{-12} y{:}p - k_{14} y{:}p. \tag{A.1h}$$

Adiabatic elimination of the protease activity as well as mRNA species, leads to Equation (2.1). Where the parameter conversion is:

$$a_1 = \frac{k_1 k_5}{k_7}\Omega^{-1}, \quad b_1 = \frac{k_2 k_5}{k_7}\Omega^{-1}, \quad a_2 = \frac{k_3 k_6}{k_8}\Omega^{-1},$$

$$b_3 = \frac{k_4 k_6}{k_8}\Omega^{-1}, \quad c_1 = \frac{p \ \ k_{+11}}{k_{-11} + k_{13}}, \quad c_2 = \frac{p \ \ k_{+12}}{k_{-12} + k_{14}},$$

$$_1 = \frac{k_{-11} + k_{13}}{k_{+11}}, \qquad _2 = \frac{k_{-12} + k_{14}}{k_{+12}}, \quad d_1 = k_7, \quad d_2 = k_8.$$

In this system, the total amount of protease is conserved, thus $p \ \ =$ $p + x{:}p + y{:}p = \mathrm{P} \ /\Omega$ is a constant.

Derivation of Circuit B equations

Circuit B described in Chapter 2 is an activator-repressor system where the inhibition of the activator species X by the repressor Y takes place at the post-translational level. This repression can be thought of as the Y-mediated degradation of X. Table A.2 lists the reactions defining this system. Again, reactions (a) to (d) describe the unregulated and regulated expression of X and Y transcripts, reactions (e) and (f) protein translation, and reactions (g) to (j) the degradation of mRNA and protein species. The last reaction (k) expresses the mentioned Y-mediated degradation of X. In this circuit, both promoter input functions are assumed to be activation Hill functions:

$$f(x, k_2, K_X, n) = k_2 \frac{x^n}{x^n + k_x{}^n} \quad \text{and} \quad (x, k_4, k\ \ , m) = k_4 \frac{x^m}{x^m + k\ \ ^m}.$$

The full set of reaction rate equations for the concentrations, derived from the reaction set, is:

$$\frac{dm_X}{dt} = k_1\Omega^{-1} + k_2\Omega^{-1}\frac{x^n}{x^n + k_x{}^n} - k_7 m_X, \tag{A.2a}$$

$$\frac{dm}{dt} = k_3\Omega^{-1} + k_4\Omega^{-1}\frac{x^m}{x^m + k_y{}^m} - k_8 m\ \ , \tag{A.2b}$$

$$\frac{dx}{dt} = k_5 m_X - k_9 x - k_{11} x \cdot y \tag{A.2c}$$

$$\frac{dy}{dt} = k_6 m\ \ - k_{10} y \tag{A.2d}$$

Again, assuming fast equilibrium for the mRNA species, we obtain Equation (2.2), with the following parameter conversions:

$$a_1 = \frac{k_1 k_5}{k_7}\Omega^{-1}, \quad b_1 = \frac{k_2 k_5}{k_7}\Omega^{-1}, \quad a_2 = \frac{k_3 k_6}{k_8}\Omega^{-1},$$
$$b_3 = \frac{k_4 k_6}{k_8}\Omega^{-1}, \qquad = k_{11}/\Omega, \quad d_1 = k_7, \quad d_2 = k_8.$$

Table A.2: **Set of reactions of Circuit B.**

Transcription	
$\mathrm{P_X} \xrightarrow{k_1} \mathrm{P_X} + \mathrm{m_X},$	(a)
$\mathrm{P_X} \xrightarrow{f(x,k_2,k_x,n)} \mathrm{P_X} + \mathrm{m_X},$	(b)
$\mathrm{P} \xrightarrow{k_3} \mathrm{P} + \mathrm{m}\ ,$	(c)
$\mathrm{P} \xrightarrow{(x,k_4,k_y,p)} \mathrm{P} + \mathrm{m}\ ,$	(d)
Translation	
$\mathrm{m_X} \xrightarrow{k_5} \mathrm{m_X} + \mathrm{X},$	(e)
$\mathrm{m} \xrightarrow{k_6} \mathrm{m} + \mathrm{Y},$	(f)
mRNA degradation	
$\mathrm{m_X} \xrightarrow{k_7} \varnothing,$	(g)
$\mathrm{m} \xrightarrow{k} \varnothing,$	(h)
Protein degradation	
$\mathrm{X} \xrightarrow{k} \varnothing,$	(i)
$\mathrm{Y} \xrightarrow{k_{10}} \varnothing,$	(j)
Enzymatic degradation	
$\mathrm{X} + \mathrm{Y} \xrightarrow{k_{11}/} \mathrm{Y}.$	(k)

Derivation of Circuit C equations

Circuit C in Chapter 2 is a standard transcriptional activator-repressor system where a species X activates both itself X its transcriptional inhibitor Y. Table A.3 shows the reactions involved in this circuit. The kinetics of the regulated activation of the P_X promoter, shown in reaction (b), is now described by a Hill function with cooperativity n:

$$f(x, y, k_2, k_{\text{eff}}, n) = k_2 \frac{x^n}{x^n + k_{\text{eff}}^{\ n}},$$

where x is the activator concentration in the cell, and the effective activation threshold k_{eff} depends on the concentration y of repressor through the expression $k_{\text{eff}}^{\ m} = k_x^{\ n}\left(1 + \left(\frac{y}{k_i}\right)^m\right)$. In this equation, the term k_i accounts for the competitive inhibition exerted by Y, the net effect of which is to increase the promoter's activation threshold. The kinetics of regulated transcription of the repressor mRNA is described by a simple activation Hill equation

$$(x, k_4, k_y, p) = k_4 \frac{x^p}{x^p + k_y^{\ p}}.$$

with constant activation threshold k_y.

Again, applying standard kinetics rules to the reactions leads in a straightforward way to the following coupled ordinary differential equations:

$$\frac{dm_X}{dt} = k_1\Omega^{-1} + k_2\Omega^{-1}\frac{x^n}{x^n + k_x^{\ n} + \gamma^m} - k_7 m_X, \tag{A.3a}$$

$$\frac{dm}{dt} = k_3\Omega^{-1} + k_4\frac{x^m}{x^p + k_y^{\ p}} - k_8 m \ , \tag{A.3b}$$

$$\frac{dx}{dt} = k_5 m_X - k_9 X, \tag{A.3c}$$

$$\frac{dy}{dt} = k_6 m \ - k_{10} \ . \tag{A.3d}$$

where $\gamma^m = \frac{k_x^{\ n}}{k_i^{\ m}}$.

Further assuming a separation of mRNA and protein time-scales, the former ones (being much smaller) can be adiabatically eliminated, and the system can be reduced to Equation (2.3) with

$$a_1 = \frac{k_1 k_5}{k_7}\Omega^{-1}, \quad b_1 = \frac{k_2 k_5}{k_7}\Omega^{-1}, \quad a_2 = \frac{k_3 k_6}{k_8}\Omega^{-1},$$
$$b_2 = \frac{k_4 k_6}{k_8}\Omega^{-1}, \quad d_1 = k_9, \quad d_2 = k_{10}.$$

The parameter values used in the three circuits in Chapter 2 are re-

Table A.3: **Set of reactions of Circuit C.**

Transcription	
$P_X \xrightarrow{k_1} P_X + m_X,$	(a)
$P_X \xrightarrow{f(x,k_2,k_{\text{eff}},n)} P_X + m_X,$	(b)
$P \xrightarrow{k_3} P + m \ ,$	(c)
$P \xrightarrow{(x,k_4,k\ ,p)} P + m \ ,$	(d)
Translation	
$m_X \xrightarrow{k_5} m_X + X,$	(e)
$m \xrightarrow{k_6} m + Y,$	(f)
mRNA degradation	
$m_X \xrightarrow{k_7} \varnothing,$	(g)
$m \xrightarrow{k} \varnothing,$	(h)
Protein degradation	
$X \xrightarrow{k} \varnothing,$	(i)
$Y \xrightarrow{k_{10}} \varnothing.$	(j)

sumed in Table A.4. The values for the stochastic simulation parame-

	Circuit A		Circuit B		Circuit C		
Par.	type I	type II	type I	type II	type I	type II	**Units**
a_1	0.008 75	0.016	0.008 75	0.005	0.15	0.147	nM s^{-1}
a_2	0	0	0.075	0.025	0.05	0	nM s^{-1}
b_1	7.5	6	7.5	15	7.5	7	nM s^{-1}
b_2	0.06	0.3	2.5	0.8	10	0.8	nM s^{-1}
c_1	0.001	0.001	–	–	–	–	nM s^{-1}
c_2	0.001	0.001	–	–	–	–	nM s^{-1}
d_1	0.0001	0.0001	0.0001	5×10^{-5}	0.000 386	0.0004	s^{-1}
d_2	0.0001	0.0001	0.0001	5×10^{-5}	3.86×10^{-5}	4×10^{-5}	s^{-1}
	–	–	4×10^{-8}	2.5×10^{-7}	–	–	nM^{-1} s^{-1}
k_x	5000	4000	5000	3000	5000	5000	nM
k_y	833	500	2500	750	75 000	2000	nM
$_1$	25 000	20 000	–	–	–	–	–
$_2$	20	100	–	–	–	–	–
γ	–	–	–	–	2	0.5	–
n	2	2	2	2	2	2	–
m	5	2	2	2	2	2	–
p	–	–	–	–	2	2	–

Table A.4: **Parameter values used in the deterministic equations of the three gene regulatory circuits.** For these sets of parameters, the circuits exhibit either type I or type II excitability.

ters are derived from this table assuming $\Omega = 1\,\text{molec}\,\text{nM}^{-1}$, mRNA degradation rates $k_7 = k_8 = 0.005\,\text{molec}\,s^{-1}$, and translation rates $k_5 = k_6 = 0.2\,s^{-1}$.

Circuit C is also investigated in the study of the interplay of polymodality and stochastic coherence in Chapter 3. In this case, the parameter values used are summarised in table Table A.5.

Par.	**Value**	**Units**	**Par.**	**Value**	**Units**
k_1	0.006 25	nM s^{-1} $molec^{-1}$	K_X	5000	nM
k_2	0.5	nM s^{-1} $molec^{-1}$	K	9000	nM
k_3	0.0	nM s^{-1} $molec^{-1}$	K	5000	nM
k_4	0.05	nM s^{-1} $molec^{-1}$	n	2	–
k_5	2	s^{-1}	m	2	–
k_6	2	s^{-1}	p	4	–
k_7	0.05	s^{-1}			
k_8	0.05	s^{-1}			
k_9	0.001	s^{-1}			
k_{10}	0.0001	s^{-1}			

Table A.5: **Values of the reaction rates used in the stochastic simulations of the circuit analysed in Chapter 3.**

A.2 Mathematical models of pluripotency

A model of pluripotency with three states

In Section 4.1 we describe a simple circuit for the mutual regulation of Oct4 and Nanog in pluripotency that accounts for the three experimentally described phenotypes of ES cells. The continuous 2d system describing the model dynamics (Equation (4.1)) can be derived from the set of reactions listed in Table A.6.

From this set of reactions we can derive the complete set of reaction rates for the species concentrations:

$$\frac{dm_N}{dt} = k_1\Omega + f(N, A, \ ; k_2, k_3)\Omega - k_{11}m_N, \tag{A.4a}$$

$$\frac{dm_A}{dt} = k_4\Omega + \ (N, A; k_5, k_6)\Omega - k_{12}m_A, \tag{A.4b}$$

$$\frac{dm}{dt} = k_7 O - k_{13}m \ , \tag{A.4c}$$

$$\frac{dN}{dt} = k_8 m_N - k_{14}N, \tag{A.4d}$$

$$\frac{dA}{dt} = k_9 m_O - k_{15}O. \tag{A.4e}$$

where the input functions are

$$f(N, A, \ ) = \frac{k_2(N/K_{nn})^{p_n} + k_3(A/K_{na})^{p_a}}{1 + (N/K_{nn})^{p_n} + (A/K_{na})^{p_a} + (\ /K_{nb})^{p_b}} \tag{A.5a}$$

$$(N, A) = \frac{k_5(\mathrm{A}/K_{aa})^{q_a} + k_6(N/K_{an})^{q_n}}{1 + (A/K_{aa})^{q_a} + (N/K_{an})^{q_n}} \tag{A.5b}$$

Assuming fast dynamics for the mRNA species and also for B, we can again apply dimensionality reduction:

$$\mathrm{m}_N = \frac{k_1}{k_{11}}\Omega^{-1} + \frac{f(N, A, \ )}{k_{11}}\Omega^{-1},$$

$$\mathrm{m}_O = \frac{k_3}{k_{12}}\Omega^{-1} + \frac{(N, A)}{k_{12}}\Omega^{-1},$$

$$\mathrm{B} = \frac{k_{10}k_7}{k_{13}k_{16}}.\mathrm{A}$$

Using these algebraic relations we can reduce Equation (A.4) to Equation (4.1). The parameter conversion between these two sets of equations

Table A.6: **Set of reactions of the 3 states model of pluripotency.**

Transcription	
$\mathrm{P}_N \xrightarrow{k_1} \mathrm{P}_N + \mathrm{m}_N$	(a)
$\mathrm{P}_N \xrightarrow{f(N,A,\ )} \mathrm{P}_N + \mathrm{m}_N$	(b)
$\mathrm{P}_A \xrightarrow{k_4} \mathrm{P}_A + \mathrm{m}_A$	(c)
$\mathrm{P}_A \xrightarrow{(N,A)} \mathrm{P}_A + \mathrm{m}_A$	(d)
$\mathrm{P} \xrightarrow{k_7\mathrm{B}} \mathrm{P} + \mathrm{m}$	(e)
Translation	
$\mathrm{m}_N \xrightarrow{k} \mathrm{m}_N + \mathrm{N}$	(f)
$\mathrm{m}_A \xrightarrow{k} \mathrm{m}_A + \mathrm{A}$	(g)
$\mathrm{m} \xrightarrow{k_{10}} \mathrm{m} + \mathrm{B}$	(h)
mRNA degradation	
$\mathrm{m}_N \xrightarrow{k_{11}} \varnothing$	(i)
$\mathrm{m}_A \xrightarrow{k_{12}} \varnothing$	(j)
$\mathrm{m} \xrightarrow{k_{13}} \varnothing$	(k)
Protein degradation	
$\mathrm{N} \xrightarrow{k_{14}} \varnothing$	(l)
$\mathrm{A} \xrightarrow{k_{15}} \varnothing$	(m)
$\mathrm{B} \xrightarrow{k_{16}} \varnothing$	(n)

are:

$$a_n = \frac{k_1 k_8}{k_{11}}\Omega^{-1}, \quad b_{nn} = \frac{k_2 k_8}{k_{11}}\Omega^{-1}, \quad b_{na} = \frac{k_3 k_8}{k_{11}}\Omega^{-1},$$
$$a_a = \frac{k_4 k_9}{k_{12}}\Omega^{-1}, \quad b_{an} = \frac{k_5 k_9}{k_{12}}\Omega^{-1}, \quad b_{aa} = \frac{k_6 k_9}{k_{12}}\Omega^{-1},$$
$$= \frac{k_{10} k_7}{k_{13} k_{16}}, \qquad d_n = k_{14}, \qquad d_a = k_{15},$$
$$K_{nb} = K_{ni}/\ ,$$

and the parameter values used in simulations are summarised in Table A.7.

