TAMPEREEN TEKNILLINEN YLIOPISTO
TAMPERE UNIVERSITY OF TECHNOLOGY

Ilya Potapov

Switching and Oscillatory Dynamics of Genetic Networks

Julkaisu 1104 • Publication 1104

Tampere 2013

Tampereen teknillinen yliopisto. Julkaisu 1104
Tampere University of Technology. Publication 1104

Ilya Potapov

Switching and Oscillatory Dynamics of Genetic Networks

Thesis for the degree of Doctor of Philosophy to be presented with due permission for public examination and criticism in Tietotalo Building, Auditorium TB111, at Tampere University of Technology, on the 18th of January 2013, at 12 noon.

Tampereen teknillinen yliopisto - Tampere University of Technology
Tampere 2013

ISBN 978-952-15-2995-5 (printed)
ISBN 978-952-15-2998-6 (PDF)
ISSN 1459-2045

Abstract

Bi-stable and periodic behaviors are present in living beings. Mechanisms capable of bi-stability provide organisms with the means to choose between states, while cyclic systems can be used to track time, and thus regulate periodic processes in cells, allowing, for example, adaptability to the day-night cycle.

In general, the switching behavior reflects the ability of decision making, which is vital in changing environments. On the other hand, processes with periodic dynamics, such as the circadian clock and the cell cycle, are necessary to time the occurrence of specific actions and processes. Complex behaviors, such as switching between states and oscillatory dynamics, are made possible by networks of interactions between cellular components, at the molecular level.

Many essential functions in living cells are regulated by complex networks. One such network, perhaps the most vital one, is the genetic regulatory network, composed of genes and the products of their expression. This network is capable of responding to other chemical substances present in a cell as well. Associations between genes, correlating their expression, are able to generate complex dynamical patterns, not possible by individual genes.

Aside from characterizing the structure of natural genetic networks, another effort that is being made to better understand them consists of engineering artificial genetic motifs that are commonly found in the natural networks. In this way, it is possible to study their dynamics without the interference of many external signals that the native motifs are subject to. Simultaneously, *in silico* models matching the kinetics of the artificial motifs have been developed, so as to explore their dynamics in various conditions.

In silico models can be used to predict the behavior of the natural system. Recently, their complexity has increased significantly, so as to account realistically for novel findings on the role that a gene's sequence and events in transcription and translation elongation have on the dynamics of genetic circuits. Similar developments have been made so as to study the effect of cell-to-cell communication on the dynamics of cells living within communi-

ties.

In this thesis we study the dynamics of *in silico* models of genetic toggle switches and oscillators. Namely, we implement delayed stochastic models at the nucleotide and codon levels to study sequence dependent properties of these circuits. Also, we implement a deterministic model of a genetic oscillator in the context of cell populations to investigate how, from the coupling between cells, emerges novel collective dynamics, including the synchronization of processes. The results impact on the ongoing efforts to better understand how cells regulate and modulate their key physiological processes and provide clues on how to better engineer synthetic circuits to regulate processes in cells.

Preface

> "I have never debated with a knowledgeable person except that I won the debate, and I have never debated with an ignorant person except that I lost."
>
> Imam Al-Shafi'i

This work was carried out in the Department of Signal Processing at Tampere University of Technology, Finland, and in the Department of Theoretical Physics at Lebedev Physical Institute of Russian Academy of Sciences, Moscow, Russia. I would like to express the sincere gratitude to the Center for International Mobility (CIMO) for the financial support of the major part of this work. Two persons from CIMO, Tarja Mäkelä and Hannele Ahti, deserve special mentioning for their warm words and valuable advice.

I am sincerely thankful to Andre Sanches Ribeiro for honest, straightforward, and fruitful discussions which accompanied the work presented in this thesis. I am truly thankful to Evgenii Volkov who supervised a good half of this work and supported me all the way. He is also one making complicated things clear and understandable. I would like to express gratitude to Olli Yli-Harja for providing me an opportunity to join the Computational Systems Biology Research Group and for creating a diverse multi-disciplinary research environment.

My sincere thanks go to all members of the Laboratory of Biosystem Dynamics, who are an eternal source of ideas, and, especially, to Jarno Mäkelä and Jason Lloyd-Price for the beneficial discussions and great help in fulfillment of this work.

I express my grateful acknowledgment to Virve Larmila for her constant help and support, which range from setting me up in Tampere to useful advice, not to mention a plenty of other things that Virve makes for people without a hint of fatigue. Frank gratitude goes to Elina Orava for her help in submitting and processing this thesis for the final examination. Additionally, I wish to thank all personnel members of the Department of Signal Process-

ing, who make our lives easier and minds fully concentrated on scientific matters.

I am deeply grateful to the reviewers of this thesis, Dr. Krasimira Tsaneva-Atanasova and Prof. Marc Roussel, for their useful comments and suggestions to improve the content of the thesis. I keenly appreciate time and effort they have devoted to reviewing the work.

For those people who know me closer:

"O Lord that lends me life,
Lend me a heart replete with thankfulness!"

William Shakespeare, King Henry VI.

Ilya Potapov
November 2012

Abbreviations

Frequently used abbreviations are presented below.

DNA	Deoxyribonucleic Acid
RNA	Ribonucleic Acid
mRNA	Messenger RNA
tRNA	Transfer RNA
RBS	Ribosome Binding Site
ODE	Ordinary Differential Equation
QS	Quorum Sensing
AI	Autoinducer
SSA	Stochastic Simulation Algorithm
RNAp	RNA polymerase
GRN	Gene Regulatory Network
HB	Andronov-Hopf bifurcation
LP	Limit point bifurcation
BP	Pitchfork (symmetry breaking) bifurcation
TR	Invariant torus emergence bifurcation
HSS	Homogeneous steady state
IHSS	Inhomogeneous steady state
IHLC	Inhomogeneous limit cycle

Contents

List of Publications

This thesis is a compound of the following publications. In the text, these publications are referred to as "Publication-I", "Publication-II", etc. Enumeration of the publications is made chronologically, according to the time when they were published.

I. I. Potapov, E. Volkov and A. Kuznetsov, *Dynamics of coupled repressilators: the role of mRNA kinetics and transcription cooperativity*, Phys. Rev. E **83**, 031901, 2011.

II. I. Potapov, J. Lloyd-Price, O. Yli-Harja and A.S. Ribeiro, *Dynamics of a genetic switch at the nucleotide and codon levels*, 8^{th} International Workshop on Computational Systems Biology (WCSB), Switzerland, Zurich, 2011.

III. I. Potapov, J. Lloyd-Price, O. Yli-Harja and A.S. Ribeiro, *Dynamics of a genetic toggle switch at the nucleotide and codon levels*, Phys. Rev. E **84**, 031903, 2011.

IV. I. Potapov, B. Zhurov and E. Volkov, *"Quorum sensing" generated multistability and chaos in a synthetic genetic oscillator*, Chaos **22**, 023117, 2012.

V. I. Potapov, J. Mäkelä, O. Yli-Harja and A.S. Ribeiro, *Effects of codon sequence on the dynamics of genetic networks*, J. Theor. Biol. **315**, 17–25, 2012.

The author of the thesis contributed to these publications as follows. In Publication-I, Publication-II and Publication-III, the author conducted most of the simulations and of the analysis of the results. Also, he participated actively in the writing of the manuscripts. In Publication-IV, the simulations' design and the study were conceived by the author and by E. Volkov. The author performed most simulations and the analysis of the results, and participated in the writing of the manuscript. In Publication-V, the author

performed the simulations and the data analysis along with J. Mäkelä. The author also participated in devising the study and writing the manuscript.

Chapter 1

Introduction

1.1 Background and Motivation

Living organisms are utterly complex. Nevertheless, they exhibit highly ordered behaviors and can even organize themselves in complex structures in an extremely efficient fashion (Murray 2002). The organization within each organism is made possible, to some extent, by decision making mechanisms based on bi-stable genetic circuits (Jacob and Monod 1961) able to respond to various stimuli, some generated internally and some arising from the surrounding environment.

This decision-making is inherent in a broad range of processes: from adopting one of several possible pathways of differentiation (Chang et al. 2006, 2008) and choosing a cell's fate (Gupta 2002; Matsuzawa and Ichijo 2001), following some internal change, to switch-like responses to cell–cell signals (Dockery and Keener 2001; Goryachev 2011; Miller and Bassler 2001; Waters and Bassler 2005).

The temporal order of the internal processes of an organism needs to be regulated by some internal clock — a mechanism that relies on a periodic process. These periodic processes are present in widely varying biological and medical contexts, such as the heart beat, certain neuronal activities in the brain, breathing, and the 24-hour periodic emergence of fruit flies from their pupae (Keener and Sneyd 1998; Murray 2002; Rinzel 1981; Winfree 1987, 2000). At the cellular level, there are also well-known examples of the oscillatory behaviors, such as the one driven by the circadian clock (which follows the 24-hour rhythm of day and night) and the cell cycle, driven by the gene regulatory network (Chance et al. 1964; Dunlap 1999; Gekakis et al. 1998; Goldbeter 1996; Nurse 2000)

A gene regulatory network is a key concept of modern biology. It consists

of genes, whose products of expression are capable of regulating the expression of other genes. Additionally, gene networks can also respond to external signals. Recently, several artificial genetic networks have been engineered, to implement novel biological functions and to investigate existing ones in nature (Andrianantoandro et al. 2006; Benner and Sismour 2005; Heinemann and Panke 2006).

Switching and oscillatory functioning are, so far, the most important aims in synthetic biology efforts (Atkinson et al. 2003; Danino et al. 2010; Elowitz and Leibler 2000; Gardner et al. 2000; Mondragón-Palomino et al. 2011; Stricker et al. 2008; Tigges et al. 2009). The first behavior is necessary for decision making, while the latter is needed to track time.

Usually, these studies, while relying on experimental efforts, are also accompanied with the examination of models, which aid the engineering of these circuits by predicting which underlying properties are necessary (García-Ojalvo et al. 2004; McMillen et al. 2002; Ullner et al. 2007). The synthetic circuits serve as prototypes of naturally existing mechanisms. For example, the genetic switch engineered by Gardner and colleagues (Gardner et al. 2000) aims to mimic similar circuits that allow bacteria to opt between opposite phenotypes. On the other hand, the three-gene oscillator (Elowitz and Leibler 2000) was based on theoretical ring oscillators, and aims to serve as a biological clock. Following these simpler circuits, such as switches and oscillators, the next future step is to assemble larger scale systems from simple single-cell units (Purnick and Weiss 2009).

While single live cells are already highly complex, there are much higher orders of complexity in nature, due to their ability to communicate and form multi-cellular populations, whose behavior cannot be solely explained by the behavior of each individual. This is evident for higher-order organisms. However, there is an increasing number of reports on various communication mechanisms in prokaryotes that indicate that they behave in a collective way, rather than individually as previously thought (Goryachev 2011; Miller and Bassler 2001; Waters and Bassler 2005). These findings open possibilities for synthetic biologists to advance to population level studies and also to aim for understanding and engineering more complex biological devices (Levskaya et al. 2005; Liu et al. 2011).

Populations of coupled genetic switches and oscillators have been widely studied for many years in the context of coupled nonlinear dynamical systems. These studies assist a better understanding of these systems in living organisms (Murray 2002). Although the analytical problems associated to these systems are far from trivial, the understanding of coupled dynamical systems has advanced (see e.g. (Keener and Sneyd 1998; Lindner et al. 2004; Murray 2002; Strogatz et al. 1992; Winfree 1967, 2000)), demonstrating that

the combination of nonlinear dynamics approaches with methods from synthetic biology aids in a better understanding of organisms and in generating new applications in biotechnology.

1.2 Thesis Objectives

In this thesis, we study the dynamics of two small genetic circuits, namely oscillators and switches, with the use of models.

The first circuit is studied at the population level, assuming coupling between circuits in individual cells. In this context, we investigated two coupling mechanisms and how these lead to various sorts of synchronization of behaviors. The main objective of this study, which is the first objective of this thesis, was thus to investigate the dynamics of models of coupled oscillators with phase-attractive (García-Ojalvo et al. 2004) and phase-repulsive (Ullner et al. 2007) coupling, by means of the bifurcation analysis. In particular, we aimed to determine the role of the network structure, if any, on the type of synchronization.

The second objective, which follows from the previous study, was to determine how the coupling modes affect the dynamics of the individual component oscillators.

Next, making use of more detailed models of gene expression, we focused on the study of the effects of sequence-dependent events during transcription and translation elongation on the dynamics of gene expression. Namely, we studied how various sequence-dependent features of gene expression affect the dynamics of genetic toggle switches, using a model of prokaryotic gene expression at the single nucleotide and codon levels (Mäkelä et al. 2011).

Finally, we studied in great detail one of these sequence-dependent mechanisms of gene expression regulation. Different codons, coding for the same amino-acid (thus, named 'synonymous codons'), have different speed of incorporation of the amino-acid (Sørensen and Pedersen 1991). It is thus possible to code for the same protein using different codon sequences that will cause their process of translation elongation to differ significantly in duration. Relevantly, given the sequence-dependence, the profile of codon translation rates can be subject to selection (Tuller et al. 2010).

In this last study, we investigated the effects of using different codon sequences on translation rate profiles and, consequently, on the dynamics of gene networks, using the model of gene expression at the single nucleotide and codon levels (Mäkelä et al. 2011). In particular, we compared the dynamics of single gene expression and the dynamics of genetic switches and oscillators with differing randomly generated codon sequences. Also, we studied the

effects of adding to these models slow ramps of codons (Tuller et al. 2010) at the start and at the end of the elongation regions of the genes as these sequences were found significantly above chance in several organisms from different kingdoms of life.

1.3 Thesis Outline

This thesis is organized as follows. Chapter 2 briefly introduces the biological background with emphasis on gene expression and gene networks in bacteria. Also introduced are the fundamental aspects of quorum sensing communication in bacteria. Next, Chapter 3 introduces the theory of bifurcation analysis and stochastic modeling, and how these are used in the study of gene expression and genetic circuits dynamics. This chapter presents illustrative examples of switching and oscillatory gene networks and their main features. Also presented are the main advantages and disadvantages of applying bifurcation analysis and stochastic modeling in the study of their dynamics. Subsequently, Chapter 4 presents the models and simulation techniques used in the publications that compose the thesis, namely, the models of genetic repressilator with quorum sensing and the model of gene expression at the single nucleotide and codon levels. Finally, in Chapter 5 we present the conclusions and final discussion.

Chapter 2

Biological Background

This chapter is an overview of the fundamental biological concepts and processes that are studied in this thesis. These include gene expression kinetics at the sequence level, gene regulatory networks, and quorum sensing in bacteria.

2.1 Gene Expression in Bacteria

Gene expression is the process of synthesis of a functional RNA or protein, from the information encoded in a gene's sequence (Lewin 2008). This process constitutes the central dogma of molecular biology, which states that the information stored in the DNA (genes) is transcribed into a corresponding RNA sequence which, finally, in case of protein-coding genes, is translated into an amino-acid sequence, that will become, after post-translational modifications, a protein. There are several known exceptions to the central dogma as described here, but they do not apply to bacterial gene expression (Crick 1970; Lewin 2008), which is the core process of the models studied in this thesis.

Transcription is the process by which a complementary mRNA strand is created according to a gene's DNA sequence, by the RNA polymerase enzyme. The subsequent process, *translation*, is the process by which the amino-acid sequence is assembled according to the mRNA sequence, by ribosomes. This 'conversion' of ribo-nucleotides of the mRNA to amino-acids is based on a triplet-wise code, that is each triplet (codon) of nucleotides encodes for an amino-acid. The recognition of codons is done by tRNAs. The genetic code is said to be degenerated or redundant, in that some amino-acids are encoded by more than one codon sequence (Alberts et al. 2002).

Transcription and translation in bacteria are coupled, in that translation

can start as soon as the initial region of the mRNA, the Ribosome Binding Site (RBS), is formed. The RBS region is sufficient for a ribosome to recognize and bind to the growing mRNA strand and initiate translation. This coupling only occurs in prokaryotes, which lack nuclei. In eukaryotes, transcription only occurs inside the nucleus, while translation only occurs outside, which implies that the two processes are not dynamically coupled (Alberts et al. 2002; Lewin 2008).

This coupling has many relevant consequences. For example, the dynamics of protein production can be, to some extent, regulated by events that occur in transcription. Due to this, in many instances, the coupling between transcription and translation cannot be ignored in dynamical models of genetic circuits in bacteria.

2.2 Gene Networks

Genes interact with each other rather than acting as independent units. These interactions can occur via the protein products or by other means. The interactions result in changes of the rates of synthesis of genes' products. The set of genes and connections between them constitute the gene regulatory network (GRN). Relevantly, genes are not connected at random, rather, the GRN structure reflects functional, evolutionary conserved relationships within the genome (Cheng et al. 2011; de la Fuente et al. 2002; Thieffry et al. 1998).

The use of next generation high-throughput technologies has led to many studies that begin to reveal the structure of natural gene regulatory networks (Cheng et al. 2011; de la Fuente et al. 2002; Park 2009). Some of these studies have revealed that there are characteristic one-, two- and three-gene motifs, which are structures that exist in an amount 'higher than by chance' in the GRN (Cheng et al. 2011; de la Fuente et al. 2002).

The interactions between genes can be 'positive', meaning that the expression rate of the target gene is increased when the modulator is present. This process is referred to as 'activation' and, in the case of protein coding genes, the modulator is an activator protein. The interactions can also be 'negative', in that the binding of the modulator represses the expression rate of the target gene. In this case, the modulator is a repressor molecule.

The structure formed by the various regulatory links, including the underlying motifs, defines the architecture, or topology, of the network. In the case of motifs, it is this structure that makes possible switching and temporal oscillatory behaviors of protein numbers of the constituent genes. Commonly observed motifs are self-repressing genes, two- and three-gene networks with

negative feedbacks, among others (Cheng et al. 2011; de la Fuente et al. 2002).

Generally, cyclic structures of unidirectional connections with odd number of negative relationships generate oscillatory dynamics, while an even number of negative relationships produces a switching behavior (Müller et al. 2006). These theoretical predictions have been realized experimentally by using the examples of two- and three-gene networks (Elowitz and Leibler 2000; Gardner et al. 2000).

As the knowledge about natural networks increases, several approaches of increasing complexity have been proposed to model them, namely: Boolean networks (Aldana 2003; Kauffman 1969), Ordinary Differential Equations (see e.g. (Süel et al. 2006)), Bayesian networks (Perrin et al. 2003), stochastic models (Arkin et al. 1998; McAdams and Arkin 1997), and delayed stochastic models (Ribeiro et al. 2006; Roussel and Zhu 2006). Finally, recently, supported by new experimental findings, the stochastic models of gene expression have evolved to include, explicitly, the processes of transcription and translation elongation (Mäkelä et al. 2011; Ribeiro et al. 2009a), along with many of the events that can occur at these stages, such as transcriptional pauses (Rajala et al. 2010).

2.3 Quorum Sensing

Bacteria communicate between them by means of signaling molecules (Miller and Bassler 2001). The information supplied by these signals allow, among other things, to synchronize large populations of bacteria of the same species. However, there is also communication between different species (Waters and Bassler 2005).

The signaling hormone-like molecules are referred to as autoinducers (AI) (Waters and Bassler 2005) and their concentration in the extracellular space is dependent on, among other factors, the population's density, hence the name of the coupling mechanism — Quorum Sensing (QS). Bacteria are able to assess the concentrations of AI molecules in their surrounding and, when above certain thresholds, they can alter gene expression levels and, hence, behavior in response.

Gram-negative bacteria possess the LuxIR-systems, whose name derives from the first thoroughly described example, the QS system in bacterium *Vibrio fischeri*, which is considered as a paradigm for all gram-negative bacteria (Nealson and J. 1979). LuxI is the AI synthase, which produces the acyl-homoserine lactone (AHL) AI (Engebrecht and Silverman 1984), while LuxR is the cytoplasmic AI receptor/DNA-binding transcriptional activator.

Once produced, AHL diffuses freely in and out of the cell and, as it is produced by each bacterium, its concentration increases with the population's density (Kaplan and Greenberg 1985). When the signal reaches a critical concentration it becomes very likely to be bound by LuxR. Once this occurs, this complex activates the transcription of target genes, e.g., in the case of *V. fischeri*, the operon encoding luciferase (Stevens et al. 1994), which produces luminescence. The LuxR-AHL complex also induces the expression of *luxI*, encoded in the luciferase operon, providing a positive feedback loop. Therefore, eventually, virtually all cells in the population become involved in the generation of light, a classical demonstration of quorum sensing behavior.

The majority of the gram-negative bacteria possesses this LuxIR-system of QS. The AHL signals differ among species and exhibit high specificity to the cognate LuxR proteins. LuxI proteins synthesize specific AHL signaling molecules with high fidelity (Waters and Bassler 2005). Here, we chose to omit the description of the QS system of gram-positive bacteria, since most synthetic biology studies are performed on *Escherichia coli*, which belongs to the gram-negative bacteria group.

The components of the QS network can be divided into three broad functional groups: (1) signaling molecules (AI), their synthases and machinery for secretion and processing of AI; (2) optional AI receptors, signal transduction elements and intermediate transcription regulators; (3) QS transcription regulators — master transcription factors regulating the expression of target genes (Goryachev 2011).

The dose-response curves characterize the functional role of the QS network — the regulation of copy number of the master transcription factors in response to extracellular AI concentration. The network may operate as a rheostat by gradually increasing the copy number of transcription factors in response to the AI signal. This behavior is characterized by S-shaped curves with only one output value for every input value. These curves can be well fitted by the Hill function, where the steepness of the curve is accounted for by the Hill coefficient. The higher the value of the Hill coefficient, the steeper the curve. For high values of the Hill coefficient, the response is said to be "ultra-sensitive" (Goryachev 2011).

Another form of the response is termed 'bi-stability', when there are cells where the process is clearly "on", and cells where it is clearly "off". This differs from the rheostat case, where most cells exhibit intermediate states. In the case of bi-stability, the dose-response curve consists of two disjoined branches of stable steady states and the network operates as a switch (Goryachev 2011).

Chapter 3

Theoretical Background

This chapter introduces the theoretical concepts used in the thesis. These range from dynamical systems as represented by systems of ODEs or according to a stochastic formulation of the chemical kinetics, to the tools used to analyze their dynamics, e.g., bifurcation analysis.

3.1 Dynamical Systems

Living organisms are dynamical systems. Rigorous studies of these systems' kinetics have shown that they can undergo bifurcations, i.e. sudden changes in the dynamics when a parameter of the system is varied, generating new dynamical behaviors (Kuznetsov 2004). In this thesis we study the switching and oscillatory behaviors in this perspective, as these two behaviors are present in a multitude of other biological processes, including the regulation of gene expression (Jacob and Monod 1961), the adoption of a pathway of differentiation (Chang et al. 2006, 2008), periodic biochemical reactions (Boiteux et al. 1975; Chance et al. 1964; Ghosh and Chance 1964; Hess 1979; Pye and Chance 1966), brain rhythms generating animal gaits (Collins and Stewart 1993), synchrony of firefly flashes (Buck 1988; Hanson 1978), and heart periodic activity (Honerkamp 1983), among many others.

A *dynamical system* is a mathematical formalization that models physical processes (including in living organisms) that change in time. The dynamical system consists of a set of variables, describing the state of the system, whose values are ruled by an evolution law — an operation which allows the development of the system's state from any current moment of time to the next time moment, given the current moment state and, possibly, states of the previous time moments. Note that the definition covers both the deterministic and stochastic dynamical systems. In the latter ones the evolution rules

are not deterministic and the next time moment state is "determined" by the statistical laws. Note also that the definition is not restricted by so called "memory-less" processes, where the current time moment state depends only on the state of the previous time moment (for example, Markov processes).

The deterministic dynamical systems are usually represented as a set of differential equations, where derivatives with respect to time describe the evolution of the system. In practice, in general, first order ordinary differential equations are the most commonly used ones, if the system is to be modeled only in time domain (see e.g. (Elowitz and Leibler 2000; Gardner et al. 2000; Hodgkin and Huxley 1952; Lotka 1909; Murray 2002; Süel et al. 2006; Winfree 2000)).

An Ordinary Differential Equation (ODE) is a differential equation comprising a function of only one independent variable and the function's derivatives with respect to the variable. This is opposed to Partial Differential Equations, which possess the partial derivatives of functions of more than one independent variable.

ODEs are used to characterize many natural, biological and social dynamic phenomena. Namely, they describe the changes of an entity, e.g. as a function of time. The differential equation is the result of assembling the derivatives and different functions into an equation, that connects the entities and their rates of change.

Many fields of science widely use ODEs, for example, they are broadly used in physics (Lindner et al. 2004; Schrödinger 1926; Strogatz 2001), chemistry (Glansdorff and Prigogine 1971; Prigogine and Lefever 1968; Strogatz 2001; Zhabotinsky 1991) and biology (Hodgkin and Huxley 1952; Izhikevich 2007; Lotka 1909; Strogatz 2001).

The general explicit form of an ODE is:

$$F(x, y, y', \ldots, y^{(n-1)}) = y^{(n)}\,, \tag{3.1}$$

where F is the function of x, y and derivatives of y with respect to x ($y', \ldots, y^{(n)}$ are the first,..., n-th order derivatives). Note x, y and $y^{(n)}$ are scalars for the system of a single ODE. The solution of the ODE is the function $u(x)$ which, by being the substitution for the $y(x)$ in eq. (3.1), fulfills the equality. The order of an ODE is the order of the highest derivative associated with the equation. An autonomous ODE is such that F is independent of x. Throughout the thesis we use the term ODE when referring to first order autonomous ODEs, with time as an independent variable.

Since ODEs describe the changes of variables over time (i.e., the evolution of the state of the system), they are commonly used in the description of the temporal dynamics of chemical species (chemical kinetics) (Izhikevich 2007;

Kuznetsov 2004). The ODEs of chemical kinetics are derived from the law of mass action, which was introduced by Guldberg and Waage (1864–1879) and describes the rates of elementary reactions as well as chemical equilibrium, which is a dynamic process resultant from having identical forward and backward reaction rates.

If F is a linear function of its arguments, the ODE is termed a linear ODE. A linear ODE, or a system of these equations, can be analytically solved and the steady state is the only qualitative solution possible for such systems[1]. More interesting perhaps are the nonlinear ODEs, which have non-trivial solutions that cannot be found analytically. Thus, to solve this type of equations one needs an approximate numerical method. There are several numerical methods for solving differential equations, for example, the Euler and the Runge-Kutta methods (Butcher 2008), which have been implemented in some numerical packages for analyzing systems of ODEs (e.g. (Ermentrout 2002)).

Modeling the kinetics of chemical processes with ODEs requires a certain number of assumptions. One such assumption is that the concentrations of the species are assumed to be continuous variables. The results of these models are, in general, valid for systems with a large number of reactant molecules. Due to the deterministic nature of these models, for the same initial conditions, the resulting kinetics is always the same.

3.2 Bifurcation Analysis

As it was mentioned previously, the deterministic dynamical system consists of the state of the system and the law of evolution, and is written in the form of ODEs. The set of all states that the system undergoes, as determined by an initial state and the evolution law, is named an *orbit* in the state space. A set of orbits of a dynamical system is named *phase portrait*. Note that such geometrical representation of the dynamical system, while excluding time, is a powerful instrument for analyzing the behavior of any system, and it is, thus, commonly used in the scope of dynamical systems theory. The geometrical exploration of a dynamical system's state space allows us to interrogate the system without needing to know the details of the evolution law and the equations defining it (Izhikevich 2007).

Next, we introduce the notion of an *attractor* of the dynamical system. The attractor is the subset of states to which orbits are attracted in time,

[1] excluding the special case of closed trajectories in linear systems, so called 'center' equilibrium points, producing sinusoidal oscillations

either going to ∞ or $-\infty$. When the attractor is reached in forward time ($t \to \infty$) it is called *stable*, oppositely, when the attractor is reached in backward time ($t \to -\infty$) it is called *unstable*.

Bifurcation theory is the field of mathematical sciences that studies qualitative and topological changes of the attractors of systems of differential equations (Kuznetsov 2004). The bifurcation occurs when a topologically nonequivalent phase portrait arises under variation of parameters. The parameter being varied is called the bifurcation parameter and the parameter value for which the bifurcation takes place is called the bifurcation (critical) value. Rigorous mathematical definitions and theorems on equivalence of phase portraits can be found in (Kuznetsov 2004). A *bifurcation diagram* is a means to show the characteristic attractors and phase portraits of the system together with the bifurcation parameters. There are several numerical packages available for conducting bifurcation analyses (Dhooge et al. 2003; Doedel 1981; Ermentrout 2002).

Here, we conduct bifurcation analysis procedures for models of the genetic toggle switch and of the repressilator in the subsequent sections. These analyses demonstrate the emergence of the switching and oscillatory dynamical behaviors, which are the main subject of the thesis.