Table A.7: **Parameter values used in the simulations of the three states model of pluripotency.**

Par.	Value	Units	Par.	Value	Units
k_1	0.004	–	k_{NN}	4000	–
k_2	0.0	–	k_{NA}	7000	–
k_3	0.365	–	k_N	10 000	–
k_4	0.065	–	k_{AN}	5000	–
k_5	0.004	–	k_{AA}	10 000	–
k_6	0.004	–	p_N	2	–
k_7	2.5×10^{-5}	–	p_A	2	–
k_8	0.05	–	p	2	–
k_9	0.05	–	q_N	4	–
k_{10}	0.05	–	q_A	4	–
k_{11}	0.005	–			
k_{12}	0.005	–			
k_{13}	0.005	–			
k_{14}	3.5×10^{-4}	–			
k_{15}	3.5×10^{-6}	–			
k_{16}	0.001	–			

Minimal pluripotency transcriptional model

In Section 4.1 we derive a minimal model of pluripotency regulation, Equation (4.2), that accounts for the observed Nanog fluctuations via an excitable mechanism. The model is based on the reactions in Table A.8.

Here, the Nanog promoter is regulated by competitive inhibition dynamics controlled by the factor B:

$$f(N,\ ) = \frac{(N/k_N)^p}{1 + (N/k_n)^p + (\ /K\ )^q}.$$

We can derive the set of rate equations from the system reactions (Table A.8)

$$\frac{dm_N}{dt} = k_1\Omega + k_2\Omega\frac{\mathrm{N}^p}{k_N{}^p + N^p + (\quad)^q} - k_9 m_N, \tag{A.6a}$$

$$\frac{dm_A}{dt} = k_3\Omega + \frac{k_4}{\Omega}NA - k_{10}m_A, \tag{A.6b}$$

$$\frac{dm}{dt} = k_5 A - k_{11}m\ , \tag{A.6c}$$

$$\frac{dN}{dt} = k_6 m_N - k_{12}N, \tag{A.6d}$$

$$\frac{dA}{dt} = k_7 m_A - k_{13}A, \tag{A.6e}$$

$$\frac{d}{dt} = k_8 m\ - k_{14}\ . \tag{A.6f}$$

Assuming fast dynamics of mRNA molecules and the species B.

Where the parameters are related as:

$$a_n = \frac{k_6k_1}{k_9}\Omega, \qquad b_n = \frac{k_6k_2}{k_9}\Omega, \qquad a_a = \frac{k_7k_3}{k_{10}}\Omega, \qquad b_a = \frac{k_7k_4}{k_{10}\Omega}$$

$$d_n = k_{12}, \qquad d_a = k_{13}, \qquad = \gamma\overline{\quad}\overline{\lambda},$$

where

$$\gamma = \frac{(k_N)^{p/q}}{k}, \qquad b_b = \frac{k_8k_5}{k_{11}\Omega}, \qquad d_b = k_{14}$$

Table A.8: **Set of reactions of the Nanog transcriptional model.**

Transcription	
$\mathrm{P}_N \xrightarrow{k_1} \mathrm{P}_N + \mathrm{m}_N,$	(a)
$\mathrm{P}_N \xrightarrow{f(N,k_2,k_N,n)} \mathrm{P}_N + \mathrm{m}_N,$	(b)
$\mathrm{P}_A \xrightarrow{k_3} \mathrm{P}_A + \mathrm{m}_A,$	(c)
$\mathrm{P}_A + \mathrm{N} + \mathrm{A} \xrightarrow{k_4} \mathrm{P}_A + \mathrm{N} + \mathrm{A} + \mathrm{m}_A,$	(d)
$\mathrm{P}\ + \mathrm{A} \xrightarrow{k_5} \mathrm{P}\ + \mathrm{A} + \mathrm{m}\ ,$	(e)
Translation	
$\mathrm{m}_N \xrightarrow{k_6} \mathrm{m}_N + \mathrm{N},$	(f)
$\mathrm{m}_A \xrightarrow{k_7} \mathrm{m}_A + \mathrm{A},$	(g)
$\mathrm{m}\ \xrightarrow{k}\ \mathrm{m}\ + \mathrm{B},$	(h)
mRNA degradation	
$\mathrm{m}_N \xrightarrow{k} \varnothing,$	(i)
$\mathrm{m}_A \xrightarrow{k_{10}} \varnothing,$	(j)
$\mathrm{m}\ \xrightarrow{k_{11}} \varnothing,$	(k)
Protein degradation	
$\mathrm{N} \xrightarrow{k_{12}} \varnothing,$	(l)
$\mathrm{A} \xrightarrow{k_{13}} \varnothing,$	(m)
$\mathrm{B} \xrightarrow{k_{14}} \varnothing,$	(n)

The NOC model

Transcription		**mRNA degradation**	
$P^{\alpha}_{N,0} \xrightarrow{k_{+1}} P^{\alpha}_{N,1}$,	(a)	$m_N \xrightarrow{k_7} \varnothing$,	(i)
$P^{\alpha}_{N,1} \xrightarrow{k_{-1}} P^{\alpha}_{N,0}$,	(b)	$m_O \xrightarrow{k} \varnothing$,	(j)
$P^{\alpha}_{O,0} \xrightarrow{k_{+2}} P^{\alpha}_{O,1}$,	(c)	**Protein degradation**	
$P^{\alpha}_{O,1} \xrightarrow{k_{-2}} P^{\alpha}_{O,0}$,	(d)	$N \xrightarrow{k} \varnothing$,	(k)
$P^{\alpha}_{N,1} \xrightarrow{k_3} P^{\alpha}_{N,1} + m_N$,	(e)	$O \xrightarrow{k_{10}} \varnothing$,	(l)
$P^{\alpha}_{O,1} \xrightarrow{k_4} P^{\alpha}_{O,1} + m_O$,	(f)	$O{:}N \xrightarrow{k_{11}} \varnothing$,	(m)
Translation		**Complex formation**	
$m_N \xrightarrow{k_5} m_N + N$,	(g)	$N + O \xrightarrow{k_{+12}} O{:}N$,	(n)
$m_O \xrightarrow{k_6} m_O + O$,	(h)	$O{:}N \xrightarrow{k_{-12}} N + O$,	(o)

Table A.9: **Set of reactions of the NOC model.**

In Section 4.2 we study the regulation of pluripotency in ES cells at the post-translational level. In that Section we consider a minimal motif based on the formation of a protein complex between Nanog and Oct4. The reactions describing the complex association and dissociation dynamics together with unregulated transcriptional events that bring new free proteins to the system are detailed in Table A.9 and the parameter values used in the stochastic simulations are shown in Table A.10 below.

Par.	Value	Units	Description
k_{+1}	0.04 (S+L)	h^{-1}	Nanog promoter activation rate
k_{+1}	0.1 (2i+L)		
k_{-1}	1	h^{-1}	Nanog promoter inactivation rate
k_{+2}	0.1	h^{-1}	Oct4 promoter activation rate
k_{-2}	1	h^{-1}	Oct4 promoter inactivation rate
k_3	25	transc h^{-1}	Nanog mRNA transcription rate
k_4	25	transc h^{-1}	Oct4 mRNA transcription rate
k_5	100	molec transc^{-1} h^{-1}	Nanog translation rate
k_6	25	molec transc^{-1} h^{-1}	Oct4 translation rate
k_7	0.1	h^{-1}	Nanog mRNA degradation rate
k	0.1	h^{-1}	Oct4 mRNA degradation rate
k (λ_N)	5	h^{-1}	Free Nanog degradation rate
k_{10}(λ_O)	0.1	h^{-1}	Free Oct4 degradation rate
k_{11}(λ_{ON})	0.1	h^{-1}	Nanog-Oct4 complex degradation rate
k_{+12}(k_+)	0.2	molec^{-1} h^{-1}	Nanog-Oct4 complex association rate
k_{-12}(k_-)	0.5	h^{-1}	Nanog-Oct4 complex dissociation rate

Table A.10: **Parameter values for the NOC model.**

The TBON model

The second (larger) model investigated in Section 4.3 consists of multiple reactions of binary complex association and dissociation. The full set of

reactions considered is detailed in Table A.11.

Table A.11: **Set of reactions of the NOC model.**

Nanog promoters	
$P^{\alpha}_{N,0} \xrightarrow{k_{+1}} P^{\alpha}_{N,1}$,	$P^{\alpha}_{N,1} \xrightarrow{k_{-1}} P^{\alpha}_{N,0}$,
$P^{\alpha}_{N,0} + T \xrightarrow{k_{+2}} P^{\alpha}_{N,2}$,	$P^{\alpha}_{N,2} \xrightarrow{k_{-2}} P^{\alpha}_{N,0} + T$,
Oct4 promoters	
$P^{\alpha}_{O,0} \xrightarrow{k_{+3}} P^{\alpha}_{O,1}$,	$P^{\alpha}_{O,1} \xrightarrow{k_{-3}} P^{\alpha}_{O,0}$,
Tcf3 promoters	
$P^{\alpha}_{,0} \xrightarrow{k_{+4}} P^{\alpha}_{,1}$,	$P^{\alpha}_{,1} \xrightarrow{k_{-4}} P^{\alpha}_{O,0}$,
mRNA transcription	
$P^{\alpha}_{N,1} \xrightarrow{k_5} P^{\alpha}_{N,1} + m_N$,	$P^{\alpha}_{O,1} \xrightarrow{k_6} P^{\alpha}_{O,1} + m_O$,
$P^{\alpha}_{,1} \xrightarrow{k_7} P^{\alpha}_{,1} + m$,	
Protein translation	
$m_N \xrightarrow{k} m_N + N$,	$m_O \xrightarrow{k} m_O + O$,
$m \xrightarrow{k_{10}} m + T$,	
mRNA degradation	
$m_N \xrightarrow{k_{11}} \varnothing$,	$m_O \xrightarrow{k_{12}} \varnothing$,
$m \xrightarrow{k_{12}} \varnothing$,	
β-catenin **dynamics**	
$\varnothing \xrightarrow{k_{+\beta}} \beta$,	$\beta \xrightarrow{k_{-\beta}} \varnothing$,
Protein degradation	
$N \xrightarrow{k_{14}} \varnothing$,	$\beta{:}N \xrightarrow{k_{-\beta}} \varnothing$,
$O \xrightarrow{k_{15}} \varnothing$,	$\beta{:}O \xrightarrow{k_{-\beta}} \varnothing$,
$T \xrightarrow{k_{16}} \varnothing$,	$\beta{:}O \xrightarrow{k_{-\beta}} \varnothing$,
$O{:}N \xrightarrow{k_{17}} \varnothing$,	
Complex formation	
$N + O \xrightarrow{k_{+1}} O{:}N$,	$O{:}N \xrightarrow{k_{-1}} N + O$,
$N + \beta \xrightarrow{k_{+1}} \beta{:}N$,	$\beta{:}N \xrightarrow{k_{-1}} N + \beta$,
$O + \beta \xrightarrow{k_{+20}} \beta{:}O$,	$\beta{:}O \xrightarrow{k_{-20}} O + \beta$,
$T + \beta \xrightarrow{k_{+21}} \beta{:}T$,	$\beta{:}T \xrightarrow{k_{-21}} T + \beta$,

And the parameter values used for the simulation of wild type ES cells under S+L and 2i+L pluripotency conditions are shown in Table A.12.

Par.	Value	Units	Par.	Value	Units
k_{+1}	2 (S+L)	h^{-1}	k_{11}	0.2	h^{-1}
	8 (2i+L)	h^{-1}	k_{12}	0.2	h^{-1}
k_{-1}	2	h^{-1}	k_{13}	0.2	h^{-1}
k_{+2}	0.005	$transc^{-1}\ h^{-1}$	k_{14}	6	h^{-1}
k_{-2}	2	$transc^{-1}\ h^{-1}$	k_{15}	0.4	h^{-1}
k_{+3}	0.4	h^{-1}	k_{16}	0.15	h^{-1}
k_{-3}	2	h^{-1}	k_{17}	0.3	h^{-1}
k_{+4}	4	h^{-1}	k_{+18}	0.4	$molec^{-1}\ h^{-1}$
k_{-4}	20	h^{-1}	k_{-18}	1	h^{-1}
k_5	50	$transc\ h^{-1}$	k_{+19}	0.1	$molec^{-1}\ h^{-1}$
k_6	50	$transc\ h^{-1}$	k_{-19}	0.4	h^{-1}
k_7	50	$transc\ h^{-1}$	k_{+20}	100	$molec^{-1}\ h^{-1}$
k_8	50	$molec\ transc^{-1}\ h^{-1}$	k_{-20}	0.02	h^{-1}
k_9	20	$molec\ transc^{-1}\ h^{-1}$	k_{+21}	20	$molec^{-1}\ h^{-1}$
k_{10}	8	$molec\ transc^{-1}\ h^{-1}$	k_{-21}	0.2	h^{-1}
			$k\ \_$	0.1	h^{-1}

Table A.12: **Parameter values for the TBON model.**

Simulation of mutant ES cells with the TBON model

Nanog mutant cells In the absence of Nanog, the pluripotent state is unstable (Muñoz Descalzo, Rué, Faunes, et al., 2013). In these simulations, the activity of both Nanog promoters is eliminated, while all other parameter values are taken from S+L. The TBON model is able to reproduce the increased Oct4 and β-catenin levels in both culture conditions (Figure 4.20a and ibid.) and predicts an increased number of free Oct4 molecules that is reduced upon stabilisation of β-catenin. In the TBON we consider that PD03 acts promoting Nanog expression; in the absence of Nanog, this means that PD03 would have little effect as we observe experimentally (ibid.).

β-catenin mutant cells. In this case we assume the following: in the heterozygous cells the supply of β-catenin is approximately halved ($\mu_+ = 150$ and $^2_+ = 0.1$), and all other parameter values are taken from 2i+L; in the mutant cells, β-catenin supply is completely removed ($k_+=0$), and all other parameter values are taken from 2i+L (Figure 4.20c-f).