3.2.1 Switching Kinetics

Let us consider the model of a genetic toggle switch as in (Gardner et al. 2000). The system of equations is the following:

$$\begin{aligned} \frac{du}{dt} &= \frac{\alpha}{1+v^n} - u \\ \frac{dv}{dt} &= \frac{\alpha}{1+u^n} - v \end{aligned} \tag{3.2}$$

The variables u and v are the continuous, dimensionless concentrations of the two proteins of the toggle switch model, α is the rate of synthesis in the absence of the repressor, and n is the Hill coefficient, which determines the cooperative effect of the repressor proteins. Negative terms denote the process of degradation of each protein. Note that introducing identical parameter values (α and n) for the two equations of the system, renders the system symmetric (Edelstein-Keshet 1988; Kaplan and Glass 1995; Rubinow 1975; Yagil and Yagil 1971).

Let us choose α as a bifurcation parameter and fix the other parameters ($n = 2$). We start the bifurcation analysis from small values of $\alpha = 0.1$ and proceed until $\alpha = 10$. The results are shown in Fig. 3.1.

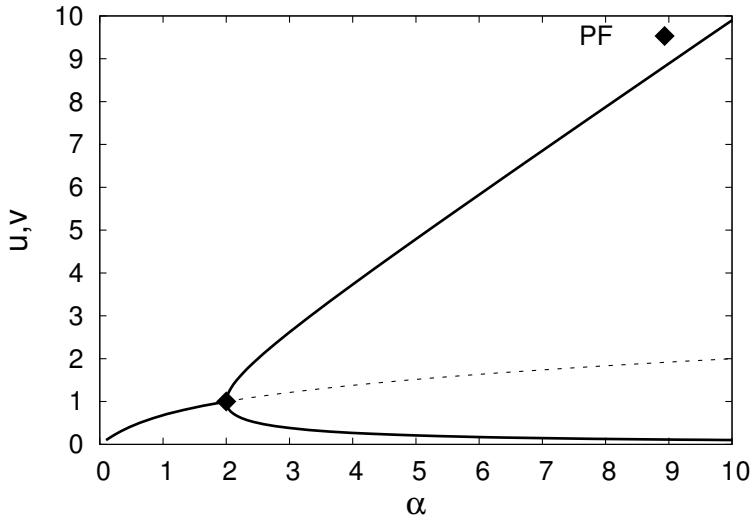


Figure 3.1: Bifurcation diagram of the system (3.2).

In Fig. 3.1, stable and unstable fixed point attractors are denoted as solid and dashed lines, respectively. From the figure, one can see that the system has a single stable steady state, for small values of α. As α increases, the system undergoes the pitchfork bifurcation (PF in the figure) at which point the initially stable attractor becomes unstable. Instead, after the PF point, two new stable fixed point attractors emerge. They differ significantly in protein levels and this difference increases with increasing α.

Consider now the regime of small values of α, before the pitchfork bifurcation (Fig. 3.1). The kinetics of the proteins numbers for one initial condition and the phase portrait of the system are shown in Fig. 3.2 for $\alpha = 1$. Both graphs in this figure show the existence of a single fixed point attractor. The stability of the attractor can be seen from the kinetics plot (Fig. 3.2, left). Proteins concentrations, u and v, eventually settle down at the attractor, which, thus, is called *equilibrium*, since the variables remain there forever. In the phase space, different orbits, corresponding to different initial conditions, lead to this stable attractor (Fig. 3.2, right).

The structure of the phase plane changes for α values following the pitchfork bifurcation. The kinetics for a single initial condition and the phase portrait of the system are shown in Fig. 3.3 for $\alpha = 8$. In this regime, the system possesses two stable equilibrium points. In each of the two equilibrium points, one variable has a high value and the other has a low value, which can be seen from the phase portrait of the system (Fig. 3.3, right). Due to the symmetry either of the variables can be in the 'high' or 'low' levels. Therefore, two equilibrium points arise. In every single realization of

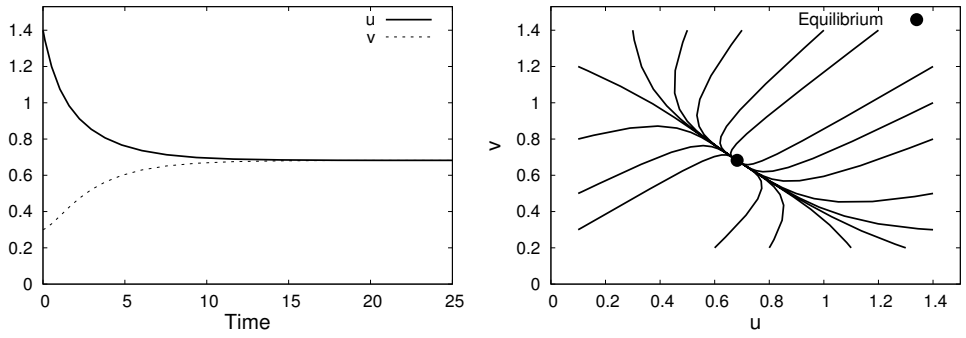


Figure 3.2: The kinetics of protein concentrations and the corresponding phase portrait of the system (3.2) for $\alpha = 1$, before the pitchfork bifurcation value.

the process, the system goes to one of the attractors, as shown in the kinetics plot (Fig. 3.3, left).

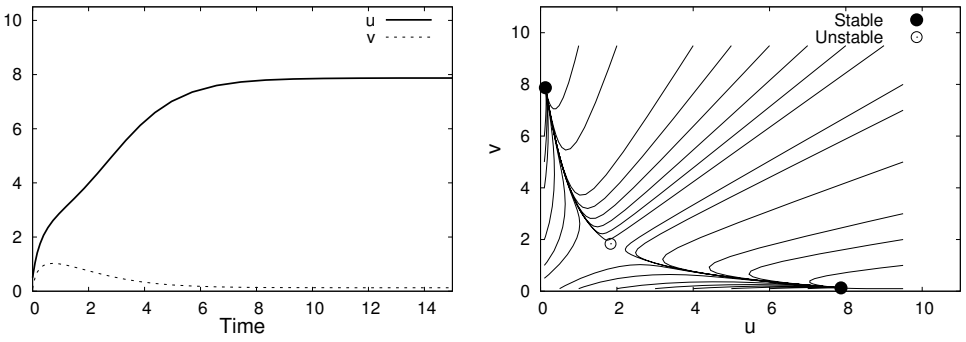


Figure 3.3: The kinetics of protein concentrations and the corresponding phase portrait of the system (3.2) for $\alpha = 8$, after the pitchfork bifurcation value. The initial conditions for the kinetics plot are $u(0) = 0.5$ and $v(0) = 0.1$. "Stable" and "Unstable" stand for stable and unstable fixed point attractors, respectively.

The set of initial conditions leading to a certain attractor is called its *basin* of attraction. From the phase portrait depicted in Fig. 3.3 one can see that the initial conditions can be divided into two sets, depending on which of the two attractors they lead to. These two sets of initial conditions are the basins of attraction of the corresponding attractors. The system also possesses the additional unstable fixed point attractor, which endows it with a topological confine between the two basins of attraction of the stable equilibrium points (Fig. 3.3).

Each of the stable equilibria of the system correspond to a state of the

toggle switch. At any particular moment in time, only one of the states is realized. The switching between these states can occur by fluctuations in the numbers of proteins u and v. These can be due to external signals, or caused by internal fluctuations. Depending on the system that is being modeled, the model switch can also be made dependent on temperature changes, 'pulses' of inducers, light conditions, etc (Gardner et al. 2000; Grams and Thiel 2002; Levskaya et al. 2005).

3.2.2 Oscillatory Kinetics

Oscillations in dynamical systems generated by ODEs correspond to closed trajectories in phase space (usually, limit cycle attractors). For obvious reasons, the dimension of the phase space must be equal to or larger than two, for the system to have a limit cycle attractor. When the equilibrium loses its stability, a limit cycle can emerge via the Andronov-Hopf bifurcation (Hopf bifurcation for short) (Kuznetsov 2004; Marsden and McCracken 1976).

There are two types of the Hopf bifurcation — supercritical and subcritical. The supercritical Hopf bifurcation takes place when the stable equilibrium loses stability and gives rise to a small amplitude, stable limit cycle attractor. In the case of subcritical Hopf bifurcation, initially three attractors persist: a large amplitude, stable limit cycle, a small amplitude, unstable limit cycle and, finally, a stable equilibrium. The small amplitude, unstable limit cycle shrinks to a stable equilibrium and makes it lose stability. Thus, initially, the system has two stable attractors separated in the phase space by an unstable limit cycle. After the subcritical bifurcation, only large amplitude, stable limit cycle persists. However, instead of the large amplitude, stable limit cycle there can be any other stable attractor in the system (Izhikevich 2007).

The two types of Hopf bifurcations are shown in Fig. 3.4. In the case of the supercritical bifurcation, only one of the dynamical regimes is stable: either the steady state or the oscillation. In the case of the subcritical bifurcation, there is a region of coexistence of two stable dynamical regimes: the steady state and the oscillations (or another kinetics, corresponding to any other stable attractor). The unstable limit cycle divides the phase space into two basins of attraction, corresponding to the two stable dynamical behaviors (Fig. 3.4).

As an example of the oscillatory dynamics, let us consider the model of a genetic oscillator presented in (Elowitz and Leibler 2000). This oscillator, named *repressilator*, consists of three genes, each synthesizing repressor protein, which inhibits the transcription process from the neighboring gene. The genetic circuit is arranged so that the inhibitory links form a cyclic,

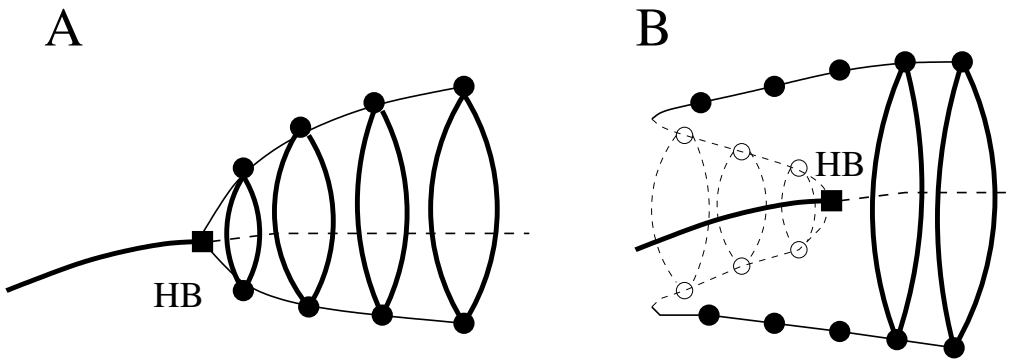


Figure 3.4: Two types of Andronov-Hopf bifurcations (HB): supercritical (A) and subcritical (B). Stable and unstable fixed point attractors are denoted as solid and dashed line, respectively. Stable and unstable limit cycles are denoted as solid and empty circles, respectively.

unidirectional structure, as shown in Fig. 3.5.

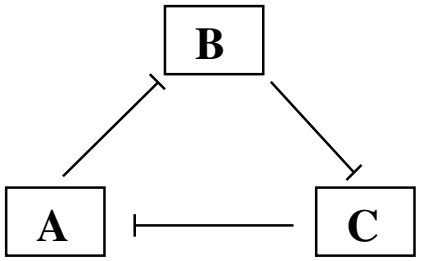


Figure 3.5: Three genes inhibit the transcription of each other in a unidirectional cycle.

The model in (Elowitz and Leibler 2000) describes the dynamics of mRNA and protein species and includes the ODEs for each of these. The following notations are used: lowercase letters a, b and c stand for mRNA concentrations and uppercase letters A, B and C are the corresponding protein concentrations (eqs. (3.3-3.4)).

$$\begin{aligned} \frac{da}{dt} &= -a + \frac{\alpha}{1+C^n} + \alpha_0 \\ \frac{db}{dt} &= -b + \frac{\alpha}{1+A^n} + \alpha_0 \\ \frac{dc}{dt} &= -c + \frac{\alpha}{1+B^n} + \alpha_0 \end{aligned} \tag{3.3}$$

mRNA degradation is reflected by the negative terms in the right hand side of the eqs. (3.3). Repression is reflected by the nonlinear function $\frac{\alpha}{1+p_j^n}$, where p_j is the concentration of the corresponding repressor protein, n is the Hill (cooperativity) coefficient, α is the maximal transcription rate. Note that the same repression function was used in the model of the toggle switch in Section 3.2.1. The leakiness of the promoters is modeled by a constant rate of transcription, α_0.

$$\begin{aligned} \frac{dA}{dt} &= \beta(a - A) \\ \frac{dB}{dt} &= \beta(b - B) \\ \frac{dC}{dt} &= \beta(c - C) \end{aligned} \tag{3.4}$$

Proteins are synthesized and degraded linearly, according to eqs. (3.4), where β is the ratio between protein and mRNA degradation rate constants.

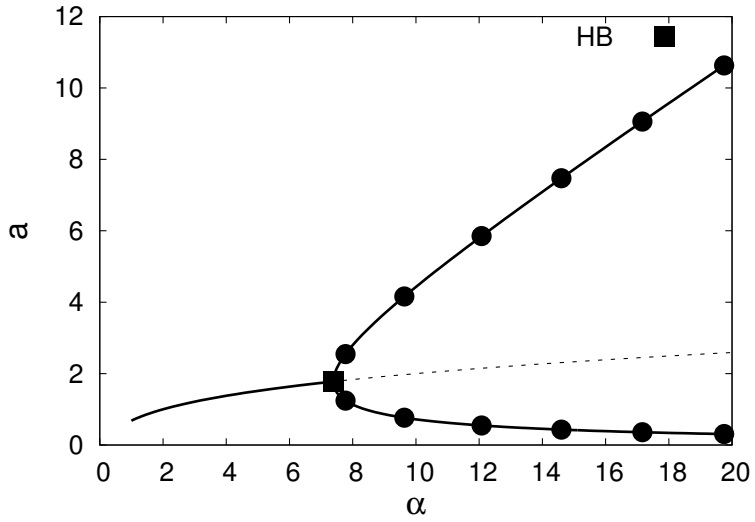


Figure 3.6: Bifurcation diagram of the system (3.3-3.4). Parameters used in this analysis were: $n = 2$, $\beta = 0.2$ and $\alpha_0 = 0$.

The bifurcation diagram of the repressilator is shown in Fig. 3.6. The transcription rate α is used as a bifurcation parameter,while the other parameters are fixed (Fig. 3.6). Solid and dashed lines denote the stable and unstable stationary (fixed point attractor) solutions of the system, respectively. The oscillatory solution is represented with solid circles. Only one

variable, corresponding to mRNA a, is shown. One can see in Fig. 3.6 that the system possesses the stationary dynamics for small enough values of the transcription rate α, for which the phase portrait and the kinetics of the system do not qualitatively differ from those shown in Fig. 3.2 for the genetic toggle switch.