Tcf3 mutant cells. To simulate Tcf3 mutant cells, the two Tcf3 promoters are removed from the system, while all other parameter values are taken from S+L. We additionally included observed experimental effects, in particular the reduction of Oct4 and β-catenin production ($k_{+3} = 0.2$ and $\mu_+ = 150$). One striking feature of the Tcf3 mutants is the existence a stable population of cells in the lower O/N ratio boundary, a population that does not normally exist due to the instability of free Nanog (Figure 4.20i). This suggests a scenario in which the absence of Tcf3 allows an increased stability of Nanog, therefore in our simulations we increased the stability of free Nanog ($k_{14} = 0.8$, Figure 4.20g-j). All other parameter values are taken from S+L.

Oct4 loss of function. ZHBTc4 cells are Oct4 heterozygous cells in which the functional allele has been modified so that it is inactive and the other one is under doxycycline control (Niwa, S. Masui, et al., 2002). Addition of doxycycline to a culture of these cells results in a rapid decrease in the levels of Oct4 (Figure 4.21a). For the simulations we consider only one Oct4 promoter. Upon start of doxycycline treatment, Oct4 transcription is halted ($k_6 = 0$), while all other parameter values are taken from S+L. The inability of maintaining the pluripotent state in these cells comes

from the fact the Nanog levels constantly decrease as Oct4 disappears from the system. The decrease of Oct4 that we observe experimentally is well reproduced by the model simulation (compare panels B,C with E,F from Figure 4.21a-c).

Oct4 overexpression. AG23191 cells overexpress Oct4 from a TetOFF system, so that removal of doxycycline from the culture leads to overexpression of Oct4 (Nishiyama, Xin, et al., 2009). We simulate the behaviour of the TBON in these cells by considering an additional gene of Oct4 with continuous expression ($P_{tet} \xrightarrow{k} P_{tet} + m_O$) with $k = 5$, while all other parameter values are taken from S+L or Chiron conditions. Extra production of Oct4 in S+L conditions leads to stabilisation of Nanog (Figure 4.21), which is not able to mop up all excess free Oct4 (Figure 4.21g). Extra β-catenin supply to the system via Chiron compensates the excess of Oct4 and reduces the levels of free protein to its normal condition (Figure 4.21d-g).

A.3 Details of the Boolean Network model

Boolean networks are discrete dynamical systems composed by individual units, called nodes, connected with each other by directed links. In a BN, each of the network nodes is assumed to be either active (1) or inactive (0), and to evolve in a discrete-time basis according to a set of node-specific logic rules. Specifically, in a BN with N nodes, $i = 1, \dots, N$, the state of each node i, x_i, evolves according to its corresponding logic rule, f_i, which depends on the current states of its k_i incident nodes, $x_i^1, \dots, x_i^{k_i}$:

$$x_i(t+1) = f_i(x_i^1(t), \dots, x_i^{k_i}(t)). \tag{A.7}$$

In the classical approach to the dynamics of BNs, all the states of the network are updated in parallel at every time iteration, resulting in the dynamical evolution of the activity of the nodes. Time thus evolves in a discrete manner, with no quantitative meaning being associated to it. Hence the model represents dynamical behaviours in terms of sequences of biochemical events rather than actual time, so that all references to time and frequency in what follows are given in terms of model iterations. For a given initial condition of the network, the states of all nodes evolve in a deterministic manner according to the iteration map defined by the logic rules.

1	EGF	25	alpha-q-R	49	Crk	73	WASP	97	Tiam	121	PDK1
2	ExtPump	26	alpha-s-R	50	RalGDS	74	PLD	98	Graf	122	RasGRF-GRP
3	Stress	27	TAO-1-2	51	PLC-g	75	Mekk1	99	p115RhoGEF	123	Raf
4	alpha-i-lig	28	Ca	52	Rap1	76	Mekk3	100	Talin	124	MLCK
5	alpha-q-lig	29	EGFR	53	PI3K	77	PLA2	101	B-Arrestin	125	MKPs
6	alpha-12-13-lig	30	AC	54	PKC	78	MLK1	102	RhoGDI	126	GRK
7	alpha-s-lig	31	Nck	55	Grb2	79	MLK2	103	p190RhoGAP	127	Tab-1-2
8	ECM	32	Gas	56	p120RasGAP	80	MLK3	104	RGS	128	Sos
9	IL1-TNF	33	Gaq	57	Src	81	CaMKK	105	Raf-DeP	129	TAK1
10	Akt	34	ASK1	58	PIP2-45	82	SHP2	106	DAG	130	Actin
11	Erk	35	Trafs	59	Fak	83	PTPa	107	PI5K	131	Rho
12	Rac	36	CaM	60	PIP-4	84	Ral	108	PKA	132	RhoK
13	Cdc42	37	PP2A	61	MKK3	85	Tpl2	109	NIK	133	Ras
14	SAPK	38	PI4K	62	MKK6	86	Cbp	110	Vinc	134	B-Parvin
15	p38	39	IP3R1	63	Palpha-q-R	87	Csk	111	PKC-primed	135	RalBP1
16	alpha-i-R	40	Mekk2	64	PIP3-345	88	AND-3-4	112	ARF	136	p90RSK
17	Integrins	41	Ga-12-13	65	Palpha-i-R	89	GCK	113	Sek1	137	PDE4
18	PTP1b	42	Gbg-12-13	66	Palpha-s-R	90	CaMK	114	Myosin	138	Raf-Loc
19	Cas	43	Gbg-i	67	PIP2-34	91	RKIP	115	MKK7	139	Raf-Rest
20	Gai	44	Gbg-s	68	cAMP	92	PLC-B	116	Mek		
21	IL1-TNFR	45	Gbg-q	69	Palpha-12-13-R	93	DOCK180	117	Arp-2-3		
22	alpha-12-13-R	46	Gab1	70	Shc	94	ILK	118	AA		
23	Trx	47	DGK	71	PAK	95	IP3	119	MLCP		
24	PTEN	48	PTPPEST	72	Mekk4	96	Pix-Cool	120	PA		

Table A.13: **Node labels of the annotated fibroblast network.** Specific information about the update rule for each node can be found at `http://mathbio.unomaha.edu/Database`.

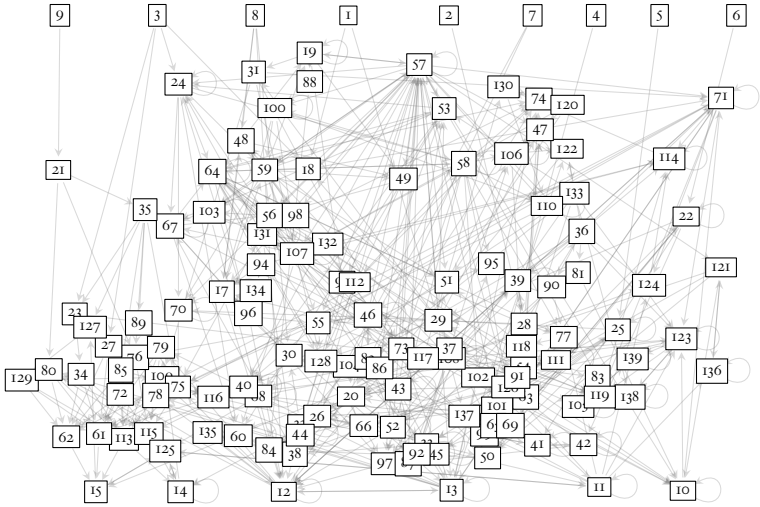

Figure A.1: **Annotated fibroblast network.** Node labels described in Table A.13.

Boolean network modelling

BNs, introduced by Kauffman in 1969, are defined as sets of dynamical units, called nodes, which are connected by interaction edges. The dynamic evolution of a Boolean network is characterised by logic operations which act upon binary state variables defined in each node. Thus, the state of every node i at time t, $x_i(t)$, is either active ($x_i(t) = 1$) or inactive ($x_i(t) = 0$) and all nodes are updated synchronously according to a set of node-specific logic rules (Kauffman, 1969). Here we will work with non-autonomous BNs, in which two types of nodes exist according to its dynamics: internal nodes and input nodes. In particular, every internal node i has an associated logic rule, f_i, which determines the new state $x_i(t)$ at time t from the states of its k_i incident nodes at time $t-1$:

$$x_i(t) = f_i(x_i^1(t-1), \ldots, x_i^{k_i}(t-1)) \tag{A.8}$$

The state of the input nodes, in contrast, does not depend on the state of other nodes but is given externally. Here we set these states using periodic and random value sequences. The update rules, $f_i, i = 1, \ldots, N$, can be specified either as a truth table or as a composition of basic logic functions (ANDs, ORs, NOTs …). Simulations obtained using these functions are, in general, more efficient than those obtained using truth tables. To speed up the simulation runtime, we transformed the original truth tables (Helikar et al., 2008) to the corresponding equivalent logic functions using an adapted version of the Quine-McCluskey algorithm (Quine, 1952). We implemented the Boolean network simulator in Python.

In order to generate the stochastic input sequences, we use Bernoulli distributions with success probability equal to q, which we define as the chatter level. Bernoulli distribution sequences are obtained using a standard Mersenne twister pseudo-random number generator (Matsumoto and Nishimura, 1998).

A.4 Mesoscopic model of bacterial cellular death

Model integration scheme and numerical details

The reaction diffusion model described in Chapter 7 was integrated in a square lattice of mesh size $\times = 500 \times 500$ and spacing $x, y = 10\mu m$ [early- and late-burst dynamics experiments (Figure 7.7)] or $x, y = 5\mu m$ [dilute single colony experiment (Figure 7.8)]. Thus, the integrated

domain spans an area of $|\Omega| = \quad x \quad y \quad ^2 = 25\,\mathrm{mm}^2$ in the early-/late-burst simulations and $|\Omega| = 6.25\,\mathrm{mm}^2$ in the single colony assay. Model equations were integrated using a second order Runge-Kutta (Heun) method (Garcia-Ojalvo and Sancho, 1999) with standard finite differences, written in custom-made C code [2].

[2] The code is publicly available from `https://bitbucket.org/prue/cdp`.

The numerical scheme used in the integration of the system thus approximates the quantities $\rho(n \quad t, (i \quad x, \quad y))$, $m(n \quad t, (i \quad x, \quad y))$, $w(n \quad t, (i \quad x, \quad y))$, and $k(n \quad t, (i \quad x, \quad y))$ by $\rho_{i,}(n)$, $m_{i,}(n)$, $w_{i,}(n)$, and $k_{i,}(n)$ respectively, where

$$\rho_{i,}(n+1) = \rho_{i,}(n) + \frac{t}{2}\left[R^{\rho}_{i,}(n) + R^{\rho}_{i,}(n) + D^{\rho}_{i,}(n) + D^{\rho}_{i,}(n)\right] , \tag{A.9a}$$

$$m_{i,}(n+1) = m_{i,}(n) + \frac{t}{2}\left[R^{m}_{i,}(n) + R^{m}_{i,}(n)\right] , \tag{A.9b}$$

$$w_{i,}(n+1) = w_{i,}(n) + \frac{t}{2}\left[\rho_{i,}(n) + \rho_{i,}(n)\right] , \tag{A.9c}$$

$$K(n+1) = K(n) + \frac{t}{2}\left[R^{K}(n) + R^{K}(n)\right] , \tag{A.9d}$$

$$s_{i,}(n+1) = s_{i,}(n) + \frac{t}{2}\left[R^{s}_{i,}(n) + R^{s}_{i,}(n)\right] , \tag{A.9e}$$

and

$$D_{i,}(n) = D_0^{\ -m_{i,}(n)/k_d}, \tag{A.10a}$$

$$k_{i,}(n) = \frac{K(n) + k_1 \cdot \ w_{i,}(n)/k_k^{\ n_k}}{1 + \ w_{i,}(n)/k_k^{\ n_k}}, \tag{A.10b}$$

$$R^{\rho}_{i,}(n) = \alpha_{i,} \cdot \rho_{i,}(n) \ \ 1 - \frac{\rho_{i,}(n)}{k_{i,}(n)} \ \ , \tag{A.10c}$$

$$D^{\rho}_{i,}(n) = + \frac{D_{i,}(n)}{x \ \ y}\Big(\rho_{i+1,}(n) + \rho_{i-1,}(n) + \rho_{i,+1}(n) + \rho_{i,-1}(n) - 4\rho_{i,}(n)\Big) \tag{A.10d}$$

$$- \frac{D_{i,}(n)}{4 \ \ x \ \ yk_d}\Big(\ \rho_{i+1,}(n) - \rho_{i-1,}(n) \quad m_{i+1,}(n) - m_{i-1,}(n) \tag{A.10e}$$

$$+ \ \ \rho_{i,+1}(n) - \rho_{i,-1}(n) \quad m_{i,+1}(n) - m_{i,-1}(n) \ \Big), \tag{A.10f}$$

$$R^{m}_{i,}(n) = \lambda_m \frac{\rho_{i,}(n)/k_m^{\ n_m}}{1 + \ \rho_{i,}(n)/k_m^{\ n_m}} \tag{A.10g}$$

$$R^{K}(n) = -\lambda_k \langle \rho_{i,}(n) \rangle_{i,}, \tag{A.10h}$$

$$R^{s}_{i,}(n) = \alpha_s \frac{\rho_{i,}(n)/k_s^{\ n_s}}{1 + \ \rho_{i,}(n)/k_s^{\ n_s}} - \lambda_s s_{i,}(n), \tag{A.10i}$$

$$\rho_{i,}(n) = \rho_{i,}(n) + \ \ t \cdot \ R^{\rho}_{i,}(n) + D^{\rho}_{i,}(n) \ \ , \tag{A.10j}$$

$$m_{i,}(n) = m_{i,}(n) + \ \ t \cdot R^{m}_{i,}(n), \tag{A.10k}$$

$$w_{i,}(n) = w_{i,}(n) + \ \ t \cdot \rho_{i,}(n), \tag{A.10l}$$

$$K(n) = K(n) + \ \ t \cdot R^{K}(n), \tag{A.10m}$$

$$s_{i,}(n) = s_{i,}(n) + \ \ t \cdot R^{s}_{i,}(n), \tag{A.10n}$$

configure the Euler step and

$$D_{i,}(n) = D_0^{\ -m_{i,}(n)/k_d}, \tag{A.100}$$

$$k_{i,}(n) = \frac{K(n) + k_1 \cdot \ w_{i,}(n)/k_k^{\ n_k}}{1 + \ w_{i,}(n)/k_k^{\ n_k}}, \tag{A.10p}$$

$$R^{\rho}_{i,}(n) = \alpha_{i,}\,\rho_{i,}(n)\ \ 1 - \frac{\rho_{i,}(n)}{k_{i,}(n)}\ \ , \tag{A.10q}$$

$$D^{\rho}_{i,}(n) = +\frac{D_{i,}(n)}{x\ \ y}\Big(\rho_{i+1,}(n) + \rho_{i-1,}(n) + \rho_{i,+1}(n) + \rho_{i,-1}(n) - 4\rho_{i,}(n)\Big) \tag{A.10r}$$

$$- \frac{D_{i,}(n)}{4\ \ x\ \ yk_d}\Big(\ \rho_{i+1,}(n) - \rho_{i-1,}(n)\quad m_{i+1,}(n) - m_{i-1,}(n) \tag{A.10s}$$

$$+\ \ \rho_{i,+1}(n) - \rho_{i,-1}(n)\quad m_{i,+1}(n) - m_{i,-1}(n)\ \Big) \tag{A.10t}$$

$$R^{m}_{i,}(n) = \lambda_m \frac{\rho_{i,}(n)/k_m^{\ n_m}}{1 + \ \rho_{i,}(n)/k_m^{\ n_m}} \tag{A.10u}$$

$$R^{K}(n) = -\lambda_k \langle \rho_{i,}(n) \rangle_{i,}, \tag{A.10v}$$

$$R^{s}_{i,}(n) = \alpha_s \frac{\rho_{i,}(n)/k_s^{\ n_s}}{1 + \ \rho_{i,}(n)/k_s^{\ n_s}} - \lambda_s s_{i,}(n), \tag{A.10w}$$

Model integration conditions

Random growth profile and initial conditions were set up using the Mersenne Twister pseudo-random number generator from the open-source GNU Scientific Library. The integration time step used was $t = 0.01\,\mathrm{h}$ ($t = 0.0025\,\mathrm{h}$ for the single colony simulation) and periodic boundary conditions were assumed. Simulation time was 40 h. As initial condition for the cell density in the early-/late- burst dynamics experiment, 2500 micro-colonies of diameter 30 μm were randomly placed within a ring of high density (coffee ring (Figure 7.6) of diameter 2.5 mm and amplitude 0.0005. In the single colony experiment, a single colony was plated at the centre of the grid. Waste and matrix fields where started at zero value and the initial global carrying capacity was set to $k_0 = 5\,\mathrm{cells\,mm^{-2}}$.

The parameter values used in the simulations of the model presented in Chapter 7 are summarised in Table A.14 below.

Cluster analysis of simulation results

We applied the clustering procedure of the experimental data to the model simulations (detailed in Asally et al., 2012)). The 500 × 500 sim-

Par.	Value	Units	Description
α_0	1	cells $\mu m^{-2} h^{-1}$	Maximum logistic growth rate
k_0	5	cells μm^{-2}	Initial maximum carrying capacity
k_1	1.5	cells μm^{-2}	Maximum carrying capacity at saturating waste levels
γ_k	0.05	h^{-1}	Carrying capacity decay coefficient
k_k	50	cells h μm^{-2}	Local carrying capacity waste threshold
n_k	2	–	Local carrying capacity waste Hill coefficient
D_0	2500	cells h^{-1}	Maximum cell diffusivity
λ_m	10	h^{-1}	Matrix production
k_m	0.05	cells μm^{-2}	Matrix accumulation density threshold
	∞		(epsH mutant)
n_m	2	–	Matrix accumulation density Hill coefficient
α_s	20	molec h^{-1}	Sytox accumulation rate
k_s	3.75	cells μm^{-2}	Sytox accumulation density threshold
n_s	3	–	Sytox accumulation density Hill coefficient
λ_s	20	h^{-1}	Sytox degradation rate

Table A.14: **Parameter values used in the simulations of wild type and *epsH* mutant colony dynamics.**

ulation grid was cropped to 360×270 points, an area corresponding to the experimental field of view (3.6 mm × 2.7 mm at $x = 10\,\mu m$). The centred cropping procedure also ensures that the resulting area is not affected by the periodic boundary conditions used in the numerical integration scheme. Frames were sampled every 30 min for the length of the entire simulation (40 h). This data was then fed to the *k*-means clustering algorithm. Indeed, the time thresholds for the early-/late-burst classification were set to the experimental ones (18.7 h for the wild type) and 23.3 h for the epsH mutant). The amplitude threshold for the late burst Sytox signal was set to 0.002 in all cases. For further details see Asally et al. (ibid.).

B

Bioinformatics analysis of Oct4 and Nanog targets

A prediction of the TBON model is that an excess of Oct4 in ES cells is inversely related to their degree of pluripotency (see Section 4.3 for details and Figure 4.19). Our working hypothesis is that, when not interacting with Nanog, Oct4 promotes differentiation through the interactions with lineage determinants, and the activation of markers and enforcers of lineage commitment. This notion finds some support in several ES cell expression (K. M. Loh and Lim, 2011; Nishiyama, Xin, et al., 2009; A. Sharov, S. Masui, et al., 2008) and DNA binding (X. Chen et al., 2008) data sets (see in particular `http://www.maayanlab.net/ESCAPE`, a database integrating high-throughput data of pluripotency in m/hESCs). In search of further evidence and support of our contention, we performed a comprehensive bioinformatics analysis of publicly available data looking at the genes regulated by Nanog and Oct4, jointly or individually and, most importantly, looking at the functional categories associated with these genes (gene-ontology analysis). To do this, we fused genomic gene expression profiling data from loss of function (LoF) and gain of function (GoF) experiments of Nanog and Oct4 with genome-wide chromatin-immunoprecipitation (ChIP) evidence of DNA binding. Table B.1 summarises the data sets included in the reanalysis.

As the experiments were performed in different platforms, the first step in our analysis was to identify genes that were tested in all experiments. Using the Mouse Genome Informatics (MGI) Marker Accession IDs or, alternatively, their Official Gene Symbol, we matched 17 927 genes that were then tested for expression changes in both Nanog/Oct4 LoF and GoF microarray experiments. We applied specific statistical criteria to this list of genes (background) based on filtering by fold-change and thresholding p-values with the false-discovery rate method, to identify differentially expressed genes (see Table B.1 for details of specific

Experiment	Publication	Description	
Oct4 and Nanog Gain of Function	Nishiyama, Xin, et al. (2009)	Conditions	MC1 cells with Tet-inducible Oct4 or Nanog transgene (knocked-in for ROSA-TET) in S+L conditions. Doxicycline removed 48h before harvesting.
		Platform	Agilent Mouse 44k mircorarray.
		Data source	NIA Array Analysis.[1]
		Stat. criterion	Fold-Change greater than 1.5.
Nanog Loss of Function	Y.-H. Loh et al. (2006)	Conditions	E14 cells in S+L transfected with a plasmid expressing a Nanog siRNA sequence.
		Platform	Affymetrix Mouse Genome 430 2.0 microarray.
		Data source	Y.-H. Loh et al. (2006).
		Stat. criterion	Same Significance Analysis of Microarray (SAM) as in Y.-H. Loh et al. (2006), with median FDR lower than 0.001.
Oct4 Loss of Function	A. Sharov, S. Masui, et al. (2008)	Conditions	ZHBTc4 cells in S+Lwith Tetracycline. Harvested at 0, 3, 6, 12, 24, 48, 72, 96 and 108 hours.
		Platform	Agilent Mouse 44k mircorarray.
		Data source	NIA Array Analysis.[1]
		Stat. criterion	Genes with a Fold-Change of the response magnitude (max. FC of the time-course) greater than 1.5, and FDR<0.05.
Nanog and Oct4 binding to DNA	X. Chen et al. (ibid.)	Conditions	E14 cells in S+L. Chromatin Immunoprecipitation against Oct4/Nanog coupled with massively parallel short-tag-based sequencing (ChIP-seq) was used to map DNA binding locations.
		Platform	Affymetrix Mouse Genome 430 2.0 microarray.
		Data source	X. Chen et al. (ibid.).
		Stat. criterion	Genes with a positive TF-Gene Association Score (ibid.).

[1] http://lgsun.grc.nia.nih.gov/ANOVA/

Table B.1: **Data sets included in the Bioinformatics analysis**

statistical criteria used in each data set). In addition, we used ChIP-Seq data (X. Chen et al., 2008, and Table B.1) to tag genes as being bound by Nanog and/or Oct4. Figure B.1a shows the number of genes detected with each method.

We consider a gene to be an Oct4 (Nanog) target if it is differentially expressed in either LoF or GoF experiments and, in addition, if there is evidence that Oct4 (Nanog) binds close to its transcription start, as established by a positive TF-Gene Association Score computed in (ibid.). Using this criterion we found 2716 Oct4 and 1328 Nanog target genes (see Figure B.1a) with a significant overlap of target genes (438 common targets, Fisher's exact test $p-\text{value} < 1 \times 10^{-15}$, see Figure B.1b), which we deem to be jointly regulated, but also a large amount of genes that are specific targets of either Nanog (890, 67 %) or Oct4 (2278, 83 %).

Given the significant but incomplete overlap of Nanog and Oct4 target genes detected, we wanted to explore if each of the three groups identified is associated with a particular biological function. To that end we used gene annotations from MGI (www.informatics.jax.org) and performed a standard Gene Ontology (GO) analysis with p-values of Biological Process GO terms assessed by Fisher's exact test. Figure B.1c shows the general Biological Process terms that are more strongly over-represented by Oct4 and Nanog common and specific target genes, with its associated significance value (measured by $-\log_{10} p$, where p is the corresponding p-value). This Figure shows that, while Cellular and Metabolic processes are highly over-represented by Oct4 targets, Developmental processes are at least as well represented amidst the Oct4 targets as they are amidst the Nanog targets (756 of 2716 genes, $p-\text{value} = 1.1 \cdot 10^{-15}$ for Oct4 while for Nanog 391 of 1328 genes, $p-\text{value} = 2.2 \cdot 10^{-11}$). This analysis indicates that while both factors target a large number of genes involved in developmental processes, Oct4 also has a very large number of targets associated with basic biochemical and cellular events. The significance of this targeting will have to be addressed experimentally. One other feature of our analysis is that neither Nanog nor Oct4 targets are associated neither with late differentiation events nor with specialised activities of differentiated cells.

Given the large number of GO terms over-represented by, in particular, the list of Oct4 target genes, we focused our further analysis only on those terms of particular interest in pluripotency and early lineage differentiation. Specifically, we analysed GO terms that in their description contained at least one of the following words: "development", "differ-

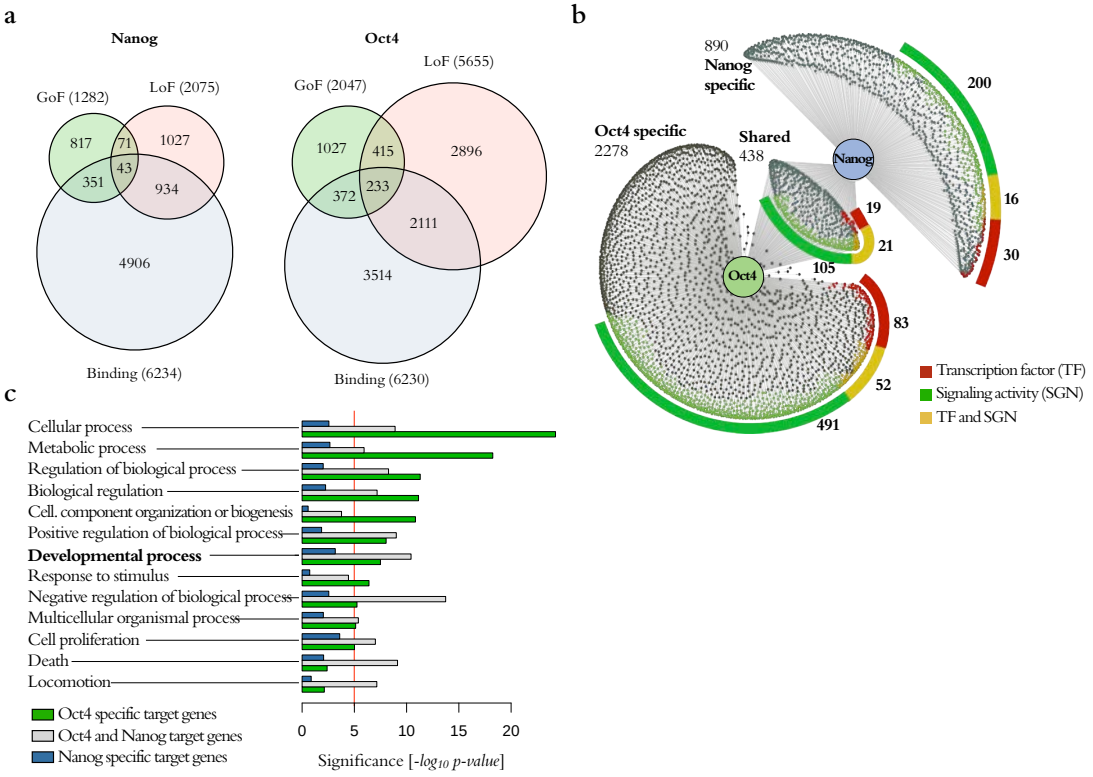

Figure B.1: **Results of the bioinformatic analysis.** (a) Venn diagrams describing the sizes of the sets of Nanog and Oct target genes detected in LoF, GoF and DNA binding experiments (see text for details). (b) Schematic representation of the shared and specific Nanog and Oct4 targets. Genes identified as either transcription factors or involved in signaling are highlighted. (c) Gene Ontology cathegories enriched for the sets of Nanog- and Oct4-specific target genes as well as for shared targets.

entiation", "ectoderm", "embryo", "endoderm", "epithelial", "mesoderm", "morphogenesis", "neural", "organ", "stem" and "tissue". We further filtered the search to those GO terms with at least 100 genes annotated, a depth on the GO graph of at most 3 (this is, either their parents or the parents of their parents are the "Biological Process" root term) and that at least one of the $p-$values of Nanog or Oct4 target genes was below 10^{-6}. This procedure yielded a meaningful and hierarchical list of terms, which is shown in Figure B.1c, and that reveals that the set of Oct4 targets is, in part, made up of many development/differentiation-related genes, which, unlike Nanog targets, are not involved in very specific differentiation processes.

To conclude the comparison of Oct4 and Nanog targets, we tried to identify those genes with transcription factor (TF) and signalling (SGN) activity. For this purpose, we rely on the GO Molecular Function annotation from MGI (term GO:0003700: "sequence-specific DNA binding transcription factor activity") or the Biological Process annotation (term GO:0023052: "signalling"). We were able to identify 175 Oct4 target genes and 86 Nanog target genes with TF activity, of which 40 are common targets. In the case of SGN activity, we recognised 669 Oct4 and

342 Nanog targets involved in signalling activity, 126 of which were shared (TF and SGN activity annotated in Supplementary File 1).