With increasing α, the stationary mRNA and protein levels increase until the supercritical Hopf bifurcation occurs (HB in Fig. 3.6). After the HB, the fixed point attractor loses stability and a small amplitude limit cycle emerges. For the limit cycle, maximum and minimum values of the variable are shown in Fig. 3.6. As α increases, the amplitude of the limit cycle grows (Fig. 3.6). The corresponding kinetics plot and the phase portrait are depicted in Fig. 3.7.

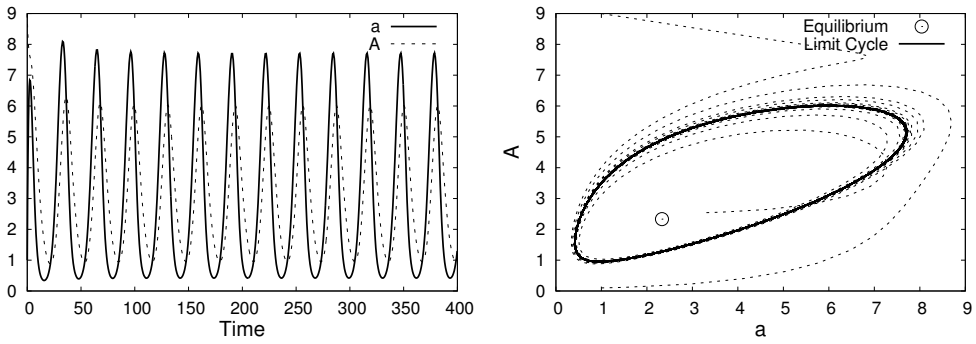


Figure 3.7: Example of the kinetics of variables a and A and the phase portrait of the system (3.3-3.4), after the Hopf bifurcation. Parameter values: $\alpha = 15$, $n = 2$, $\beta = 0.2$, $\alpha_0 = 0$. On the phase portrait (right), the limit cycle and several trajectories are shown in solid and dashed lines, respectively.

The oscillatory behavior of the variables of the system with a certain periodicity and amplitude can be seen from Fig. 3.7. The kinetics plot (Fig. 3.7, left) shows the periodic changes in concentrations of mRNA a and protein A. The phase portrait (Fig. 3.7, right) shows the limit cycle corresponding to this oscillatory behavior. From this plot, it can be seen that all orbits (dashed lines) end up at the limit cycle attractor (solid line). Note that the equilibrium of the system is unstable for the given parameter set (Fig. 3.7).

3.3 Stochastic Kinetics

The deterministic approach can capture qualitative changes in dynamical systems, for example, the structural rearrangements of phase portraits, the trajectories of entities in space, changes in rhythms or the generation of novel dynamics through bifurcations. However, in many situations, the "stochasticity" or "noise" underlying the kinetics of the processes in real systems cannot be neglected. Due to its significance, realistic dynamical models of genetic circuits must include this component.

Noise in a system's dynamics has two major sources. One is "intrinsic", i.e. the system under study and its internal properties are the cause for this type of noise. For example, it arises when micro-constituents of the system undergo thermal fluctuations that affect the macro-scale constituents. The finite number of macro-variables used in mathematical modeling inevitably leads to the existence of this type of noise (van Kampen 2007). Additionally, fluctuations in copy numbers of the key participating molecules of the chemically reacting system are also a source of intrinsic noise (Gillespie 1977). The second source of noise is the surrounding environment of the system in question. The environmental conditions are themselves stochastic. This stochasticity can arise from fluctuations in temperature, nutrient supplies for cells, solar radiation, etc (Horsthemke and Lefever 1983).

Stochastic differential equations in general and the Langevin equation (Lemons and Gythiel 1997) in particular account for noise. The Langevin equation can be written in the following general form:

$$\frac{d\mathbf{x}}{dt} = f(\mathbf{x}(t), t) + g(\mathbf{x}(t), \xi(t), t)\,, \tag{3.5}$$

where $\mathbf{x}(t) = [x_1(t), \ldots, x_n(t)]$ is the state vector of the dynamical system; $\xi(t) = [\xi_1(t), \ldots, \xi_n(t)]$ is a set of random forces affecting the system. The differential equation (3.5) can be divided in two parts: deterministic $f(\mathbf{x}, t)$ and stochastic part $g(\mathbf{x}, \xi, t)$, such that $g(\mathbf{x}, \xi = 0, t) = 0$. Note that, if $g(\mathbf{x}, \xi, t) = 0$ equation (3.5) becomes ODE. If $g(\mathbf{x}, \xi, t) \neq 0$, the trajectories for the same initial condition differ from each other and represent different realizations of the same stochastic process. Thus, one needs to consider the statistical ensemble consisting of many trajectories for an adequate description.

In the case of intrinsic noise, the intensities of the stochastic fluctuations satisfy specific relationships according to the fluctuation-dissipation theorem (Callen and Welton 1951; Nyquist 1928), which correlate the intensities of the fluctuations to the dissipative properties of the system (Kubo et al. 1985). For example, the Einstein-Smoluchowski relationship connects

the diffusion coefficient (stochastic component) with the viscosity coefficient (deterministic component). Similarly, the Nyquist theorem states that the noise in an electrical circuit depends on the resistance of the circuit.

In the case of extrinsic noise, the stochastic and deterministic components are independent from one another and, thus, the features of the fluctuations can be considered as additional parameters of the system (Hänggi and Thomas 1982). Another approach to account for extrinsic noise is to add uncertainty to the kinetic parameters of the deterministic system, by drawing them randomly from particular distributions (see, e.g. (Mondragón-Palomino et al. 2011)).

An alternative definition of a random process can be made in terms of probability distributions and probability density functions (p). In this approach, the evolution of the system is described by the deterministic equations that have probabilities as variables, rather than the state of the system. One can use differential equations to describe the evolution of the system with continuous time as follows,

$$\frac{\partial}{\partial t}p = \mathbf{L}p\,, \tag{3.6}$$

where $\mathbf{L}$ is a linear evolution operator and p is the probability density function. Usually, $\mathbf{L}$ is set so that it describes the Markov process. That is, a process where each next state is determined only by the previous state, i.e. the system possesses no memory. In this case, one can derive the Chapman-Kolmogorov equation, which is the general equation for describing the evolution of a Markov process. One particular form of this equation is the master equation (van Kampen 2007).

These two, alternative, definitions of a stochastic process, based on the statistical ensemble of the realizations and based on the evolution of the probability density functions, provide two alternative approaches for analyzing stochastic systems.

3.3.1 The Stochastic Simulation Algorithm

In chemical reactive systems, the effects of fluctuations in molecule numbers are most 'visible' when the number of reactive elements is small. For example, gene expression is considered to be a highly stochastic process due to the small number of chemical species involved in this process (McAdams and Arkin 1997; Ozbudak et al. 2002). Namely, genes are usually represented by one to a few copies in the whole genome, there are only a few RNA-polymerases available for transcription at any given moment in the cell, etc. Due to this, it was proposed (Arkin et al. 1998; McAdams and Arkin 1997)

that the simulation of the dynamics of gene expression and of gene regulatory networks should be done according to the Stochastic Simulation Algorithm (SSA) (Gillespie 1977), initially developed for simple chemical reactive events.

The SSA is a Monte-Carlo simulation of the chemical master equation and thus is an exact procedure for the numerical simulation of the temporal evolution of a well-stirred chemically reacting system (Gillespie 1977, 2007; Karlebach and Shamir 2008). Each chemical species entity is treated separately and is a variable of integer type. Each reaction is simulated explicitly. Time advances in discrete time steps and, at every step, a reaction occurs and the numbers of all species involved in the reaction (reactants and products) are updated according to the reaction formula. The time step and the reaction at each step are determined according to formulas, each of which taking as input one of the two random numbers drawn from uniform distributions. At each step of the algorithm, the random numbers are drawn independently from any events occurring at previous time steps. Therefore, the algorithm can simulate Markov processes.

The algorithm requires the system to be well-stirred, meaning that it is assumed that there is a direct mixing of the reacting space and/or that the number of non-reactive collisions of molecules exceeds significantly the number of reactive collisions, allowing all components of the system to be distributed homogeneously in the space at all times (Gillespie 1977), immediately prior and immediately following any reaction event.

Each reaction rate c_μ in the SSA depends on the reactive radii of the molecules involved in the reactions and their relative velocities (Gillespie 1977, 2007). The velocities depend on the temperature and molecular masses. The algorithm computes propensities for each of the reaction as $a_\mu = c_\mu \cdot h_\mu$, where h_μ is the number of distinct molecular reactants combinations available at a given moment, h_μ depends on the number of molecules X_i of each of the species involved in the reaction μ. The SSA uses the unit-interval uniform random generator to compute τ, the time interval until the next reaction occurs, and μ, that determines the reaction that occurs. Finally, time is increased by τ and all molecular species X_i involved in the reaction occurred are updated according to the reaction formula. This process is repeated until no more reactions can occur or for a predefined time interval. The SSA runs as follows:

Step 0 (Initialization) Input the system's M reaction constants $c_1, \ldots, c_M$ and N numbers of initial molecular population numbers $X_1, \ldots, X_N$. Set time to zero and initialize the unit-interval uniform random number generator.

Step 1 Calculate and store M quantities $a_1 = c_1 \cdot h_1, \ldots, a_M = c_M \cdot h_M$ for the current molecular population numbers. Calculate and store the sum of all a_i values as a_0.

Step 2 Generate two random numbers r_1 and r_2 from a unitary uniform distribution and calculate τ and μ as follows: $\tau = \frac{1}{a_0} \cdot \ln(\frac{1}{r_1})$, and integer μ that satisfies the inequality $\sum_{i=1}^{\mu-1} a_i < r_2 \cdot a_0 \leq \sum_{i=1}^{\mu} a_i$.

Step 3 Using values of τ and μ obtained in Step 2, increase t by τ and adjust the molecular population numbers to reflect changes caused by the chosen reaction μ. If there is no stop condition satisfied, return to Step 1.

Several algorithms were proposed to increase the speed of the original SSA or to add additional features. For this, these algorithms usually assume restrictions and approximations (e.g., the Gibson-Bruck algorithm (Gibson and Bruck 2000)) or introduce new procedures as, for example, in the delayed SSA (Roussel and Zhu 2006).

As a side note, several of the players involved in the process of gene expression cannot be considered to be well-stirred as in a chemical reactor. First, some of these elements, the genes, exist in very small numbers and are integrated in a larger structure whose location in the cell is far from random. Also, the cell is a highly-ordered system with rigid temporal and spatial hierarchy of processes. Similarly, some evidence suggests that the RNA polymerases and the transcription factors do not freely diffuse through the cell, even in bacteria (Alberts et al. 2002; Lewin 2008). Nevertheless, the SSA has been used successfully to predict the kinetics of gene expression, particular in organisms such as *Escherichia coli* (Arkin et al. 1998; Karlebach and Shamir 2008; McAdams and Arkin 1997; Ozbudak et al. 2002).

3.3.2 Delayed stochastic models of gene expression

While the first stochastic models of gene expression treated transcription and translation as instantaneous processes (Arkin et al. 1998), it takes non-negligible time to produce mRNA and protein molecules, particularly when compared to the lifetime of the cell and other associated processes. Namely, following the initialization of these multi-stepped processes, several subsequent steps, such as transcripts assembly, mRNA translation, post-translational modifications, and folding, unavoidably follow in sequence. All these events are time consuming, particularly protein folding and activation, and events at the promoter region (Ribeiro 2010). To account for this consecutiveness and the duration of the processes, models were proposed that introduced

delays in the appearance of the expression products following the occurrence of the first step of these events (Bratsun et al. 2005; Ribeiro et al. 2006; Roussel and Zhu 2006). These delays ought to be drawn from distributions, each time an expression event occurs, to account for the variance in the durations.

In order to implement such models, a novel algorithm was necessary to drive the dynamics, since the SSA assumes all reactive events as instantaneous. One of such algorithms, allowing multiple delays in a single reaction, was proposed in (Roussel and Zhu 2006). This algorithm was the first to allow implementing the modeling strategy for gene networks proposed in (Ribeiro et al. 2006) and it was developed precisely with this aim.

The 'delayed SSA' (Roussel and Zhu 2006) uses a waiting list to store delayed output events. The waiting list contains a list of molecular species (e.g., proteins and RBS being produced) along with defined delays, time intervals for which they should stay in the list after the production reaction and before the release into the system. The algorithm proceeds as follows (Ribeiro 2010):

Step 0 (Initialization) Set time t to zero and define the stop time. Set initial numbers of molecules. Create an empty wait-list L.

Step 1 Generate the SSA step that gets the next reaction event R_1 and the corresponding occurrence time $t + t_1$.

Step 2 Compare t_1 with the least delay time in L, $\tau_{\min}$. If $t_1 < \tau_{\min}$ or L is empty, set $t \leftarrow t + t_1$. Subtract t_1 from all the delays in L. Update the number of molecular species in accordance with R_1, adding to L both the delayed products of the reaction and the time delays for which they have to stay in L.

Step 3 If L is not empty and if $t_1 \geq \tau_{\min}$, set $t \leftarrow t + \tau_{\min}$. Update the number of molecules by releasing the products from L with the delay equal to $\tau_{\min}$. Remove those products from L. Subtract $\tau_{\min}$ from all the remaining delays in L.

Step 4 Continue from Step 1, if time t is less than the stop time; otherwise stop.

Note that, after Step 3, the algorithm goes directly to Step 4 without performing the R_1 reaction generated in Step 1. In Step 4, it then jumps to Step 1, if time has not exceeded the stop time, where it generates a new reaction event rather than performing the previous reaction event that was not executed due to the delayed products release that occurred first. The previously generated reaction event R_1 is not performed since the delayed events

introduce new substances to the system and the reaction propensities might change. The algorithm instead generates new reaction event. This sequence of steps does not introduce errors as the process is memoryless (Ribeiro 2010).

Multi-delayed model of transcription and translation

The general model for transcription and translation with multiple delays was initially proposed in (Ribeiro et al. 2006). The proposed set of chemical reactions constituting the stochastic model of single gene expression was:

$$\mathrm{Pro} + \mathrm{RNAp} \xrightarrow{k_t} \mathrm{Pro}(\tau_1) + \mathrm{RBS}(\tau_1) + \mathrm{RNAp}(\tau_2) \tag{3.7}$$

$$\mathrm{Rib} + \mathrm{RBS} \xrightarrow{k_{tr}} \mathrm{RBS}(\tau_3) + \mathrm{Rib}(\tau_4) + \mathrm{P}(\tau_5) \tag{3.8}$$

$$\mathrm{RBS} \xrightarrow{d_{RBS}} \emptyset \tag{3.9}$$

The reactions (3.7) and (3.8) model prokaryotic transcription and translation, respectively. Pro represents the promoter region of a gene, RNAp is the RNA polymerase enzyme, and RBS is the ribosome binding site, i.e. the initial region of mRNA that the ribosome Rib can bind to and start translation. This explicit modeling of the RBS is required in prokaryotes, since translation can be initialized immediately after the RBS region is formed (τ_1 seconds after transcription starts).