Both TF and SGN activity was significantly over-represented in Nanog and Oct4 target genes (Fisher exact test $p-$ values $< 2 \cdot 10^{-10}$ for TFs in Oct4, $< 2 \cdot 10^{-5}$ for TFs in Nanog, $< 2 \cdot 10^{-9}$ for SGN in Oct4, $< 5 \cdot 10^{-7}$ for SGN in Nanog). Inspection of the list of TFs (Muñoz Descalzo, Rué, Faunes, et al., 2013) reveals striking functional differences between the three groups: the jointly regulated TFs include elements of the core of the pluripotency network (Esrrb, Nanog, Oct4, Klf4, 6, Otx2), while targets of Nanog, and more clearly Oct4, include genes involved in lineage specification (members of the En, Fox, Gata, Hox, T-box, Zic families). The list of signalling molecules (ibid.) also highlights the differences between the different targets: joint targets are not significantly enriched in any particular signalling class, though we notice the presence of Fgf4 and Fgfr2 in this list, both of which have been implicated in fate decisions in the early embryo (Guo et al., 2011). When looking at the list of specific targets of Nanog and Oct4, we observe an interesting symmetry with several Fgf genes as targets of Nanog, and many Fgf receptors and elements of Ras/ERK signalling as targets of Oct4, also including a number of Wnt genes (Muñoz Descalzo, Rué, Faunes, et al., 2013).

Taken together, the analysis of GO terms and the identified TFs and SGN component targets, support the notion that an excess of Oct4 over Nanog will lead to the activation of differentiation targets.

Contributions

Journal articles within the scope of this Thesis

- S. Muñoz Descalzo, P. Rué, J. Garcia-Ojalvo, et al. (Dec. 2012). "Correlations Between the Levels of Oct4 and Nanog as a Signature for Naïve Pluripotency in Mouse Embryonic Stem Cells". In: *Stem Cells* 30.12, pp. 2683–2691. DOI: 10.1002/stem.1230
- M. Asally, M. Kittisopikul, P. Rué, et al. (Nov. 2012). "Localized cell death focuses mechanical forces during 3D patterning in a biofilm". In: *Proceedings of the National Academy of Sciences* 109.46, pp. 18891–18896. DOI: 10.1073/pnas.1212429109
- P. Rué and J. Garcia-Ojalvo (May 2011). "Gene circuit designs for noisy excitable dynamics". In: *Mathematical Biosciences* 231.1, pp. 90–97. DOI: 10.1016/j.mbs.2011.02.013
- P. Rué, G. M. Süel, and J. Garcia-Ojalvo (June 2011). "Optimizing periodicity and polymodality in noise-induced genetic oscillators". In: *Physical Review E* 83.6, p. 61904. DOI: 10.1103/PhysRevE.83.061904
- P. Rué, N. Domedel-Puig, J. Garcia-Ojalvo, et al. (2012). "Integration of cellular signals in chattering environments". In: *Progress in Biophysics and Molecular Biology*. DOI: 10.1016/j.pbiomolbio.2012.05.003
- P. Rué, A. J. Pons, N. Domedel-Puig, et al. (Jan. 2010). "Relaxation dynamics and frequency response of a noisy cell signaling network". In: *Chaos* 20.4, p. 45110. DOI: 10.1063/1.3524908
- N. Domedel-Puig, P. Rué, A. J. Pons, et al. (Dec. 2011). "Information Routing Driven by Background Chatter in a Signaling Network". In: *PLoS Comput Biol* 7.12, e1002297. DOI: 10.1371/journal.pcbi.1002297
- S. Muñoz Descalzo, P. Rué, F. Faunes, et al. (2013). "A competitive protein interaction network buffers Oct4-mediated differentiation to promote pluripotency in embryonic stem cells". In: *Mol Syst Biol (under review)*

Review articles within the scope of this Thesis

- P. Rué and J. Garcia-Ojalvo (May 2013). "Modeling Gene Expression in Time and Space". In: *Annual Review of Biophysics* 42.1, pp. 605–627. DOI: 10.1146/annurev-biophys-083012-130335

Contributed software

bnsim a simple Boolean network simulator with synchronous and asynchronous update schemes and implementation of chatter. Available at https://code.google.com/p/bnsim

dssim an efficient and versatile software for the simulation of stochastic chemical reactions. Available at https://code.google.com/p/dstochsim

cdp numerical scheme for the integration of the reaction diffusion model presented in Asally et al. (2012). Available at https://bitbucket.org/prue/cdp

Bibliography

Abu-Remaileh, M., A. Gerson, M. Farago, G. Nathan, I. Alkalay, S. Zins Rousso, M. Gur, A. Fainsod, and Y. Bergman (Oct. 2010). "Oct-3/4 regulates stem cell identity and cell fate decisions by modulating Wnt/[beta]-catenin signalling". In: *EMBO J* 29.19, pp. 3236–3248 (cit. on pp. 33, 86).

Alonso, D., A. J. McKane, and M. Pascual (June 2007). "Stochastic amplification in epidemics". In: *J R Soc Interface* 4.14, pp. 575–582. DOI: 10.1098/rsif.2006.0192 (cit. on p. 70).

Asally, M., M. Kittisopikul, P. Rué, Y. Du, Z. Hu, T. Çagatay, A. B. Robinson, H. Lu, J. Garcia-Ojalvo, and G. M. Süel (Nov. 2012). "Localized cell death focuses mechanical forces during 3D patterning in a biofilm". In: *Proceedings of the National Academy of Sciences* 109.46, pp. 18891–18896. DOI: 10.1073/pnas.1212429109 (cit. on pp. 39, 140, 145, 149, 176, 177, 185).

Ashall, L., C. A. Horton, D. E. Nelson, P. Paszek, C. V. Harper, K. Sillitoe, S. Ryan, D. G. Spiller, J. F. Unitt, D. S. Broomhead, D. B. Kell, D. A. Rand, V. Sée, and M. R. H. White (Apr. 2009). "Pulsatile stimulation determines timing and specificity of {NF- B}-dependent transcription". In: *Science* 324.5924, pp. 242–246. DOI: 10.1126/science.1164860 (cit. on p. 122).

Atkinson, M. R., M. A. Savageau, J. T. Myers, and A. J. Ninfa (2003). "Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in Escherichia coli". In: *Cell* 113, pp. 597–607 (cit. on pp. 45, 56).

Bauer, A. L., T. L. Jackson, Y. Jiang, and T. Rohlf (June 2010). "Receptor cross-talk in angiogenesis: mapping environmental cues to cell phenotype using a stochastic, Boolean signaling network model". In: *J. theor. Biol.* 264.3, pp. 838–846. DOI: 10.1016/j.jtbi.2010.03.025 (cit. on p. 36).

Berg, D. L. C. van den, T. Snoek, N. P. Mullin, A. Yates, K. Bezstarosti, J. Demmers, I. Chambers, and R. A. Poot (Apr. 2010). "An Oct4-Centered Protein Interaction Network in Embryonic Stem Cells". In: *Cell stem cell* 6.4, pp. 369–381 (cit. on pp. 77, 80, 86).

Berg, D. L. C. van den, W. Zhang, A. Yates, E. Engelen, K. Takacs, K. Bezstarosti, J. Demmers, I. Chambers, and R. A. Poot (Oct. 2008). "Estrogen-Related Receptor Beta Interacts with Oct4 To Positively Regulate Nanog Gene Expression". In: *Molecular and Cellular Biology* 28.19, pp. 5986–5995 (cit. on p. 77).

Bornholdt, S. and K. Sneppen (Nov. 2000). "Robustness as an evolutionary principle". In: *Proc. Royal Soc. B* 267.1459, pp. 2281–2286. DOI: 10.1098/rspb.2000.1280 (cit. on pp. 35, 36).

Bornholdt, S. (Aug. 2008). "Boolean network models of cellular regulation: prospects and limitations". In: *J R Soc Interface* 5 Suppl 1, S85–94. DOI: 10.1098/rsif.2008.0132.focus (cit. on pp. 35, 105).

Branda, S. S., J. E. González-Pastor, S. Ben-Yehuda, R. Losick, and R. Kolter (Sept. 2001). "Fruiting body formation by Bacillus subtilis". In: *Proceedings of the National Academy of Sciences* 98.20, pp. 11621–11626 (cit. on p. 38).

Buckley, S. M., B. Aranda-Orgilles, A. Strikoudis, E. Apostolou, E. Loizou, K. Moran-Crusio, C. L. Farnsworth, A. A. Koller, R. Dasgupta, J. C. Silva, M. Stadtfeld, K. Hochedlinger, E. I. Chen, and I. Aifantis (Dec. 2012). "Regulation of Pluripotency and Cellular Reprogramming by the Ubiquitin-Proteasome System". In: *Cell stem cell* 11.6, pp. 783–798 (cit. on p. 79).

Çagatay, T., M. Turcotte, M. B. Elowitz, J. Garcia-Ojalvo, and G. M. Süel (2009). "Architecture-Dependent Noise Discriminates Functionally Analogous Differentiation Circuits". In: *Cell* 139.3, pp. 512–522 (cit. on pp. 45, 47, 49, 50, 56, 61, 70, 154).

Campbell, P. A., C. Perez-Iratxeta, M. A. Andrade-Navarro, and M. A. Rudnicki (June 2007). "Oct4 Targets Regulatory Nodes to Modulate Stem Cell Function". In: *PLoS ONE* 2.6, e553 (cit. on p. 76).

Carey, B. W., S. Markoulaki, J. H. Hanna, D. A. Faddah, Y. Buganim, J. Kim, K. Ganz, E. J. Steine, J. P. Cassady, M. P. Creyghton, G. G. Welstead, Q. Gao, and R. Jaenisch (Dec. 2011). "Reprogramming Factor Stoichiometry Influences the Epigenetic State and Biological Properties of Induced Pluripotent Stem Cells". In: *Cell Stem Cell* 9.6, pp. 588–598. DOI: http://dx.doi.org/10.1016/j.stem.2011.11.003 (cit. on p. 84).

Chambers, I., J. Silva, D. Colby, J. Nichols, B. Nijmeijer, M. Robertson, J. Vrana, K. Jones, L. Grotewold, and A. Smith (Dec. 2007). "Nanog safeguards pluripotency and mediates germline development". In: *Nature* 450.7173, pp. 1230–1234 (cit. on pp. 30, 85, 95).

Chambers, I. and S. R. Tomlinson (July 2009). "The transcriptional foundation of pluripotency". In: *Development* 136.14, pp. 2311–2322 (cit. on p. 30).

Chen, X., H. Xu, P. Yuan, F. Fang, M. Huss, V. B. Vega, E. Wong, Y. L. Orlov, W. Zhang, J. Jiang, Y.-H. Loh, H. C. Yeo, Z. X. Yeo, V. Narang, K. R. Govindarajan, B. Leong, A. Shahab, Y. Ruan, G. Bourque, W.-K. Sung, N. D. Clarke, C.-L. Wei, and H.-H. Ng (June 2008). "Integration of External Signaling Pathways with the Core Transcriptional Network in Embryonic Stem Cells". In: *Cell* 133.6, pp. 1106–1117. DOI: http://dx.doi.org/10.1016/j.cell.2008.04.043 (cit. on pp. 179–181).

Chew, J.-L., Y.-H. Loh, W. Zhang, X. Chen, W.-L. Tam, L.-S. Yeap, P. Li, Y.-S. Ang, B. Lim, P. Robson, and H.-H. Ng (July 2005). "Reciprocal Transcriptional Regulation of Pou5f1 and Sox2 via the Oct4/Sox2 Complex in Embryonic Stem Cells". In: *Molecular and Cellular Biology* 25.14, pp. 6031–6046 (cit. on pp. 76, 77).

Chickarmane, V., C. Troein, U. A. Nuber, H. M. Sauro, and C. Peterson (Sept. 2006). "Transcriptional Dynamics of the Embryonic Stem Cell Switch". In: *PLoS Comput Biol* 2.9, e123 (cit. on p. 31).

Conrad, E., A. E. Mayo, A. J. Ninfa, and D. B. Forger (2008). "Rate constants rather than biochemical mechanism determine behaviour of genetic clocks". In: *J R Soc Interface* 5.Suppl 1, S9–15 (cit. on p. 56).

Costa, Y., J. Ding, T. W. Theunissen, F. Faiola, T. A. Hore, P. V. Shliaha, M. Fidalgo, A. Saunders, M. Lawrence, S. Dietmann, S. Das, D. N. Levasseur, Z. Li, M. Xu, W. Reik, J. C. R. Silva, and J. Wang (Mar. 2013). "NANOG-dependent function of TET1 and TET2 in establishment of pluripotency". In: *Nature* 495.7441, pp. 370–374 (cit. on p. 33).

Cotterell, J. and J. Sharpe (2010). "An atlas of gene regulatory networks reveals multiple three-gene mechanisms for interpreting morphogen gradients". In: *Mol Syst Biol* 6, p. 425 (cit. on p. 56).

Covert, M. W., E. M. Knight, J. L. Reed, M. J. Herrgard, and B. O. Palsson (2004). "Integrating high-throughput and computational data elucidates bacterial networks". In: *Nature* 429.6987, pp. 92–96 (cit. on p. 36).

Danino, T., O. Mondragón-Palomino, L. Tsimring, and J. Hasty (2010). "A synchronized quorum of genetic clocks". In: *Nature* 463, pp. 326–330 (cit. on p. 45).

Davidich, M. I. and S. Bornholdt (Jan. 2008). "Boolean network model predicts cell cycle sequence of fission yeast". In: *Plos ONE* 3.2, e1672. DOI: 10.1371/journal.pone.0001672 (cit. on p. 36).

Derrida, B. and Y. Pomeau (Jan. 1986). "Random networks of automata: a simple annealed approximation". In: *Europhysics Letters* 1, pp. 45–+. DOI: 10.1209/0295-5075/1/2/001 (cit. on p. 35).

Domedel-Puig, N., P. Rué, A. J. Pons, and J. Garcia-Ojalvo (Dec. 2011). "Information Routing Driven by Background Chatter in a Signaling Network". In: *PLoS Comput Biol* 7.12, e1002297. DOI: 10.1371/journal.pcbi.1002297 (cit. on pp. 106, 112, 113, 118, 185).

Drossel, B. (July 2005). "Number of attractors in random Boolean networks". In: *Phys. Rev. E* 72.1, p. 16110. DOI: 10.1103/PhysRevE.72.016110 (cit. on p. 35).