The delays in the model are denoted as $X(\tau)$, meaning that the product X is delayed for τ seconds. This implies that the product X is released after τ seconds after the moment when the reaction occurs according to the procedure of the delayed SSA. Note that τ can be a random variable following a predefined distribution, thus varying from one reaction event to another.

The RBS undergoes degradation according to the reaction (3.9). The similar degradation reaction can be included for the protein, but it does not serve any function in the system, thus the reaction was omitted. However, this reaction is crucial when studying the dynamics of genetic networks and must be present in the model.

As genes are arranged in networks (see Sec. 2.2) the model of prokaryotic gene expression can be supplemented with additional chemical reactions reflecting interactions between genes (Ribeiro 2010). For example, one could include reactions for the repression where the protein (here referred to as 'Rep') coded by one gene exerts on another gene's activity:

$$\begin{aligned} \mathrm{Pro} + \mathrm{Rep} &\xrightarrow{k_{rep}} \mathrm{Pro} \cdot \mathrm{Rep} \\ \mathrm{Pro} \cdot \mathrm{Rep} &\xrightarrow{k_{unrep}} \mathrm{Pro} + \mathrm{Rep} \end{aligned} \tag{3.10}$$

The first reaction in (3.10) models the binding of the repressor Rep protein, while the second reaction models its unbinding. The formation of the complex Pro · Rep temporarily blocks the promoter Pro. Only when the repressor unbinds the promoter new transcription events become possible.

Note that due to these reactions express real events that occur in cells, all the rate constants and time delays can be directly obtained from experimental data (Ribeiro 2010), while ODE models, for example, have parameters that only indirectly can be extracted from the same data (e.g. the Hill coefficient).

The model accounts for the bursty production of proteins, which was reported experimentally in (Yu et al. 2006) for a highly repressed *lac* promoter. In this context, 'burst' implies that from one RNA, which is produced rarely in time, several proteins are translated, during the short life time of the RNA.

In our model, the strong repression can be accounted for by setting k_{rep} to a value large enough that is much bigger than k_{unrep}, by at least one order of magnitude. For example, one could set $k_{rep} = 1\ \mathrm{s}^{-1}$ and $k_{unrep} = 0.1\ \mathrm{s}^{-1}$. Relevantly, in this model, in accordance to the measurements, the distribution of the number of transcription initiation events fits a Poisson distribution and that of the number of translation reactions fits an exponential distribution (Ribeiro 2010).

Subsequent studies of this model (Zhu and Salahub 2008) focused on its dynamical properties in various conditions and studied what parameters most affect noise in RNA and protein levels. In particular, the transcription and translation initiation rates, the rates of mRNA and protein decay, the promoter and RBS delays were studied in this perspective (Ribeiro 2010). Usually, these studies use methodologies to measure noise following the approach proposed in (Paulsson 2005).

Ensemble approach for delayed stochastic gene networks

The significant progress in high-throughput technologies made it possible to identify most genes of organisms in reasonably short time. For example, these efforts revealed that the human genome has ∼30,000 genes. These technologies, which include, e.g. ChIP–seq and microarray, further allow studies of how the various genes interact. They revealed that genes interact in an intricate way, forming regulatory networks, by means of their products or via other substances in the cell, with complex large scale topologies (Chua et al. 2006; Ribeiro 2010).

However, these procedures, as well as the algorithms used for processing the data still lack accuracy. Thus, in order to understand the dynamics of the large scale topology gene networks the ensemble approach (Kauffman 2004) is still needed. This approach relies on the average topological features, rather

than detailed topologies, and provides average expected behaviors, rather than the behavior of a specific circuit. The use of large scale stochastic modeling is, nevertheless, also complicated, as it has high computational demands on the simulation of the network dynamics and it is not easy to establish the rules that the ensemble follows.

Several of these rules have been established. For example, it is known that the regulation of gene expression occurs via transcription factors or their oligomeric structures. These interactions define the topology of the gene regulatory networks. The genes have operator regions to which the transcription factors bind and, when doing so, change the rate of transcription of the gene.

This knowledge was used to establish a modeling strategy of gene regulatory networks (Ribeiro et al. 2006). In this strategy, the following notation is used for the promoter of a gene i: $\mathrm{Pro}_{i,\{\vec{\mathrm{op}}\}}$. The array $\{\vec{\mathrm{op}}\}$ is based on the set of the operator sites of the gene i, and it represents the state of each of the operator sites, i.e., if and which transcription factor is bound to the site (Ribeiro 2010).

Depending on the occupancy of the operator sites of the promoter, the gene is either repressed or activated at a certain level (in comparison to its 'basal level'). A fraction of genes can be assigned to have only a basal level of expression. For each combination of input states, which constitutes the promoter's state, a regulating function that determines the gene's expression rate is assigned (Ribeiro 2010).

The model of gene regulatory networks (Ribeiro et al. 2006) includes the following generalized chemical reactions:

$$\mathrm{Pro}_i + \mathrm{RNAp} \xrightarrow{k_{t,i}} \mathrm{Pro}_i(\tau_i^1) + \mathrm{RBS}_i(\tau_i^1) + \mathrm{RNAp}(\tau_i^2) \tag{3.11}$$

$$\mathrm{Pro}_{i,\{\vec{\mathrm{op}}\}} + \mathrm{RNAp} \xrightarrow{k_{t1,i,\{\vec{\mathrm{op}}\}}} \mathrm{Pro}_{i,\{\vec{\mathrm{op}}\}}(\tau_i^1) + \mathrm{RBS}_i(\tau_i^1) + \mathrm{RNAp}(\tau_i^2) \tag{3.12}$$

$$\mathrm{Rib} + \mathrm{RBS}_i \xrightarrow{k_{tr,i}} \mathrm{Rib}(\tau_i^3) + \mathrm{RBS}(\tau_i^4) + \mathrm{p}_i(\tau_i^5) \tag{3.13}$$

$$\mathrm{Pro}_{i,(\ldots,j,\ldots)} + \mathrm{p}_w \underset{k_{w,ij}}{\overset{k_{ij,w}}{\rightleftharpoons}} \mathrm{Pro}_{i,(\ldots,p_w,\ldots)} \tag{3.14}$$

$$\mathrm{Pro}_{i,(\ldots,p_z,\ldots)} \xrightarrow{d_{iz}} \mathrm{Pro}_{i,(\ldots,0,\ldots)} \tag{3.15}$$

$$\mathrm{Pro}_{i,(\ldots,p_z,\ldots)} + \mathrm{p}_w \xrightarrow{d_{iz,w}} \mathrm{Pro}_{i,(\ldots,0,\ldots)} \tag{3.16}$$

$$\mathrm{RBS} \xrightarrow{d_{RBS,i}} \emptyset\,, \qquad \mathrm{p}_i \xrightarrow{d_{p,i}} \emptyset \tag{3.17}$$

$$\mathrm{p}_i + \mathrm{p}_j \underset{k_{undim,ij}}{\overset{k_{dim,ij}}{\rightleftharpoons}} \mathrm{p}_{i,j} \tag{3.18}$$

For genes $i = 1, \ldots, N$ there is a basal transcription reaction of promoter Pro_i (3.11). The model includes transcription reactions for each promoter, with a specific set of transcription factors bound (3.12), and the translation of RNA (RBS) by ribosomes (Rib) into proteins (p) (3.13). The products of these reactions are all time-delayed. The delays are different between the products within each reaction and between the similar reactions for different genes.

The binding and unbinding of a transcription factor from operator site j of a gene i are represented by the reversible reaction (3.14). If the complex of promoter and transcription factor does not allow transcription to initialize, the two reactions represent repression and derepression. Derepression can also occur due to removing of the repressor from the operator site by another transcription factor via reaction (3.16). Decay of RNA/RBS and proteins are represented by uni-molecular reactions (3.17). Decay of a protein while bound to a promoter occurs via reaction (3.15). Finally, protein polymerizes (here, only dimerization is considered, for simplicity) and dissociate via the reversible reaction (3.18).

Ensembles of gene regulatory networks (Kauffman 2004) can be generated by randomly assigning integers for all indexes in the reactions modeling interactions between genes i, j, z, and w. Each different set of assignments corresponds to a unique gene regulatory network.

3.3.3 Stochastic models of small genetic circuits

The Genetic Toggle Switch

Different studies of genetic toggle switches suggest that the bi-stability arises from the cooperative binding of repressor proteins. However, in (Lipshtat et al. 2006) the authors concluded that the combination of the relevant biological conditions, suitable network features and stochastic effects give rise to bi-stability without the need for cooperative repression. Namely, three models that exhibit bi-stability were analyzed. The first circuit contains two promoter sites which cannot be occupied simultaneously. The second circuit has degradation of the repressor when bound to promoter. In the third circuit, two repressor proteins can form a complex which is not active as a transcription factor. Such models, when following the deterministic approach, as in Sec. 3.2.1, show bi-stability only in the two latter cases, while the first circuit exhibits only a single, stable steady state. When using the Master Equation approach, it was observed that all three circuits exhibited bi-stability (Lipshtat et al. 2006).

The bi-stability of the stochastic toggle switch is tightly associated to the

concept of noisy attractors (Ribeiro and Kauffman 2007), as this model was the one first used to exemplify this concept. The bi-stable genetic switch has two noisy attractors corresponding to high and low protein levels of either of the genes of the switch. Note that in each of these noisy attractors the number of proteins fluctuates, which implies that, technically, these are not attractors in a strict sense. These noisy attractors relate to the conventional concept of an attractor of the dynamical system (see Sec. 3.2), being constrained regions of the state space that the system remains in for a very long time.

Interestingly, even with the same structure of interactions, given sufficient noise in the protein numbers, the two genes of the switch can express simultaneously for long periods of time. Thus, within a specific region of parameter values, the toggle switch can be tri-stable (Ribeiro 2008). Both the bi-stability without cooperative repression and the tri-stability of the genetic toggle switch are now commonly used examples of noise-induced emergence of novel dynamics in genetic circuits.

The genetic toggle switch circuit was long ago hypothesized to be a potential decision circuit for cell differentiation (Jacob and Monod 1961). This hypothesis has been tested making use of delayed stochastic models. Namely, in (Ribeiro et al. 2009b), the authors studied how the stochasticity of the toggle switch can be used to regulate patterns of cell differentiation. Each cell contained a genetic switch whose dynamics was driven by the delayed SSA (Sec. 3.3.2). Assuming that the protein levels in each cell determine the cell's fate regarding differentiation at a given point in time, the authors discriminated between four possible differentiation pathways of the cell. The four states are: (1,0), (0,1), (1,1), and (0,0), where, for example, (1,0) implies that the first protein level is 'high' while the second protein level is 'low'. Differentiation between these levels was possible by the K-means algorithm due to their bi-stable nature.

The delayed stochastic model of the genetic toggle switch used in this study can be derived from the general model of an ensemble of gene regulatory networks (3.11-3.18). Taking into account only simplest interactions of transcription factors and DNA one can write the model of the genetic switch as follows:

$$\mathrm{Pro}_i + \mathrm{RNAp} \xrightarrow{k_t} \mathrm{Pro}_i(\tau_1) + \mathrm{RNAp}(\tau_2) + \mathrm{RBS}_i(\tau_1) \tag{3.19}$$

$$\mathrm{RBS}_i + \mathrm{Rib} \xrightarrow{k_{tr}} \mathrm{RBS}_i(\tau_3) + \mathrm{Rib}(\tau_4) + \mathrm{Rep}_i(\tau_5, \tau_{5,std}) \tag{3.20}$$

$$\mathrm{RBS}_i \xrightarrow{d_{RBS}} \emptyset\,, \qquad \mathrm{Rep}_i \xrightarrow{d_p} \emptyset \tag{3.21}$$

$$\mathrm{Pro}_i + \mathrm{Rep}_j \underset{k_{unrep}}{\overset{k_{rep}}{\rightleftharpoons}} \mathrm{Pro}_i \cdot \mathrm{Rep}_j \tag{3.22}$$

$$\mathrm{Pro}_i \cdot \mathrm{Rep}_j \xrightarrow{d_p} \mathrm{Pro}_i \tag{3.23}$$

In reactions (3.19-3.23) $i, j = 1, 2$ and $i \neq j$; Pro_i is the promoter of the i-th gene; RNAp is RNA polymerase; Rib is a ribosome and RBS_i is the RBS of the i-th gene. The reaction (3.19) models transcription. The translation reaction (3.20) accounts for variability in time needed to get functional protein Rep_i (translation elongation, folding, activation, etc.), given that the delay for Rep_i follows the normal distribution with the mean τ_5 and the standard deviation $\tau_{5,std}$ (Zhu et al. 2007). Each protein represses the opposite gene's promoter: the reaction (3.22) models binding and unbinding of the repressor from the promoter; this reaction defines the topology of the network, i.e. its switching properties. Reactions (3.21) and (3.23) model the degradation of proteins (including the degradation on the promoter (3.23)) and RBS.

The study (Ribeiro et al. 2009b) revealed that cells could potentially tune their pluripotency and distribution of lineage choice in a population by varying the transcription rate k_t and/or the protein decay d_p, suggesting that the stochastic switch has high plasticity regarding differentiation pathway choice regulation. For example, as the noise level changed, the significant changes in the pattern of cell differentiation were observed.

Genetic Repressilator

The genetic repressilator was the first synthetic circuit that was shown to exhibit periodic oscillations in the protein numbers (Elowitz and Leibler 2000). The first attempt to model the repressilator circuit accounting for stochasticity was made in the original work (Elowitz and Leibler 2000). The model was driven by the SSA (Gillespie 1977) and was used as a comparison to the deterministic model considered in this thesis (see Sec. 3.2.2) (Elowitz and Leibler 2000). The cooperativity, crucial for oscillations to occur, was modeled by including two binding sites for the repressor molecules on each promoter.

The extensive analysis and comparison between different realizations of stochastic and deterministic models of the repressilator was performed in (Loinger and Biham 2007). One of the results of that work is that the oscillations of the deterministic model persist, even without cooperative binding, in presence of degradation of the repressor when bound to the promoter. Additionally, the deterministic model does exhibit oscillations, while the stochastic model does not, provided cooperative binding, absence of the bound repressor degradation, and the explicit inclusion of mRNA species in the model (Loinger and Biham 2007).

The detailed model of the repressilator and the conditions for sustained

oscillations are provided for three known natural repressors in (Tuttle et al. 2005). It is worthwhile noting that the authors did not consider the effects of delays in the model (as in (Zhu et al. 2007)) which implies that it is expected that some of the conditions will differ in the real system.