Eldar, A. and M. B. Elowitz (Sept. 2010). "Functional roles for noise in genetic circuits". In: *Nature* 467.7312, pp. 167–173. DOI: 10.1038/nature09326 (cit. on p. 70).

Epstein, A. K., B. Pokroy, A. Seminara, and J. Aizenberg (Jan. 2011). "Bacterial biofilm shows persistent resistance to liquid wetting and gas penetration". In: *Proceedings of the National Academy of Sciences* 108.3, pp. 995–1000 (cit. on p. 38).

Ermentrout, B., Author, A. Mahajan, and Reviewer (2003). "Simulating, Analyzing, and Animating Dynamical Systems: A Guide to XPPAUT for Researchers and Students". In: *Applied Mechanics Reviews* 56.4, B53–B53. DOI: doi:10.1115/1.1579454 (cit. on p. 159).

Faunes, F., P. Hayward, S. Muñoz Descalzo, S. S. Chatterjee, T. Balayo, J. Trott, A. Christoforou, A. Ferrer-Vaquer, A.-K. Hadjantonakis, R. Dasgupta, and A. Martinez Arias (Mar. 2013). "A membrane-associated -catenin/Oct4 complex correlates with ground-state pluripotency in mouse embryonic stem cells". In: *Development* 140.6, pp. 1171–1183 (cit. on pp. 33, 79, 85, 86, 90, 92).

Faure, A., A. Naldi, C. Chaouiya, and D. Thieffry (2006). "Dynamical analysis of a generic Boolean model for the control of the mammalian cell cycle". In: *Bioinformatics* 22.14, e124–131. DOI: 10.1093/bioinformatics/btl210 (cit. on p. 36).

Festuccia, N., R. Osorno, F. Halbritter, V. Karwacki-Neisius, P. Navarro, D. Colby, F. Wong, A. Yates, S. Tomlinson, and I. Chambers (Oct. 2012). *Esrrb Is a Direct Nanog Target Gene that Can Substitute for Nanog Function in Pluripotent Cells* (cit. on p. 76).

Fidalgo, M., F. Faiola, C.-F. Pereira, J. Ding, A. Saunders, J. Gingold, C. Schaniel, I. R. Lemischka, J. C. R. Silva, and J. Wang (Oct. 2012). "Zfp281 mediates Nanog autorepression through recruitment of the NuRD complex and inhibits somatic cell reprogramming". In: *Proceedings of the National Academy of Sciences* 109.40, pp. 16202–16207 (cit. on pp. 80, 86).

Fischer-Friedrich, E., G. Meacci, J. Lutkenhaus, H. Chaté, and K. Kruse (2010). "Intra- and intercellular fluctuations in Min-protein dynamics decrease with cell length". In: *Proc Natl Acad Sci USA* 107, pp. 6134–6139 (cit. on p. 60).

Gang, H., T. Ditzinger, C. Z. Ning, and H. Haken (1993). "Stochastic resonance without external periodic force". In: *Phys. Rev. Lett.* 71, pp. 807–810 (cit. on pp. 29, 60).

Garcia-Ojalvo, J. and A. Martinez Arias (Dec. 2012). "Towards a statistical mechanics of cell fate decisions". In: *Current Opinion in Genetics & Development* 22.6, pp. 619–626. DOI: http://dx.doi.org/10.1016/j.gde.2012.10.004 (cit. on p. 95).

Garcia-Ojalvo, J. and J. M. Sancho (1999). *Noise in Spatially Extended Systems (Institute for Nonlinear Science)*. Springer (cit. on pp. 142, 174).

Gillespie, D. T. (1977). "Exact stochastic simulation of coupled chemical reactions". In: *The Journal of Physical Chemistry* 81.25, pp. 2340–2361. DOI: 10.1021/j100540a008 (cit. on pp. 47, 62, 74).

Glauche, I., M. Herberg, and I. Roeder (June 2010). "Nanog Variability and Pluripotency Regulation of Embryonic Stem Cells - Insights from a Mathematical Model Analysis". In: *PLoS ONE* 5.6, e11238 (cit. on p. 31).

Goldbeter, A. (Nov. 2002). "Computational approaches to cellular rhythms". In: *Nature* 420.6912, pp. 238–245. DOI: 10.1038/nature01259 (cit. on p. 28).

Gonze, D., J. Halloy, and A. Goldbeter (2002). "Robustness of circadian rhythms with respect to molecular noise". In: *Proc Natl Acad Sci USA* 99.2, pp. 673–678. DOI: 10.1073/pnas.022628299 (cit. on p. 28).

Greil, F. and B. Drossel (July 2005). "Dynamics of Critical Kauffman Networks under Asynchronous Stochastic Update". In: *Physical Review Letters* 95.4, p. 48701 (cit. on p. 113).

Guantes, R. and J. F. Poyatos (2006). "Dynamical principles of two-component genetic oscillators". In: *PLoS Comput Biol* 2, e30 (cit. on p. 43).

Guo, G., Y. Huang, P. Humphreys, X. Wang, and A. Smith (Apr. 2011). "A <italic>PiggyBac</italic>-Based Recessive Screening Method to Identify Pluripotency Regulators". In: *PLoS ONE* 6.4, e18189 (cit. on pp. 93, 183).

Helikar, T., J. Konvalina, J. Heidel, and J. A. Rogers (Feb. 2008). "Emergent decision-making in biological signal transduction networks". In: *PNAS* 105.6, pp. 1913–1918. DOI: 10.1073/pnas.0705088105 (cit. on pp. 36, 37, 99, 109, 121, 124, 127, 133, 173).

St-Hilaire, M. and A. Longtin (2004). "Comparison of coding capabilities of Type I and Type II neurons". In: *J Comput Neurosci* 16, pp. 299–313 (cit. on p. 27).

Hodgkin, A. L. (1948). "The local electric changes associated with repetitive action in a non-medullated axon". In: *J. Physiol.* 107, pp. 165–181 (cit. on p. 27).

Hofer, T., J. A. Sherratt, and P. K. Maini (1995). "Dictyostelium discoideum: Cellular Self-Organization in an Excitable Biological Medium". In: *Proc. Royal Soc. London B* 259, pp. 249–257 (cit. on p. 26).

Hornung, G. and N. Barkai (Jan. 2008). "Noise Propagation and Signaling Sensitivity in Biological Networks: A Role for Positive Feedback". In: *PLoS Comput Biol* 4.1, e8 (cit. on p. 56).

Hsueh, R. C., M. Natarajan, I. Fraser, B. Pond, J. Liu, S. Mumby, H. Han, L. I. Jiang, M. I. Simon, R. Taussig, and P. C. Sternweis (Jan. 2009). "Deciphering signaling outcomes from a system of complex networks". In: *Science Signaling* 2.71, ra22. DOI: 10.1126/scisignal.2000054 (cit. on pp. 123, 135).

Huang, S. and D. E. Ingber (2000). "Shape-Dependent Control of Cell Growth, Differentiation, and Apoptosis: Switching between Attractors in Cell Regulatory Networks". In: *Exp. Cell Res.* 261.1, pp. 91–103. DOI: DOI : 10.1006/excr.2000.5044 (cit. on p. 36).

Izhikevich, E. M. (2006). *Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting*. The MIT Press (cit. on pp. 26, 27, 63).

Janes, K. A., J. G. Albeck, S. Gaudet, P. K. Sorger, D. A. Lauffenburger, and M. B. Yaffe (Dec. 2005). "A Systems Model of Signaling Identifies a Molecular Basis Set for Cytokine-Induced Apoptosis". In: *Science* 310.5754, pp. 1646–1653 (cit. on p. 135).

Janes, K. A., H. C. Reinhardt, and M. B. Yaffe (2008). "Cytokine-induced signaling networks prioritize dynamic range over signal strength." In: *Cell* 135.2, pp. 343–354. DOI: 10.1016/j.cell.2008.08.034 (cit. on p. 124).

Just, W., I. Shmulevich, and J. Konvalina (2004). "The number and probability of canalizing functions". In: *Physica D* 197.3-4, pp. 211–221 (cit. on p. 35).

Kalmar, T., C. Lim, P. Hayward, S. Muñoz-Descalzo, J. Nichols, J. Garcia-Ojalvo, and A. Martinez Arias (July 2009). "Regulated fluctuations in nanog expression mediate cell fate decisions in embryonic stem cells". In: *PLoS Biol* 7.7, e1000149. DOI: 10.1371/journal.pbio.1000149 (cit. on pp. 26, 27, 30–32, 71, 72, 74, 77, 95, 154).

Kampen, N. G. van (2007). *Stochastic processes in physics and chemistry*. North Holland (cit. on pp. 60, 61, 160).

Karagiannis, G. S., M. P. Pavlou, and E. P. Diamandis (Dec. 2010). "Cancer secretomics reveal pathophysiological pathways in cancer molecular oncology". In: *Mol Oncol* 4.6, pp. 496–510. DOI: 10.1016/j.molonc.2010.09.001 (cit. on p. 123).

Kauffman, S. A. (1969). "Metabolic stability and epigenesis in randomly constructed genetic nets". In: *J. Theor. Biol.* 22, pp. 437–467 (cit. on pp. 35, 100, 173).

Kauffman, S. A. (1993). *The origins of order: self-organization and selection in evolution*. Oxford University Press (cit. on p. 35).

Kauffman, S. A., C. Peterson, B. Samuelsson, and C. Troein (2004). "Genetic networks with canalyzing Boolean rules are always stable". In: *Proceedings of the National Academy of Sciences of the United States of America* 101.49, pp. 17102–17107. DOI: 10.1073/pnas.0407783101 (cit. on p. 109).

Kearns, D. B., F. Chu, S. S. Branda, R. Kolter, and R. Losick (Feb. 2005). "A master regulator for biofilm formation by Bacillus subtilis". In: *Molecular Microbiology* 55.3, pp. 739–749. DOI: 10.1111/j.1365-2958.2004.04440.x (cit. on p. 38).

Keller, M., J. Mazuch, U. Abraham, G. Eom, E. Herzog, H. Volk, A. Kramer, and B. Maier (Dec. 2009). "A circadian clock in macrophages controls inflammatory immune responses". In: *Proc Natl Acad Sci USA*. DOI: 10.1073/pnas.0906361106 (cit. on p. 123).

Kelly, K. F., D. Y. Ng, G. Jayakumaran, G. A. Wood, H. Koide, and B. W. Doble (Feb. 2011). " -Catenin Enhances Oct-4 Activity and Reinforces Pluripotency through a TCF-Independent Mechanism". In: *Cell stem cell* 8.2, pp. 214–227 (cit. on pp. 33, 86).

Kholodenko, B. N., J. F. Hancock, and W. Kolch (June 2010). "Signalling ballet in space and time". In: *Nat Rev Mol Cell Biol* 11.6, pp. 414–426 (cit. on p. 35).

Kim, J., J. Chu, X. Shen, J. Wang, and S. H. Orkin (Mar. 2008). "An Extended Transcriptional Network for Pluripotency of Embryonic Stem Cells". In: *Cell* 132.6, pp. 1049–1061 (cit. on p. 76).

Kittisopikul, M. and G. M. Süel (June 2010). "Biological role of noise encoded in a genetic network motif". In: *Proc Natl Acad Sci U S A* 107, pp. 1–6 (cit. on pp. 56, 70, 154).

Klemm, K. and S. Bornholdt (2005a). "Stable and unstable attractors in {Boolean} networks". In: *Phys. Rev. E* 72, pp. 101–104 (cit. on pp. 35, 113).

– (2005b). "Topology of biological networks and reliability of information processing". In: *Proc. Natl. Acad. Sci. USA* 102, pp. 414–419 (cit. on p. 35).

Klevecz, R. R. (Nov. 1976). "Quantized generation time in mammalian cells as an expression of the cellular clock". In: *Proc Natl Acad Sci USA* 73.11, pp. 4012–4016 (cit. on pp. 28, 29).

Koseska, A., E. Volkov, A. Zaikin, and J. Kurths (Aug. 2007). "Quantized cycling time in artificial gene networks induced by noise and intercell communication". In: *Physical Review E* 76.2, p. 20901. DOI: 10.1103/PhysRevE.76.020901 (cit. on p. 70).

Lahav, G., N. Rosenfeld, A. Sigal, N. Geva-Zatorsky, A. J. Levine, M. B. Elowitz, and U. Alon (Feb. 2004). "Dynamics of the p53-Mdm2 feedback loop in individual cells". In: *Nat Genet* 36.2, pp. 147–150. DOI: 10.1038/ng1293 (cit. on p. 26).

Lanner, F., K. L. Lee, M. Sohl, K. Holmborn, H. Yang, J. Wilbertz, L. Poellinger, J. Rossant, and F. Farnebo (2010). "Heparan Sulfation–Dependent Fibroblast Growth Factor Signaling Maintains Embryonic Stem Cells Primed for Differentiation in a Heterogeneous State". In: *STEM CELLS* 28.2, pp. 191–200. DOI: 10.1002/stem.265 (cit. on p. 87).

Lindner, B., J. Garcia-Ojalvo, A. Neiman, and L. Schimansky-Geier (2004). "Effects of noise in excitable systems". In: *Physics Reports* 392.6, pp. 321–424 (cit. on pp. 26, 29, 60, 66–68, 70).

Loewer, A., E. Batchelor, G. Gaglia, and G. Lahav (2010). "Basal Dynamics of p53 Reveal Transcriptionally Attenuated Pulses in Cycling Cells". In: *Cell* 142.1, pp. 89–100 (cit. on pp. 26, 154).

Loh, K. M. and B. Lim (Apr. 2011). "A Precarious Balance: Pluripotency Factors as Lineage Specifiers". In: *Cell Stem Cell* 8.4, pp. 363–369. DOI: http://dx.doi.org/10.1016/j.stem.2011.03.013 (cit. on p. 179).

Loh, Y.-H., Q. Wu, J.-L. Chew, V. B. Vega, W. Zhang, X. Chen, G. Bourque, J. George, B. Leong, J. Liu, K.-Y. Wong, K. W. Sung, C. W. H. Lee, X.-D. Zhao, K.-P. Chiu, L. Lipovich, V. A. Kuznetsov, P. Robson, L. W. Stanton, C.-L. Wei, Y. Ruan, B. Lim, and H.-H. Ng (Apr. 2006). "The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells". In: *Nat Genet* 38.4, pp. 431–440 (cit. on pp. 76, 180).

Lu, R., F. Markowetz, R. D. Unwin, J. T. Leek, E. M. Airoldi, B. D. MacArthur, A. Lachmann, R. Rozov, A. Ma/'ayan, L. A. Boyer, O. G. Troyanskaya, A. D. Whetton, and I. R. Lemischka (Nov. 2009). "Systems-level dynamic analyses of fate change in murine embryonic stem cells". In: *Nature* 462.7271, pp. 358–362 (cit. on p. 76).

Luo, Y., C. L. Lim, J. Nichols, A. Martinez Arias, and L. Wernisch (Jan. 2013). "Cell signalling regulates dynamics of Nanog distribution in embryonic stem cell populations". In: *Journal of The Royal Society Interface* 10.78 (cit. on pp. 72, 95, 154).

Lyashenko, N., M. Winter, D. Migliorini, T. Biechele, R. T. Moon, and C. Hartmann (July 2011). "Differential requirement for the dual functions of [beta]-catenin in embryonic stem cell self-renewal and germ layer formation". In: *Nat Cell Biol* 13.7, pp. 753–761 (cit. on p. 85).

Ma, W., L. Lai, Q. Ouyang, and C. Tang (Dec. 2006). "Robustness and modular design of the Drosophila segment polarity network". In: *Mol Syst Biol* 2 (cit. on p. 56).

Ma, W., A. Trusina, H. El-Samad, W. A. Lim, and C. Tang (Jan. 2009). "Defining Network Topologies that Can Achieve Biochemical Adaptation". In: *Cell* 138.4, pp. 760–773. DOI: 10.1016/j.cell.2009.06.013 (cit. on p. 56).

Mammoto, T. and D. E. Ingber (May 2010). "Mechanical control of tissue and organ development". In: *Development* 137.9, pp. 1407–1420 (cit. on p. 38).

Martello, G., T. Sugimoto, E. Diamanti, A. Joshi, R. Hannah, S. Ohtsuka, B. Göttgens, H. Niwa, and A. Smith (Oct. 2012). *Esrrb Is a Pivotal Target of the Gsk3/Tcf3 Axis Regulating Embryonic Stem Cell Self-Renewal* (cit. on pp. 77, 86, 87).

Martinez Arias, A. and J. M. Brickman (Dec. 2011). "Gene expression heterogeneities in embryonic stem cell populations: origin and function." In: *Curr Opin Cell Biol* 23.6, pp. 650–656. DOI: 10.1016/j.ceb.2011.09.007 (cit. on pp. 26, 30).

Martins, E. Q. V. and M. M. B. Pascoal (2003). "A new implementation of {Y}en's ranking loopless paths algorithm". In: *4OR: A Quarterly Journal of Operations Research* 1.2, pp. 121–133 (cit. on p. 117).

Marvasi, M., P. T. Visscher, and L. Casillas Martinez (Dec. 2010). "Exopolymeric substances (EPS) from Bacillus subtilis : polymers and genes encoding their synthesis". In: *FEMS Microbiology Letters* 313.1, pp. 1–9. DOI: 10.1111/j.1574-6968.2010.02085.x (cit. on p. 38).

Masui, S., Y. Nakatake, Y. Toyooka, D. Shimosato, R. Yagi, K. Takahashi, H. Okochi, A. Okuda, R. Matoba, A. A. Sharov, M. S. H. Ko, and H. Niwa (June 2007). "Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells". In: *Nat Cell Biol* 9.6, pp. 625–635 (cit. on p. 76).

Masui, Y. and P. Wang (Nov. 1998). "Cell cycle transition in early embryonic development of Xenopus laevis". In: *Biol Cell* 90.8, pp. 537–548 (cit. on pp. 28, 29).

Matoba, R., H. Niwa, S. Masui, S. Ohtsuka, M. G. Carter, A. A. Sharov, and M. S. H. Ko (Dec. 2006). "Dissecting Oct3/4-Regulated Gene Networks in Embryonic Stem Cells by Expression Profiling". In: *PLoS ONE* 1.1, e26 (cit. on p. 76).

Matsumoto, M. and T. Nishimura (1998). "Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator". In: *ACM Trans. Model. Comput. Simul.* 8.1, pp. 3–30. DOI: http://doi.acm.org/10.1145/272991.272995 (cit. on p. 173).

Mayo, A. E., Y. Setty, S. Shavit, A. Zaslaver, and U. Alon (Mar. 2006). "Plasticity of the <italic>cis</italic>-Regulatory Input Function of a Gene". In: *PLoS Biol* 4.4, e45 (cit. on p. 132).

McKane, A. J. and T. J. Newman (June 2005). "Predator-prey cycles from resonant amplification of demographic stochasticity". In: *Phys. Rev. Lett.* 94.21, p. 218102 (cit. on p. 70).

Mendoza, L., D. Thieffry, and E. R. Alvarez-Buylla (Jan. 1999). "Genetic control of flower morphogenesis in Arabidopsis thaliana: a logical analysis". In: *Bioinformatics* 15.7-8, pp. 593–606 (cit. on p. 36).

Meng, X., G. Huguet, and J. Rinzel (2012). "Type III excitability, slope sensitivity and coincidence detection". In: *Discrete Contin Dyn Syst Ser A* (cit. on p. 154).

Miyanari, Y. and M.-E. Torres-Padilla (Mar. 2012). "Control of ground-state pluripotency by allelic regulation of Nanog". In: *Nature* 483.7390, pp. 470–473 (cit. on pp. 80, 82, 87).

Mullin, N. P., A. Yates, A. J. Rowe, B. Nijmeijer, D. Colby, P. N. Barlow, M. D. Walkinshaw, and I. Chambers (Apr. 2008). "The pluripotency rheostat Nanog functions as a dimer". In: *Biochemical Journal* 411.2, pp. 227–231 (cit. on p. 84).

Muñoz Descalzo, S., P. Rué, F. Faunes, P. Hayward, L. M. Jakt, T. Balayo, J. Garcia-Ojalvo, and A. Martinez Arias (2013). "A competitive protein interaction network buffers Oct4-mediated differentiation to promote pluripotency in embryonic stem cells". In: *Mol Syst Biol (under review)* (cit. on pp. 81, 90, 91, 95, 170, 183, 185).

Muñoz Descalzo, S., P. Rué, J. Garcia-Ojalvo, and A. Martinez Arias (Dec. 2012). "Correlations Between the Levels of Oct4 and Nanog as a Signature for Naïve Pluripotency in Mouse Embryonic Stem Cells". In: *Stem Cells* 30.12, pp. 2683–2691. DOI: 10.1002/stem.1230 (cit. on pp. 32, 79, 92, 93, 96, 185).

Munteanu, A., M. Constante, M. Isalan, and R. Sole (2010). "Avoiding transcription factor competition at promoter level increases the chances of obtaining oscillation". In: *BMC Systems Biology* 4.1, p. 66 (cit. on p. 56).

Natarajan, M., K.-M. Lin, R. C. Hsueh, P. C. Sternweis, and R. Ranganathan (June 2006). "A global analysis of cross-talk in a mammalian cellular signalling network". In: *Nature* 8.6, pp. 571–580. DOI: 10.1038/ncb1418 (cit. on pp. 123, 124, 128).

Navarro, P., N. Festuccia, D. Colby, A. Gagliardi, N. P. Mullin, W. Zhang, V. Karwacki-Neisius, R. Osorno, D. Kelly, M. Robertson, and I. Chambers (Dec. 2012). "OCT4/SOX2-independent Nanog autorepression modulates heterogeneous Nanog gene expression in mouse ES cells". In: *EMBO J* 31.24, pp. 4547–4562 (cit. on pp. 76, 77, 80, 82).

Nishiyama, A., A. A. Sharov, Y. Piao, M. Amano, T. Amano, H. G. Hoang, B. Y. Binder, R. Tapnio, U. Bassey, J. N. Malinou, L. S. Correa-Cerro, H. Yu, L. Xin, E. Meyers, M. Zalzman, Y. Nakatake, C. Stagg, L. Sharova, Y. Qian, D. Dudekula, S. Sheer, J. S. Cadet, T. Hirata, H.-T. Yang, I. Goldberg, M. K. Evans, D. L. Longo, D. Schlessinger, and M. S. H. Ko (Mar. 2013). "Systematic repression of transcription factors reveals limited patterns of gene expression changes in ES cells". In: *Sci. Rep.* 3 (cit. on pp. 76, 90).

Nishiyama, A., L. Xin, A. A. Sharov, M. Thomas, G. Mowrer, E. Meyers, Y. Piao, S. Mehta, S. Yee, Y. Nakatake, C. Stagg, L. Sharova, L. S. Correa-Cerro, U. Bassey, H. Hoang, E. Kim, R. Tapnio, Y. Qian, D. Dudekula, M. Zalzman, M. Li, G. Falco, H.-T. Yang, S.-L. Lee, M. Monti, I. Stanghellini, M. N. Islam, R. Nagaraja, I. Goldberg, W. Wang, D. L. Longo, D. Schlessinger, and M. S. H. Ko (Oct. 2009). "Uncovering Early Response of Gene Regulatory Networks in ESCs by Systematic Induction of Transcription Factors". In: *Cell Stem Cell* 5.4, pp. 420–433. DOI: http://dx.doi.org/10.1016/j.stem.2009.07.012 (cit. on pp. 76, 90, 94, 96, 171, 179, 180).

Niwa, H., S. Masui, I. Chambers, A. G. Smith, and J.-i. Miyazaki (Mar. 2002). "Phenotypic Complementation Establishes Requirements for Specific POU Domain and Generic Transactivation Function of Oct-3/4 in Embryonic Stem Cells". In: *Molecular and Cellular Biology* 22.5, pp. 1526–1536 (cit. on pp. 93, 96, 170).

Niwa, H., J.-i. Miyazaki, and A. G. Smith (Apr. 2000). "Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells". In: *Nat Genet* 24.4, pp. 372–376 (cit. on pp. 90, 93, 94, 96).

Novak, B., Z. Pataki, A. Ciliberto, and J. J. Tyson (Mar. 2001). "Mathematical model of the cell division cycle of fission yeast". In: *Chaos* 11.1, pp. 277–286. DOI: 10.1063/1.1345725 (cit. on p. 29).

Novák, B. and J. J. Tyson (Dec. 2008). "Design principles of biochemical oscillators". In: *Nat. Rev. Mol. Cell Biol.* 9.12, pp. 981–991. DOI: 10.1038/nrm2530 (cit. on pp. 28, 37).

Oda, K. and H. Kitano (2006). "A comprehensive map of the toll-like receptor signaling network". In: *Molecular Systems Biology* 2, p. 2006.0015 (cit. on p. 125).

Oda, K., Y. Matsuoka, A. Funahashi, and H. Kitano (Jan. 2005). "A comprehensive pathway map of epidermal growth factor receptor signaling". In: *Molecular Systems Biology* 1, p. 2005.0010. DOI: 10.1038/msb4100014 (cit. on pp. 35, 155).

Padrick, S. B. and M. K. Rosen (June 2010). "Physical Mechanisms of Signal Integration by WASP Family Proteins". In: *Annual Review of Biochemistry* 79.1, pp. 707–735. DOI: 10.1146/annurev.biochem.77.060407.135452 (cit. on p. 134).

Pereira, L., F. Yi, and B. J. Merrill (Oct. 2006). "Repression of Nanog Gene Transcription by Tcf3 Limits Embryonic Stem Cell Self-Renewal". In: *Molecular and Cellular Biology* 26.20, pp. 7479–7491 (cit. on pp. 32, 72, 77, 93).

Pikovsky, A., A. Zaikin, and M. A. de la Casa (2002). "System Size Resonance in Coupled Noisy Systems and in the Ising Model". In: *Phys. Rev. Lett.* 88, p. 50601 (cit. on p. 60).

Pikovsky, A. S. and J. Kurths (Feb. 1997). "Coherence Resonance in a Noise-Driven Excitable System". In: *Phys. Rev. Lett.* 78.5, pp. 775–778. DOI: 10.1103/PhysRevLett.78.775 (cit. on pp. 29, 60, 66, 67).

Prehoda, K. E., J. A. Scott, R. D. Mullins, and W. A. Lim (Oct. 2000). "Integration of Multiple Signals Through Cooperative Regulation of the N-WASP-Arp2/3 Complex". In: *Science* 290.5492, pp. 801–806 (cit. on p. 134).

Quine, W. V. (1952). "The Problem of Simplifying Truth Functions". In: *The American Mathematical Monthly* 59.8, pp. 521–531 (cit. on p. 173).

Radzisheuskaya, A., G. Le Bin Chia, R. L. dos Santos, T. W. Theunissen, L. F. C. Castro, J. Nichols, and J. C. R. Silva (Apr. 2013). "A defined Oct4 level governs cell state transitions of pluripotency entry and differentiation into all embryonic lineages". In: *Nat Cell Biol* advance on (cit. on p. 96).

Ribeiro, A. S. and S. A. Kauffman (Aug. 2007). "Noisy attractors and ergodic sets in models of gene regulatory networks". In: *J. theor. Biol.* 247.4, pp. 743–755. DOI: 10.1016/j.jtbi.2007.04.020 (cit. on p. 35).

Rodda, D. J., J.-L. Chew, L.-H. Lim, Y.-H. Loh, B. Wang, H.-H. Ng, and P. Robson (July 2005). "Transcriptional Regulation of Nanog by OCT4 and SOX2". In: *Journal of Biological Chemistry* 280.26, pp. 24731–24737 (cit. on pp. 76, 77).

Rose, J. E., J. F. Brugge, D. J. Anderson, and J. E. Hind (July 1967). "Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey". In: *J Neurophysiol* 30.4, pp. 769–793 (cit. on p. 29).

Rué, P., N. Domedel-Puig, J. Garcia-Ojalvo, and A. J. Pons (2012). "Integration of cellular signals in chattering environments". In: *Progress in Biophysics and Molecular Biology*. DOI: 10.1016/j.pbiomolbio.2012.05.003 (cit. on p. 185).

Rué, P., A. J. Pons, N. Domedel-Puig, and J. Garcia-Ojalvo (Jan. 2010). "Relaxation dynamics and frequency response of a noisy cell signaling network". In: *Chaos* 20.4, p. 45110. DOI: 10.1063/1.3524908 (cit. on pp. 110, 111, 185).

Rué, P. and J. Garcia-Ojalvo (May 2011). "Gene circuit designs for noisy excitable dynamics". In: *Mathematical Biosciences* 231.1, pp. 90–97. DOI: 10.1016/j.mbs.2011.02.013 (cit. on p. 185).

– (May 2013). "Modeling Gene Expression in Time and Space". In: *Annual Review of Biophysics* 42.1, pp. 605–627. DOI: 10.1146/annurev-biophys-083012-130335 (cit. on p. 185).