On the other hand, the study conducted in (Zhu et al. 2007) of the repressilator model, using the mean-field deterministic approach with time delays, shows that increasing the protein delay decreases the frequency of oscillations. Moreover, for the highest considered delay for proteins, two modes of oscillations have been observed: high-frequency small-amplitude oscillations and low-frequency high-amplitude oscillations. The former oscillations are transient and correspond to the regime of the circuit where either all proteins, or none of them, are synthesized. The low-frequency high-amplitude oscillations are robust and represent the rhythm corresponding to the true limit cycle of the system (Zhu et al. 2007).

Note that the repressilator model can be written according to the general delayed stochastic model of GRN (see Sec. 3.3.2) (Ribeiro et al. 2006). This model can also be supplemented with details further expanding the key processes of gene expression. One such approach uses the model of transcription at the nucleotide level in order to study effects of transcriptional pauses on the dynamics of gene expression and gene regulatory networks (Rajala et al. 2010; Ribeiro et al. 2009a).

In these works, the model of transcription at the single nucleotide level expands the transcription process, namely, it incorporates the promoter occupancy time, pausing, arrests, misincorporations and editing, pyrophosphorolysis, premature termination, and accounts for the range occupied by an RNA polymerase when on the DNA strand (Ribeiro et al. 2009a). The model further sets the probability of RNA polymerase to pause at each nucleotide of the gene and the duration of the paused state. The study of the dependence of the period of the repressilator on the pause rate and its duration shows that the period increases with increasing both the rate and the duration of the pause (Rajala et al. 2010). The pause rate and its duration do not affect mean protein levels and period robustness, which was assessed by the 3-tuple information-entropy of the time series of the three proteins. Noise was assessed by the coefficient of variation (CV) of the sum of three proteins at each time step. The summing was done in order to remove the effect of the periodic oscillations on CV.

Interestingly, an increase in noise level at the single gene level does not significantly affect the robustness of the circuit's periodicity. This is because the negative feedback connections between the genes via their proteins act as 'noise filters' (Rajala et al. 2010). This indicates that transcriptional pauses might play a key role in adapting to environmental changes by tuning the

period without affecting the robustness of the oscillations. Given that the rate and duration of the pause depend on the sequence, the pauses can be considered as an evolvable mechanism of adaptation. In general, these recent studies show that the sequence of the genes composing the repressilator cannot be neglected when planning for a specific dynamic behavior.

Chapter 4

Models

This chapter contains a detailed description of the models of gene expression and gene regulatory networks that were used in the studies included in the thesis. Namely, a broad overview of the theoretical and experimental assessments of the dynamics of these circuits is provided. We also describe how, from these data, the models evolved. Finally, we describe results of previous studies analyzing these models.

4.1 Coupled genetic repressilators

The repressilator (Elowitz and Leibler 2000) is one of the simplest genetic constructions that is capable of nonlinear dynamical behavior, as it exhibits temporal oscillations in protein levels. This circuit is thus an excellent model for studies of cellular cyclic processes, such as the cell cycle and the circadian oscillator (see Sec. 1), and as a starting point for the developing of biotechnological applications, for example, reliable synthetic genetic clocks. One natural step is to expand this design effort to implement the mechanism of inter-cell global communication (García-Ojalvo et al. 2004), as such mechanism would serve many purposes, both in better understanding of cellular populations as well as for industrial applications.

One of the expected outcomes of this communication is to globally enhance the oscillating response of the population, which is crucial for the technological applications where the signal from single cells might not be discernible by the technological set-up (García-Ojalvo et al. 2004). However, the coupling between cells is not enough to achieve the synchronization over the population due to either oscillators' actively resisting to the synchronization (Winfree 1967) or if the coupling is too small or nonexistent (Njus et al. 1981). This coupling has to be able to actively promote the synchronization

of the population in order to serve its purpose. In this context, since the cells are subject to chemical signals from neighboring cells, it is worth studying the effects of such chemical environment on the dynamics of the coupled oscillators (Ullner et al. 2007).

Other applications are possible. For example, as opposed to enhancing coupling, one could implement a repressive coupling, which would lead to the phase-repulsive influence (Balázsi et al. 2001; Han et al. 1995; Volkov and Stolyarov 1991) on the oscillator cells of the population. The phase-repulsive coupling is used to explain the mechanisms underlying, e.g., morphogenesis in Hydra regeneration and animal coat pattern formation (Meinhardt 1982), neural activity in the brain of songbirds (Laje and Mindlin 2002), the jamming avoidance response in electrical fish (Metzner 1993), and regulation in the respiratory systems (Glass and Mackey 1988).

Due to a wide range of possible applications, coupled oscillators are of fundamental interest in various contexts (Collins and Stewart 1993; Mondragón-Palomino et al. 2011; Strogatz et al. 1992; Winfree 1967). In accordance with these aims, two principal schemes of the repressilator with quorum sensing coupling were proposed (García-Ojalvo et al. 2004; Ullner et al. 2007). Figure 4.1 shows the original repressilator circuit (compare with Fig. 3.5) and the coupling module implemented as the LuxIR-type quorum sensing network (see Sec. 2.3).

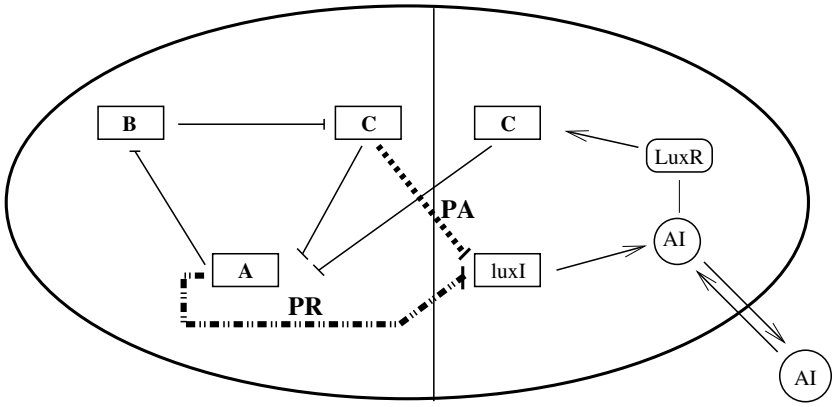


Figure 4.1: Two schemes of the repressilator (left module) with quorum sensing (right module). Negative relations (repression) are denoted as blunt end arrows. Activation and synthesis are denoted as pointed end arrows. Each of the schemes implies either PA or PR connections, i.e. phase-attractive or phase-repulsive coupling, respectively. AI stands for the autoinducer, the signal molecule of quorum sensing.

The three elements of the repressilators are denoted as A,B, and C in Fig. 4.1 (left module). The coupling module is depicted on the right in Fig. 4.1: a gene of the enzyme synthesizing the autoinducer (AI) is denoted as *luxI*, the receptor of AI is LuxR. The signal AI freely diffuses through the cell membrane and binds in the cell to its receptor LuxR. The complex of LuxR and AI activates the target gene. In all schemes of the repressilator with QS the target gene is C. The choice of the target gene does not influence subsequent results due to the symmetry of the repressilator circuit.

In order to achieve the synchronization over the population of the oscillators, *phase-attractive* coupling was introduced by placing the *luxI* synthase gene under the inhibitory control of the target element C (PA in Fig. 4.1). Thus, when the signal AI reaches the cell, the AI-LuxR complex activates the target gene whose protein, in turn, forms a negative feedback loop, in addition to the overall negative feedback of the repressilator's core. At the population level, this means that any deviation from the "average" behavior of any cell is suppressed by the coupling.

The *phase-repulsive* coupling is introduced by placing *luxI* under the inhibitory control of the gene that is repressed by the target gene (PR in Fig. 4.1). The incoming pulse of AI increases the synthesis of the target gene (C), whose product, being the element of the repressilator's core, serves its function by repressing the corresponding gene (A). The latter (A), in turn, represses the synthesis of *luxI*, providing the phase-repulsive coupling. These two negative feedbacks provide the overall positive feedback, in addition to the overall negative feedback of the repressilator's core. The response of the cell to the pulse of incoming AI is, thus, in the form of an increase of AI's synthesis. Note that in the natural examples of the quorum sensing mechanism of the LuxIR-type, the signal molecule induces its own synthesis, sustaining the phase-repulsive coupling (see Sec. 2.3) (Waters and Bassler 2005).

The phase space of the system, consisting of many oscillators, can be reduced to a smaller space, constituted with the variables of a single oscillator, given the oscillators are identical. In such reduced phase space, every state of an oscillator at any particular moment of time is represented by a point. The set of these representative points for an oscillator forms, during an evolution of the system, a closed trajectory in the space, known as limit cycle. In case of homogeneous dynamical regimes all limit cycles of the system coincide in the reduced phase space.

During an evolution of the system with phase-attractive coupling, the representative points of each oscillator are attracted to each other, while being on the limit cycle in the reduced phase space; hence the name of this form of coupling. The difference of the oscillator phases tends to a zero in the case of the phase-attractive coupling. Oppositely, during an evolution of

the system with phase-repulsive coupling, the representative points of each oscillator repulse each other, while being on the limit cycle in the reduced phase space; hence the name of this type of coupling. In this case, the difference of the oscillator phases tends to a finite limit.

Both phase-attractive and phase-repulsive systems possess the same equations governing mRNA and protein dynamics. The mRNA dynamics is described by the following Hill-type kinetics, with Hill coefficient n:

$$\begin{aligned} \frac{da_i}{dt} &= -a_i + \frac{\alpha}{1 + C_i^n}\,, \\ \frac{db_i}{dt} &= -b_i + \frac{\alpha}{1 + A_i^n}\,, \\ \frac{dc_i}{dt} &= -c_i + \frac{\alpha}{1 + B_i^n} + \kappa\frac{S_i}{1 + S_i}\,, \end{aligned} \tag{4.1}$$

where i specifies the index of the cell in the population; a_i, b_i, and c_i denote the concentrations of mRNA and A_i, B_i, and C_i are the corresponding protein concentrations. S_i is the concentration of the autoinducer inside the cell. The model is made dimensionless by measuring time in units of mRNA life time (assumed identical for all genes), and mRNA and protein levels in units of their Michaelis constants. The mRNA concentrations are additionally rescaled by the ratio of their corresponding proteins' degradation and translation rates. α is the dimensionless transcription rate in the absence of a repressor. κ is the maximum transcription rate of the LuxR promoter. The term with κ describes the activation of the target gene (C) by AI. LuxR is assumed to be constantly produced and sufficient for AI to bind and form the activation complex. Thus, only AI concentration is explicitly present in the model.

The dynamics of the proteins is described by the following equations:

$$\begin{aligned} \frac{dA_i}{dt} &= \beta_a(a_i - A_i)\,, \\ \frac{dB_i}{dt} &= \beta_b(b_i - B_i)\,, \\ \frac{dC_i}{dt} &= \beta_c(c_i - C_i)\,, \end{aligned} \tag{4.2}$$

where $\beta_{\{a,b,c\}}$ is the ratio between mRNA and protein life times (inverse degradation rates). Note that the protein equations are linear.

The AI concentration S_i in cell i is proportional to the concentration of LuxI (the protein synthesizing AI). The AI concentration S_i is scaled by its Michaelis constant. The dynamics of AI is also affected by the intracellular degradation and diffusion toward and from the intercellular space:

$$\frac{dS_i}{dt} = -k_{s0}S_i + k_{s1}[\text{LuxI}] - \eta(S_i - S_e)\,, \tag{4.3}$$

$$S_e = Q\bar{S}\,, \tag{4.4}$$

$$\bar{S} = \frac{1}{N}\sum_{i=1}^{N} S_i\,.$$

The diffusion coefficient η depends on the permeability of the membrane to the autoinducer, the membrane surface area and the cell volume. The extracellular AI concentration S_e can be replaced with the mean field $\bar{S}$ of intracellular AI concentrations S_i due to the fact that the diffusion of S_e is faster than the period of the repressilator and, thus, the quasi-steady-state approximation can be applied (Dockery and Keener 2001) (eq. (4.3)).

The parameter Q is defined as:

$$Q = \frac{\delta N/V_{ext}}{k_{se} + \delta N/V_{ext}}\,, \tag{4.5}$$

where N is the number of cells, V_{ext} is the total extracellular volume, k_{se} is the extracellular AI degradation rate, and δ is the product of the membrane permeability and the surface area (García-Ojalvo et al. 2004; Ullner et al. 2008). The coupling coefficient Q is proportional to the cell density (N/V_{ext}) of the population and can be varied in the range between 0 and 1. This can be achieved by, for example, changing the chemostat volume in the experiment (García-Ojalvo et al. 2004; Ullner et al. 2008).

Equations (4.1-4.2) are identical for both phase-attractive and phase-repulsive coupling type models. The difference between them is in the equation describing the dynamics of the autoinducer (eq. (4.3)). Assuming that the life times of the proteins of the schemes are identical, one can express the concentration of LuxI protein ([LuxI] in eq. (4.3)) as one of the repressilator's proteins, A, B, or C. Thus, in the case of the phase-attractive coupling scheme, the LuxI follows the changes in the concentration of protein A since both A and LuxI are repressed by the same protein C.

$$[\text{LuxI}] \equiv A\,. \tag{4.6}$$

Similarly, in the phase-repulsive scheme:

$$[\text{LuxI}] \equiv B\,. \tag{4.7}$$

4.1.1 Phase-attractive coupling

The main results of the population of repressilators with phase-attractive coupling were obtained in (García-Ojalvo et al. 2004). As this type of coupling is intended to enhance the oscillatory signal, the synchronization and coherence properties were investigated in a systematic way.

It was shown that the population of 10^4 oscillatory cells demonstrates the perfect phase locking as the coupling Q increases from 0 (uncoupled oscillators) to 1 (maximal coupling strength) (García-Ojalvo et al. 2004). Also, the study presents the effect of mRNA-to-protein lifetime ratio $\beta \equiv \beta_a = \beta_b = \beta_c$ variation, since β has the strongest effect on the period of the oscillators. β was set to follow a Gaussian distribution with different standard deviations $\Delta\beta$ to model to what extent the repressilator cells in the population are similar to each other. One of the results of the study is that the more similar the individual repressilators are, the smaller is the threshold of the coupling strength for synchronization to occur. Furthermore, maximal coherence in the synchronized regime was found to decrease as the differences between the oscillators increase (García-Ojalvo et al. 2004).

Thus, in (García-Ojalvo et al. 2004) the authors showed the possibility to observe the macroscopic rhythms in a population of synthetic genetic oscillators under realistic experimental conditions. This is a significant development, given that the earlier experimental studies assumed that a large degree of variability would prevent this (Elowitz and Leibler 2000). The degree of synchronization reported sufficed to generate global rhythms in a highly heterogeneous population of genetic oscillators. The resulting clock-like behavior is robust even to the stochastic drifts of the individual oscillators (García-Ojalvo et al. 2004). Finally, it was shown that quorum sensing coupling is able to synchronize the population of genetic oscillators. Similar conclusions were drawn in the experimental set up later on (Danino et al. 2010), but on another genetic oscillator with quorum sensing coupling.