Rué, P., G. M. Süel, and J. Garcia-Ojalvo (June 2011). "Optimizing periodicity and polymodality in noise-induced genetic oscillators". In: *Physical Review E* 83.6, p. 61904. DOI: 10.1103/PhysRevE.83.061904 (cit. on p. 185).

Rutkowski, J. M. and M. A. Swartz (Jan. 2007). "A driving force for change: interstitial flow as a morphoregulator". In: *Trends Cell Biol* 17.1, pp. 44–50. DOI: 10.1016/j.tcb.2006.11.007 (cit. on p. 123).

Saez-Rodriguez, J., L. Simeoni, J. A. Lindquist, R. Hemenway, U. Bommhardt, B. Arndt, U.-U. Haus, R. Weismantel, E. D. Gilles, S. Klamt, and B. Schraven (Aug. 2007). "A logical model provides insights into T cell receptor signaling". In: *PLoS Comput Biol* 3.8, e163. DOI: 10.1371/journal.pcbi.0030163 (cit. on p. 36).

Sagués, F., J. M. Sancho, and J. Garcia-Ojalvo (2007). "Spatiotemporal order out of noise". In: *Reviews of modern physics* 79.3, pp. 829–882 (cit. on p. 70).

Samaga, R., J. Saez-Rodriguez, L. G. Alexopoulos, P. K. Sorger, and S. Klamt (2009). "The logic of {EGFR/ErbB} signaling: theoretical properties and analysis of high-throughput data". In: *PLoS Comp. Biol.* 5.8, e1000438. DOI: 10.1371/journal.pcbi.1000438 (cit. on p. 36).

Samuelsson, B. and C. Troein (2003). "Superpolynomial Growth in the Number of Attractors in {Kauffman} Networks". In: *Phys. Rev. Lett.* 90.9, p. 98701. DOI: 10.1103/PhysRevLett.90.098701 (cit. on p. 35).

Sancristóbal, B., J. M. Sancho, and J. Garcia-Ojalvo (2010). "Phase-response approach to firing-rate selectivity in neurons with subthreshold oscillations". In: *Phys. Rev. E* 82, p. 41908 (cit. on p. 48).

Saxe, J. P., A. Tomilin, H. R. Schöler, K. Plath, and J. Huang (Feb. 2009). "Post-Translational Regulation of Oct4 Transcriptional Activity". In: *PLoS ONE* 4.2, e4467 (cit. on p. 84).

Schimz, A. and E. Hildebrand (Jan. 1992). "Nonrandom structures in the locomotor behavior of Halobacterium: a bifurcation route to chaos?" In: *Proc Natl Acad Sci USA* 89.2, pp. 457–460 (cit. on p. 29).

Schmid, G., I. Goychuk, and P. Hänggi (2001). "Stochastic resonance as a collective property of ion channel assemblies". In: *Europhys. Lett.* 56, p. 22 (cit. on p. 60).

Sharov, A., S. Masui, L. Sharova, Y. Piao, K. Aiba, R. Matoba, L. Xin, H. Niwa, and M. Ko (2008). "Identification of Pou5f1, Sox2, and Nanog downstream target genes with statistical confidence by applying a novel algorithm to time course microarray and genome-wide chromatin immunoprecipitation data". In: *BMC Genomics* 9.1, p. 269 (cit. on pp. 76, 179, 180).

Sharov, A., A. Nishiyama, Y. Piao, L. Correa-Cerro, T. Amano, M. Thomas, S. Mehta, and M. Ko (2011). "Responsiveness of genes to manipulation of transcription factors in ES cells is associated with histone modifications and tissue specificity". In: *BMC Genomics* 12.1, p. 102 (cit. on p. 76).

Silva, J. and A. Smith (Feb. 2008). "Capturing Pluripotency". In: *Cell* 132.4, pp. 532–536 (cit. on p. 91).

Singh, A. M., T. Hamazaki, K. E. Hankowski, and N. Terada (Oct. 2007). "A heterogeneous expression pattern for Nanog in embryonic stem cells." In: *Stem cells* 25.10, pp. 2534–2542. DOI: 10.1634/stemcells.2007-0126 (cit. on pp. 26, 30).

Smith, A. G. (Nov. 2001). "EMBRYO-DERIVED STEM CELLS: Of Mice and Men". In: *Annual Review of Cell and Developmental Biology* 17.1, pp. 435–462. DOI: 10.1146/annurev.cellbio.17.1.435 (cit. on p. 30).

Stark, J., C. Chan, and A. J. T. George (Apr. 2007). "Oscillations in the immune system". In: *Immunological Reviews* 216, pp. 213–231 (cit. on p. 123).

Steuer, R. (June 2004). "Effects of stochasticity in models of the cell cycle: from quantized cycle times to noise-induced oscillations". In: *J Theor Biol* 228.3, pp. 293–301. DOI: 10.1016/j.jtbi.2004.01.012 (cit. on p. 29).

Stricker, J., S. Cookson, M. R. Bennett, W. H. Mather, L. S. Tsimring, and J. Hasty (Nov. 2008). "A fast, robust and tunable synthetic gene oscillator". In: *Nature* 456.7221, pp. 516–519. DOI: 10.1038/nature07389 (cit. on pp. 45, 61).

Strogatz, S. H. (Jan. 1994). *Nonlinear Dynamics And Chaos: With Applications To Physics, Biology, Chemistry, And Engineering (Studies in Nonlinearity)*. 1st ed. Studies in nonlinearity. Perseus Books Group (cit. on p. 27).

Süel, G. M., J. Garcia-Ojalvo, L. M. Liberman, and M. B. Elowitz (2006). "An excitable gene regulatory circuit induces transient cellular differentiation". In: *Nature* 440.7083, pp. 545–550 (cit. on pp. 25, 26, 44, 61, 154, 159).

Suel, G. M., R. P. Kulkarni, J. Dworkin, J. García-Ojalvo, M. B. Elowitz, G. M. Süel, and J. Garcia-Ojalvo (2007). "Tunability and noise dependence in differentiation dynamics". In: *Science* 315.5819, p. 1716 (cit. on pp. 25, 26, 46, 60–62).

Sveiczer, A., A. Csikasz-Nagy, B. Gyorffy, J. J. Tyson, and B. Novak (July 2000). "Modeling the fission yeast cell cycle: quantized cycle times in wee1- cdc25Delta mutant cells". In: *Proc Natl Acad Sci USA* 97.14, pp. 7865–7870 (cit. on p. 29).

Sveiczer, A., B. Novak, and J. M. Mitchison (1996). "The size control of fission yeast revisited". In: *J Cell Sci* 109, pp. 2947–2957 (cit. on pp. 28, 29).

Taniguchi, Y., P. J. Choi, G.-W. Li, H. Chen, M. Babu, J. Hearn, A. Emili, and X. S. Xie (2010). "Quantifying E. coli proteome and transcriptome with single-molecule sensitivity in single cells". In: *Science* 329, pp. 533–538 (cit. on p. 68).

Tigges, M., T. T. Marquez-Lago, J. Stelling, and M. Fussenegger (2009). "A tunable synthetic mammalian oscillator". In: *Nature* 457, pp. 309–312 (cit. on p. 61).

Toral, R., C. R. Mirasso, and J. D. Gunton (2003). "System size coherence resonance in coupled FitzHugh-Nagumo models". In: *Europhys. Lett.* 61, p. 162 (cit. on p. 60).

Tsai, T. Y.-C., Y. S. Choi, W. Ma, J. R. Pomerening, C. Tang, and J. E. Ferrell (2008). "Robust, Tunable Biological Oscillations from Interlinked Positive and Negative Feedback Loops". In: *Science* 321, pp. 126–129 (cit. on p. 61).

Turcotte, M., J. Garcia-Ojalvo, and G. M. Süel (Oct. 2008). "A genetic timer through noise-induced stabilization of an unstable state". In: *Proceedings of the National Academy of Sciences* 105.41, pp. 15732–15737. DOI: 10.1073/pnas.0806349105 (cit. on pp. 29, 59, 63–65).

Tyson, J. J., K. Chen, and B. Novak (Dec. 2001). "Network dynamics and cell physiology". In: *Nat Rev Mol Cell Biol* 2.12, pp. 908–916. DOI: 10.1038/35103078 (cit. on p. 28).

Volkov, E. I., M. N. Stolyarov, A. A. Zaikin, and J. Kurths (2003). "Coherence resonance and polymodality in inhibitory coupled excitable oscillators". In: *Physical Review E* 67.6, p. 66202 (cit. on p. 70).

Wang, J., S. Rao, J. Chu, X. Shen, D. N. Levasseur, T. W. Theunissen, and S. H. Orkin (Nov. 2006). "A protein interaction network for pluripotency of embryonic stem cells". In: *Nature* 444.7117, pp. 364–368 (cit. on pp. 34, 80, 86).

Webb, J. S., M. Givskov, and S. Kjelleberg (Dec. 2003). "Bacterial biofilms: prokaryotic adventures in multicellularity". In: *Current Opinion in Microbiology* 6.6, pp. 578–585. DOI: http://dx.doi.org/10.1016/j.mib.2003.10.014 (cit. on p. 38).

Wieczorek, S., P. Ashwin, C. M. Luke, and P. M. Cox (May 2011). "Excitability in ramped systems: the compost-bomb instability". In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science* 467.2129, pp. 1243–1269. DOI: 10.1098/rspa.2010.0485 (cit. on p. 154).

Wiesenfeld, K. and F. Moss (Jan. 1995). "Stochastic resonance and the benefits of noise: from ice ages to crayfish and SQUIDs". In: *Nature* 373.6509, pp. 33–36. DOI: 10.1038/373033a0 (cit. on p. 124).

Wilking, J. N., V. Zaburdaev, M. De Volder, R. Losick, M. P. Brenner, and D. A. Weitz (Jan. 2013). "Liquid transport facilitated by channels in Bacillus subtilis biofilms". In: *Proceedings of the National Academy of Sciences* 110.3, pp. 848–852 (cit. on p. 38).

Wray, J., T. Kalkan, S. Gomez-Lopez, D. Eckardt, A. Cook, R. Kemler, and A. Smith (July 2011). "Inhibition of glycogen synthase kinase-3 alleviates Tcf3 repression of the pluripotency network and increases embryonic stem cell resistance to differentiation". In: *Nat Cell Biol* 13.7, pp. 838–845 (cit. on pp. 85–87).

Yi, F., L. Pereira, and B. J. Merrill (2008). "Tcf3 Functions as a Steady-State Limiter of Transcriptional Programs of Mouse Embryonic Stem Cell Self-Renewal". In: *STEM CELLS* 26.8, pp. 1951–1960. DOI: 10.1634/stemcells.2008-0229 (cit. on pp. 76, 93).

Young, R. A. (Mar. 2011). "Control of the Embryonic Stem Cell State". In: *Cell* 144.6, pp. 940–954 (cit. on p. 30).

Yunker, P. J., T. Still, M. A. Lohr, and A. G. Yodh (Aug. 2011). "Suppression of the coffee-ring effect by shape-dependent capillary interactions". In: *Nature* 476.7360, pp. 308–311 (cit. on pp. 40, 144).

Zhang, L., S. Rayner, N. Katoku-Kikyo, L. Romanova, and N. Kikyo (Sept. 2007). "Successful co-immunoprecipitation of Oct4 and Nanog using cross-linking". In: *Biochemical and Biophysical Research Communications* 361.3, pp. 611–614. DOI: http://dx.doi.org/10.1016/j.bbrc.2007.07.089 (cit. on pp. 80, 86).

Zhang, X., J. Zhang, T. Wang, M. A. Esteban, and D. Pei (Dec. 2008). "Esrrb Activates Oct4 Transcription and Sustains Self-renewal and Pluripotency in Embryonic Stem Cells". In: *Journal of Biological Chemistry* 283.51, pp. 35825–35833 (cit. on p. 76).

Glossary

2i	2 inhibitors. 26, 27, 193
2i+L	2i and LIF. 26, 27, 73, 76, 78, 81, 83, 84, 86, 162, 164
AI	altered input. 104, 105, 107
AL	altered logic. 102, 104, 105, 107
bifurcation	qualitative change in the dynamics of a non-linear dynamic system arising from a small change in one or several control parameters. 38, 39, 41–43
BN	Boolean network. 29–31, 93, 94, 103, 107, 165, 167
BT	Bodganov–Takens. 49
canalising function	A logical function such that its output value is completely determined by at least one specific value of one of its input variables, independently of the values of the remaining inputs. 103, 104
CDP	cell death pattern. 33, 34, 138–143
CFDA	carboxyfluorescein diacetate. 134
chatter	Noisy background signal.. 31, 98, 101, 122–124, 167
CR	coherence resonance. 54, 56, 60–62
CV	Coefficient of Variation. 60–63
ECM	extracellular matrix. 31–34, 95, 141
EGF	epidermal growth factor. 30
ES	Embryonic Stem. 24, 27, 28, 65, 89, 90, 162
FACS	fluorescence-activated cell sorting. 24
FLC	fold limit cycle. 46, 48
genetic competence	ability of bacterial cells to take up exogenous DNA molecules from the environment. 19
GO	Gene Ontology. 175–177
GPCR	G protein-coupled receptor. 28, 31
HB	Hopf bifurcation. 42, 43, 46, 47, 57

HN	High Nanog. 89
homoclinic orbit	a trajectory of a dynamical system which joins a saddle equilibrium point to itself. 41
ICM	inner cell mass. 24
IL1/TNF-α	interleukin-1 and tumour necrosis factor. 30, 95, 123–125
iPS	induced pluripotent stem. 24
k-means clustering	method of data clustering which classifies the observations into k groups (clusters) in which each observation belongs to the cluster with the nearest average. 139, 140
LN	Low Nanog. 89
MAD	mean absolute difference. 127
mES	mouse Embryonic Stem. 24, 26
MN	Middle Nanog. 89
mRNA	messenger RNA. 37, 62, 76, 77, 153–158, 160
PD03	PD 0325901 (MEKi), an MKK1/MEK1 and MKK2/MEK2 inhibitor. 164
PSD	power spectral density. 100, 106, 109
QIF	quantitative immunofluorescence. 26, 28, 82
RBN	random Boolean network. 29, 94, 95
RTK	receptor tyrosine kinase. 28
S+L	Serum and LIF. 24, 27, 73, 76–78, 80, 81, 83, 84, 162, 164, 165
SGN	signalling. 176, 177
SH	saddle-homoclinic. 41, 53, 58
SN	Saddle Node. 41, 43, 49, 70
SNIC	saddle node on invariant circle. 43, 49, 57
SR	stochastic resonance. 54, 60
TF	transcription factor. 176, 177
WB	Western Blot. 82