However, the phase-attractive system of coupled genetic oscillators was not shown to have any other global rhythms, except for the so called in-phase oscillations leading to synchronous dynamical behavior of the ensemble of oscillators. The principal design of the phase-attractive coupling was believed to have only the in-phase synchronous dynamical solutions, which lead to the enhancement of the dynamical rhythm of a single oscillator on the population level (since all oscillators possess the same rhythm with locked phase relative to each other).

In this thesis, we hypothesized the existence of other types of synchronization in the phase-attractive coupling system. We use the bifurcation analysis (see Sec. 3.2) to reveal the dynamical regimes of the system of two

coupled genetic oscillators of this type. The bifurcation analysis of two coupled repressilators can be applied directly to the population of the oscillators, since the type of the synchronization does not depend on the size of ensemble (Ullner et al. 2008).

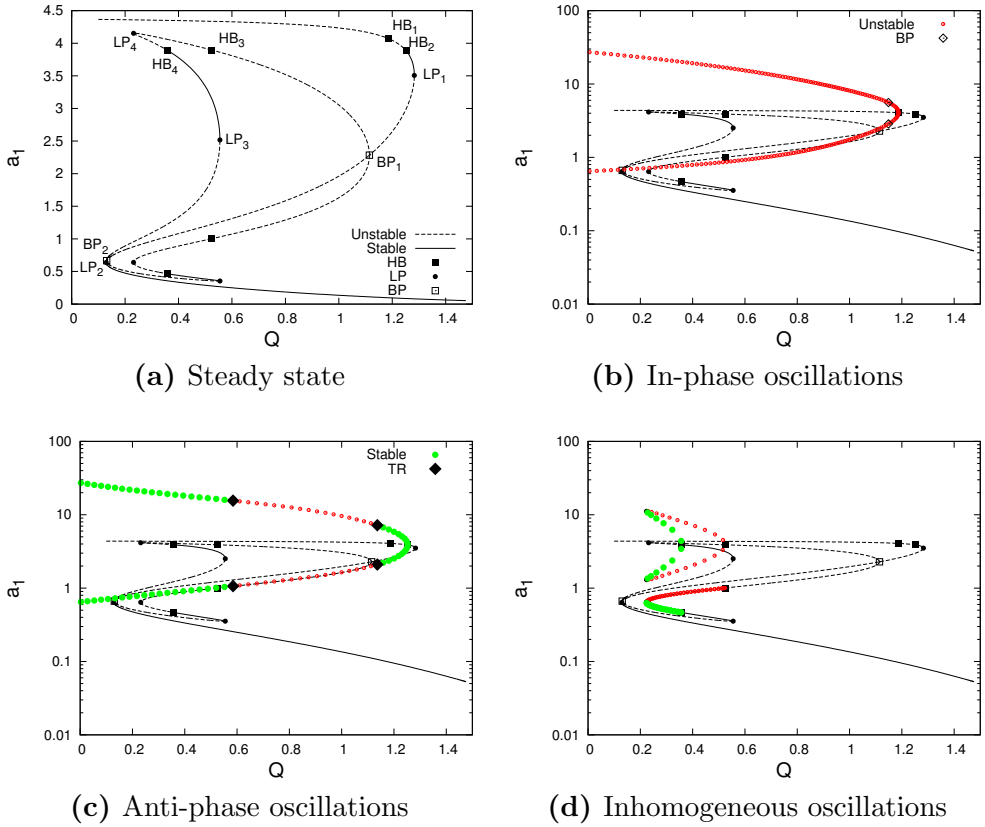
4.1.2 Phase-repulsive coupling

The system of two ($N = 2$) repressilators with phase-repulsive coupling was studied by means of the bifurcation analysis in (Ullner et al. 2008). The results are shown in Fig. 4.2. The figure shows the different dynamical solutions of the same system. Note the y-axis logarithmic scale of the panels (b), (c), and (d) of the figure. Note also that the coupling strength Q is varied within the interval $[0\,,1]$, values outside this interval do not have any biological meaning. The notations in Fig. 4.2 are the following: HB, Hopf bifurcation, LP, limit point bifurcation, BP, pitchfork (symmetry breaking) bifurcation, TR, invariant torus emergence bifurcation. The oscillatory (limit cycle) solutions are represented by the circles: solid and empty circles denote stable and unstable solutions, respectively. Lines denote steady state solutions: solid and dashed lines represent stable and unstable solutions, respectively.

The main continuation over the coupling strength parameter Q is shown in Fig. 4.2(a). This curve is characterized by the presence of the symmetry breaking bifurcations (BP) that lead to inhomogeneous solutions and by the stabilization of the homogeneous steady state solution, HSS (after LP_2 for $Q > 0.129$ in Fig. 4.2). The HSS solution is characterized by the constant levels of mRNA and protein concentrations. Another HSS is found between HB_2 and LP_1, but it is located outside of the biologically relevant interval, i.e. $Q > 1$.

The symmetry breaking bifurcations make the unstable steady states split into two additional solution branches. The latter give rise to the stable inhomogeneous steady states (IHSS). The IHSS appears as two distinct concentration levels. This regime becomes stable via the Hopf bifurcation HB_4. The IHSS coexist with HSS solution in the parameter space (Fig. 4.2(a)).

The Hopf bifurcation HB_4 gives rise to a branch of stable inhomogeneous periodic solutions, known as inhomogeneous limit cycles (IHLC) (Tyson and Kauffman 1975). In this system, the IHLC is characterized by the small amplitude oscillations of one repressilator cell and large amplitude oscillations of the other. Small amplitude oscillations are not discernible from the constant steady state level, while the large amplitude oscillations have amplitude slightly smaller than that of the isolated repressilator. The IHLC is stable for values of Q between HB_4 and LP_4 (Fig. 4.2(d)).



(a) Steady state **(b)** In-phase oscillations

(c) Anti-phase oscillations **(d)** Inhomogeneous oscillations

Figure 4.2: The dynamical regimes of the system of two repressilators coupled with phase-repulsive coupling.

The inhomogeneous solutions, IHSS and IHLC, can be occupied by each of the two oscillators with the same probability due to the symmetry of the system. The initial conditions determine the separation of the oscillators. These regimes are not characteristic to this model and were observed previously (Kuznetsov et al. 2004; Tyson and Kauffman 1975; Volkov and Romanov 1995).

For the coupling values smaller than that of LP_2 the system is described by the self-oscillatory solution. For the system of two repressilators this solution corresponds to anti-phase oscillations (Fig. 4.2(c)). This branch arises at HB_2 and loses stability at the bifurcation of the invariant torus emergence, TR. For smaller values of Q it regains the stability through the other TR bifurcation. The numerical analysis reveals the complex chaotic behavior within the region where the anti-phase limit cycle is unstable (Ullner et al. 2008).

In contrast to the phase-attractive coupling system (García-Ojalvo et al. 2004) (see Sec. 4.1.1), where the coupling was implemented to provide the coherence enhancement in the population, the system with phase-repulsive coupling does not show the existence of stable in-phase regime (synchronous oscillation over the entire population). The branch of the in-phase limit cycle emerges at HB_1 and is unstable in the whole range of the coupling strengths (Fig. 4.2(b)). The existence of anti-phase (or phase-shifted) oscillations is a manifestation of the phase-repulsive character of the autoinducer-mediated coupling (Ullner et al. 2008).

The system of two identical repressilators coupled through the phase-repulsive coupling implemented as a diffusion of the autoinducer demonstrates the big variety of the dynamical behaviors. This ranges from both homogeneous and inhomogeneous stationary solutions to homogeneous, where the phase shift is maximized (anti-phase), and inhomogeneous oscillatory solutions and even chaos, the regime whose dynamics is characterized by the unpredictable behavior though the system is governed by deterministic laws.

4.2 Model of gene expression at the nucleotide and codon levels

Transcription and translation are multi-step processes. Events occurring during the transcriptional and translational elongation are stochastic due to the probabilistic competition of the elementary processes. Thus, there are pathways leading to different states of the RNA polymerase, when it processes each nucleotide (Ribeiro 2010; Ribeiro et al. 2009a; Roussel and Zhu 2006), and of the ribosome, when it processes each codon (Mäkelä et al. 2011). For example, there are transcriptional pauses of the RNA polymerase (Greive and von Hippel 2005; Herbert et al. 2006) and trans-translation (Moore and Sauer 2005) resulting in release of all translating ribosomes from mRNA template (see p. 44).

All these events, introduce noise in each of the elongation processes, which can affect the production kinetics of transcripts and proteins. Several studies suggest that these noise sources may play a key role in bacterial adaptability in unpredictable or fluctuating conditions of the surrounding environment (Acar et al. 2008; Lee et al. 2010). To account for this in the models, one needs to make use of single nucleotide level and single codon level dynamical models of gene expression (Mäkelä et al. 2011; Ribeiro 2010).

In prokaryotes this is of particular importance, since transcription and translation are coupled, i.e. translation elongation by a ribosome can start

as soon as the initial section of mRNA, known as ribosome binding site, is formed during transcription (see Sec. 2.1). This coupling implies that, at least to some extent, fluctuations in RNA levels will propagate to protein levels. Moreover, it is largely unknown what mechanisms the bacterial cell utilizes to control the level of fluctuations. If these exist, they must act at the sequence level further enhancing the need for these models.

Recently, a model of transcription and translation at the nucleotide and codon levels was proposed in (Mäkelä et al. 2011). The model of transcription is similar to that in (Ribeiro et al. 2009a).

Tables 4.1 and 4.2 contain reactions modeling transcription and translation. Every reaction described has a reference where the kinetic parameters for the reaction were taken from. Most of the rates and delays are obtained from measurements in *Escherichia coli*, mainly for *lacZ* gene. Delays and life times of some states (e.g. arrested state) are denoted as τ_x, and if a product is delayed it is represented as $X(\tau_x)$ (see Sec. 3.3.2). Stochastic rate constants are denoted as k_x. Note that the rates and delays can vary for different genes, and only averages of reported values are shown.

The model of transcription at the nucleotide level includes transcription initiation (reaction 1 in Tab. 4.1) (Zhu et al. 2007), during which the promoter open complex is formed with delay τ_{oc}, followed by promoter clearance (reaction 2) (McClure 1980) that consumes first $\Delta_\mathrm{P} + 1$ unoccupied nucleotides ($\mathrm{U}_{[1,(\Delta_\mathrm{P}+1)]}$) of the gene, which are needed for the Rp to bind. The initiation and clearance of the promoter are followed by the elongation, which consists of nucleotide activation (reaction 3) (Proshkin et al. 2010) and step-wise elongation (reaction 4) (Proshkin et al. 2010), which requires the nucleotide next to the Rp to be unoccupied ($\mathrm{U}_{n+\Delta_\mathrm{P}+1}$), accounting for the Rp coverage size, to correctly model polymerase traffic. At the end of the elongation a complete RNA molecule is formed and the Rp is released via the reaction 12 (Tab. 4.1) (Greive et al. 2008).

Several events during transcription compete with the elongation at each nucleotide, such as transcriptional pauses (reaction 5 in Tab. 4.1) (Greive and von Hippel 2005). The pause state O_{n_p} of Rp can release spontaneously (reaction 5) or by collisions with preceding activated state of Rp, $\mathrm{A}_{n-2\Delta_\mathrm{P}-1}$ (reaction 6). Such collisions can also induce the pause state of Rp (reaction 7). There are ubiquitous arrests and release from arrests (reaction 8) (Ribeiro et al. 2009a), misincorporation and editing (reaction 9) (Greive and von Hippel 2005), premature terminations (reaction 10) (Lewin 2008) that release Rp and nucleotides it occupied ($\mathrm{U}_{[(n-\Delta_\mathrm{P}),(n+\Delta_\mathrm{P})]}$). Pyrophosphorolysis removes inserted ribonucleotide ($\mathrm{U}^\mathrm{R}_{n-\Delta_\mathrm{P}-1}$) of the growing mRNA strand via reaction 11 (Erie et al. 1993). The model accounts for the nucleotides occupied by Rp on the DNA strand. Finally, two RNA polymerases can never

Table 4.1: Chemical reactions, rate constants (k_x, in s^{-1}), and delays (τ_x, in s) used to model transcription initiation, elongation and termination, as well as mRNA degradation. Pro — promoter, Rp — RNA polymerase, O — nucleotide occupied by Rp; U — unoccupied nucleotide; A — activated nucleotide; U^R — unoccupied ribonucleotide. n denotes the index number of a nucleotide in the sequence. $(2 \cdot \Delta_P + 1)$ — range of nucleotides that Rp occupies, $\Delta_P = 12$. Reactions are explained in the text.

1.Initiation and promoter complex formation	$\mathrm{Pro} + \mathrm{Rp} \xrightarrow{k_{init}} \mathrm{Rp} \cdot \mathrm{Pro}(\tau_{oc})$	$k_{init} = 0.015$, $\tau_{oc} = 40 \pm 4$
2.Promoter clearance	$\mathrm{Rp} \cdot \mathrm{Pro} + \mathrm{U}_{[1,(\Delta_P+1)]} \xrightarrow{k_m} \mathrm{O}_1 + \mathrm{Pro}$	$k_m = 114$
3. Activation	$\mathrm{O}_n \xrightarrow{k_a} \mathrm{A}_n$	$k_a = 114, n > 10$; $k_a = 30, n \leq 10$
4. Elongation	$\mathrm{A}_n + \mathrm{U}_{n+\Delta_P+1} \xrightarrow{k_m} \mathrm{O}_{n+1} + \mathrm{U}_{n-\Delta_P} + \mathrm{U}^R_{n-\Delta_P}$	$k_m = 114$
5. Pausing	$\mathrm{O}_n \underset{1/\tau_p}{\overset{k_p}{\rightleftharpoons}} \mathrm{O}_{n_p}$	$k_p = 0.55, \tau_p = 3$
6. Pause release by collision	$\mathrm{O}_{n_p} + \mathrm{A}_{n-2\Delta_P-1} \xrightarrow{0.8k_m} \mathrm{O}_n + \mathrm{A}_{n-2\Delta_P-1}$	$k_m = 114$
7. Pause by collision	$\mathrm{O}_n + \mathrm{A}_{n-2\Delta_P-1} \xrightarrow{0.2k_m} \mathrm{O}_{n_p} + \mathrm{A}_{n-2\Delta_P-1}$	$k_m = 114$
8. Arrests	$\mathrm{O}_n \underset{1/\tau_{ar}}{\overset{k_{ar}}{\rightleftharpoons}} \mathrm{O}_{n_{ar}}$	$k_{ar} = 0.00028$, $\tau_{ar} = 100$
9. Editing	$\mathrm{O}_n \underset{1/\tau_{ed}}{\overset{k_{ed}}{\rightleftharpoons}} \mathrm{O}_{n_{corr}}$	$k_{ed} = 0.008$, $\tau_{ed} = 5$
10. Premature termination	$\mathrm{O}_n \xrightarrow{k_{pre}} \mathrm{Rp} + \mathrm{U}_{[(n-\Delta_P),(n+\Delta_P)]}$	$k_{pre} = 0.00019$
11. Pyrophosphorolysis	$\mathrm{O}_n + \mathrm{U}_{n-\Delta_P-1} + \mathrm{U}^R_{n-\Delta_P-1} \xrightarrow{k_{pyr}} \mathrm{O}_{n-1} + \mathrm{U}_{n+\Delta_P-1}$	$k_{pyr} = 0.75$
12. Completion	$\mathrm{A}_{last} \xrightarrow{k_f} \mathrm{Rp} + \mathrm{U}_{[\mathrm{last}-\Delta_P,\mathrm{last}]} + \mathrm{mRNA}$	$k_f = 2$
13. mRNA degradation	$\mathrm{mRNA} \xrightarrow{k_{dr}} \emptyset$	$k_{dr} = 0.011$

occupy simultaneously the same nucleotide. Reaction 13 models the RNA degradation (Yu et al. 2006).

The reactions modeling translation are shown in Tab. 4.2. The delayed stochastic model of translation at the codon level includes initiation (reaction 14) (Mitarai et al. 2008) that requires first $\Delta_{\mathrm{R}} + 1$ ribonucleotides to be free for the ribosome to bind. The elongation of translation consists of ribonucleotide activation (reaction 15) (Wen et al. 2008) followed by step-wise translocation by triplets (codons) of ribonucleotides (reactions 16–18) (Mitarai et al. 2008).

The reactions competing with translocation are ribosomal back-translocation (reaction 19) (Shoji et al. 2009), its drop-off (reaction 20) (Jørgensen and Kurland 1990), and trans-translation (reaction 21) (Moore and Sauer 2005). The trans-translation reaction results in degradation of the mRNA being translated and release of all ribosomes ($[\mathrm{Rib}^{\mathrm{R}}] \times \mathrm{Rib}$, where $[\mathrm{Rib}^{\mathrm{R}}]$ is the number of ribosomes Rib on the mRNA strand) participating in the translation elongation. The step-wise translocation ends with elongation completion (Mitarai et al. 2008) followed by protein folding (reaction 22) (Cormack et al. 1996), which is modeled by the delay τ_{fold} following the Gaussian distribution. The model accounts for the ribonucleotides occupied by a ribosome when on the RNA. Also accounted for are codon-specific translation rates (Sørensen and Pedersen 1991). Finally, proteins undergo degradation (reaction 23) (Andersen et al. 1998).

Trans-translation corresponds to the release of the ribosome from the RNA template after stalling, which can occur for several reasons, such as incorporation of an incorrect codon, premature mRNA degradation, or frameshifting (Keiler 2008; Moore and Sauer 2005). In the model, stalling followed by trans-translation can occur spontaneously with a given probability at any codon via reaction 21. When occurring, the mRNA strand is degraded and all translating ribosomes are released. Estimates from the observation of expression activity in *E. coli* suggest that, on average, 0.4% of translation reactions are terminated by trans-translation (Moore and Sauer 2005), meaning that the probability of occurrence of this event at each nucleotide depends on the length of the gene.

The study (Mäkelä et al. 2011) showed, within realistic parameter values (Tables 4.1 and 4.2), that the protein noise level to great extent is determined by the fluctuations in the RNA levels. Translation elongation was found to be less stochastic than transcription elongation. The degree of coupling between RNA levels and protein levels was found to be strongly dependent on translation initiation rate. Other study confirms the same result (Pedraza and Paulsson 2008). Translation initiation rate also determines how fast the RNA level changes propagate to the protein level. Additionally,

Table 4.2: Chemical reactions, rate constants (k_x, in s^{-1}), and delays (τ_x, in s) used to model translation initiation, elongation, and termination, as well as protein folding and degradation. Rib — ribosome; Rib^R — ribosome on mRNA strand; $[Rib^R]$ — number of translating ribosomes on mRNA strand; P — complete protein; O^R — ribonucleotide occupied by Rib; U^R — unoccupied ribonucleotide; A^R — activated ribonucleotide. n denotes the index number of a ribonucleotide in the sequence. $(2 \cdot \Delta_R + 1)$ — range of ribonucleotides that Rib occupies, $\Delta_R = 15$. See explanations in the text.

14. Initiation	$\text{Rib} + \text{U}^\text{R}_{[1,\Delta_\text{R}+1]} \xrightarrow{k_{tl}} \text{O}^\text{R}_1 + \text{Rib}^\text{R}$	$k_{tl} = 0.33$
15. Activation	$\text{O}^\text{R}_n \xrightarrow{k_{tr\{A,B,C\}}} \text{A}^\text{R}_n$	Codon dependent: $k_{trA} = 35$; $k_{trB} = 8$; $k_{trC} = 4.5$
16–18. Stepwise translocation	$\text{A}^\text{R}_{n-3}+\text{U}^\text{R}_{[n+\Delta_\text{R}-3,n+\Delta_\text{R}-1]} \xrightarrow{k_{tm}} \text{O}^\text{R}_{n-2}$ $\text{O}^\text{R}_{n-2} \xrightarrow{k_{tm}} \text{O}^\text{R}_{n-1}$ $\text{O}^\text{R}_{n-1} \xrightarrow{k_{tm}} \text{O}^\text{R}_n + \text{U}^\text{R}_{[n-\Delta_\text{R}-2,n-\Delta_\text{R}]}$	$k_{tm} = 1000$
19. Backtranslocation	$\text{O}^\text{R}_n + \text{U}^\text{R}_{[n-\Delta_\text{R}-2,n-\Delta_\text{R}]} \xrightarrow{k_{bt}} \text{A}^\text{R}_{n-3} + \text{U}^\text{R}_{[n+\Delta_\text{R}-3,n+\Delta_\text{R}-1]}$	$k_{bt} = 1.5$
20. Drop-off	$\text{O}^\text{R}_n \xrightarrow{k_{drop}} \text{Rib} + \text{U}^\text{R}_{[n-\Delta_\text{R},n+\Delta_\text{R}]}$	$k_{drop} = 1.14 \cdot 10^{-4}$
21. Trans-translation	$\text{mRNA} \xrightarrow{k_{tt}} [\text{Rib}^\text{R}] \times \text{Rib}$	Depends on the gene length
22. Elongation completion and protein folding	$\text{A}^\text{R}_{last} \xrightarrow{k_{tlf}} \text{Rib} + \text{U}^\text{R}_{[last-\Delta_\text{R},last]} + \text{P}(\tau_{fold})$	$k_{tlf} = 2$, $\tau_{fold} = 420 \pm 100$
23. Protein degradation	$\text{P} \xrightarrow{k_{dp}} \emptyset$	$k_{dp} = 0.0017$

decreasing of the coupling between transcription and translation, as determined by the rate of translation initiation, makes the protein levels become less noisy. Thus, this parameter can be crucial in defining the time of the cell response to environmental changes and fluctuations level in protein numbers (Mäkelä et al. 2011).

Chapter 5

Conclusions and Discussion

This thesis focused on the dynamical properties of models of bacterial gene regulatory networks, namely, genetic switches and clocks. The analysis was restricted to networks possessing these behaviors in the protein and RNA molecular species due to their relevance in the functioning of real organisms as well as their potential applications in synthetic networks (Atkinson et al. 2003; Elowitz and Leibler 2000; Gardner et al. 2000; Jacob and Monod 1961; Tigges et al. 2009). In nature, they appear in the form of simple motifs, of two to three genes, but are integrated in larger scale genetic networks, playing a key role in regulating their intricately complex dynamics. Further, some of these circuits are influenced by the behavior of identical circuits in neighbor cells.

The work presented focused on the generation of novel dynamical rhythms in cell populations containing genetic clocks and on the influence of the network topology of the clock on the dynamics of the molecular species of the circuit. Additionally, several sequence dependent events were considered, particularly, their impact on the temporal dynamics of mRNA and protein numbers of both switching and oscillatory circuits.

Briefly, in Publication-I and Publication-IV, the dynamics of genetic clocks composed of three genes with a quorum sensing coupling module was studied. Publication-II, Publication-III and Publication-V studies primarily concentrated on the investigation of the dynamics of single gene expression and of two-gene toggle switches by means of highly detailed, sequence-level stochastic models with time delays.

In detail, in Publication-I, a system of two, coupled repressilators with phase-attractive coupling (García-Ojalvo et al. 2004) was studied and shown to have an anti-phase synchronization of the oscillatory solution. That is, the coupling of the oscillators of two cells causes them to behave antagonistically, which, in turn, in large cell populations, leads to the formation of clusters

of antagonistic cells. At this stage, from the initial conditions, the number and size of these clusters cannot be predicted beforehand and need to be determined separately.

Although the anti-phase synchronization was found to be characteristic to the system with phase-repulsive coupling (Ullner et al. 2007, 2008), this behavior can also be seen in the systems with phase-attractive coupling type, provided there is strong cooperativity in the repression mechanism of the network. Therefore, it was possible to conclude that the coupling scheme does not determine the type of synchronization.

Moreover, in the phase-attractive system, the in-phase synchronization persists when the anti-phase solution emerges, causing the bi-rhythmicity in the system, i.e., the coexistence of two solutions. This coexistence is not observed in the phase-repulsive system, where the in-phase synchronization was found to be unstable (Ullner et al. 2008). In this sense, the emergence of bi-rhythmicity between in- and anti-phase synchronization types as well as its dynamical properties (e.g. the bifurcations leading to the bi-rhythmicity) were found to be more informative of the population dynamics than the network structure.

The anti-phase solution, a result of highly nonuniform motion speed along the limit cycle, is stable due to the steepness of the nonlinear function determining cooperativity and/or in the regime of strong transcription rates (n and α in eq. (4.1)). Aside from these, the synchronization properties also depend on the other parameters of the system. These influence the system by shifting the dynamical regimes, while not affecting the emergence of bi-rhythmicity, which supports the idea of its determinant role in the global behavior. Interestingly, the quorum sensing coupling parameter Q (see (4.3–4.5)) was found to not influence the system significantly, meaning that the results are valid for a wide range of population densities.

Interestingly, it was in the regime of realistic mRNA lifetimes (i.e. shorter than that of proteins) that the system exhibited a whole cascade of anti-phase complex dynamical regimes, which supersede each other as the transcription rate is increased. These regimes are the inhomogeneous anti-phase oscillations, which differs from the homogeneous regime by having non-equal limit cycle amplitudes in two cells, the inhomogeneous solution with additional frequency, which appears when the torus-like attractor emerges, and, finally, the homogeneous solution with additional frequency of oscillations. Thus, the system of two oscillator cells possesses a variety of dynamical behaviors that can, under certain conditions, coexist.

In Publication-IV, the effects of the phase-repulsive coupling module (Ullner et al. 2007) on the dynamics of single oscillator cell were studied. It was shown that this coupling changes drastically the relatively simple behavior of

the original repressilator model-cell. Namely, the system acquires a new stable steady state that is not related to the emergence of the limit cycle via the Hopf bifurcation. The appearance of the additional attractor in the system leads to a coexistence of the stationary and oscillating behaviors, leading to a hysteresis. Moreover, the system is characterized by the existence of windows in the parameter space of chaotic behaviors with repeated cascades of bifurcations.

The hysteresis provides the system with new possibilities regarding its regulation and, thus, deserved special attention. We found that, in this regime, the cell obtains the ability to choose between two types of response: the clock-like periodic or the resting stationary. This bears resemblance to the bi-stability common to toggle switches, but in this case, with one of the two states being periodic in time. The hysteresis is made possible by the nature of the phase-repulsive coupling, namely, the competition between overall negative feedback of the core of the repressilator and the additional positive feedback caused by the autoinducer mediated coupling module. The latter operates against the oscillatory functioning of the repressilator's core and, for large enough autoinducer mediated transcription activation, the oscillations cease, i.e. the limit cycle is arrested as the system falls into the steady state.

For the elevated autoinducer mediated transcription activation, in addition to the hysteresis, the system of a single cellular oscillator exhibits more complex chaotic dynamics. This is accompanied with the unusual continuation of the limit cycle in the parameter space: two period doubling bifurcations (PD) are followed by two limit point bifurcations (LP), and this four-bifurcations structure is repeated many times. Each first PD in the structure gives rise to a solution branch with the whole cascade of PDs leading to chaos (Feigenbaum scenario) (Feigenbaum 1978), while the second PD of the structure restores the regular limit cycle. Thus, the regions between two consecutive period doubling bifurcations contain chaos. As a result, the system acquires regular limit cycles, resulting from the PDs (with doubled period after each PD), and chaos, when the limit cycle becomes eventually complex enough after the cascade of PDs. Moreover, the continuation branch of the solutions turns in the parameter space at LP points, overlapping the regimes. Thus, the coexistence between regular limit cycles and chaos, and even between two chaotic regimes is found to be possible. In conclusion, this study revealed that the relatively simple repressilator in its original form (i.e. which exhibited limit cycle oscillations alone) becomes intricately complex when intercellular communication, based on diffusion of small molecules, is added to the system.

From Publication-I and Publication-IV, it is possible to draw the following

general conclusions. The analysis of the dynamics of the model of a three-gene oscillator with the communication apparatus demonstrated the presence of a multitude of stable dynamical regimes and their coexistence, under specific conditions. This allowed for multi-stability and multi-rhythmicity, which are observable in living organisms and reflect the ability of living organisms to adapt in continuously changing environment. In this sense, these results shed some light on the underlying regulation mechanisms of bacteria and similar organisms.

However, gene regulatory networks are subject to natural noise arising from the small numbers of molecules constituting the network, for example, genes, RNA and RNA polymerases (McAdams and Arkin 1997; Ozbudak et al. 2002; Swain et al. 2002). Thus, it is important to determine to what extent this noise affects the dynamics of coupled genetic networks. For instance, how the period and amplitude of the oscillations are maintained by such inherently noisy mechanisms. We expect these questions to form the basis of subsequent studies in this topic.

These studies should be able to answer, for example, which of the findings are still 'valid' in the 'real systems'. This is not easily predictable. For example, it was recently shown for a bi-stable system based on the quorum sensing that the differences between "on" and "off" states are obliterated in presence of molecular noise (Goryachev 2011). One of the reasons for this is that the attractors of the deterministic system are in close vicinity in the phase space and thus, even a weak noise source removes the demarcation between the attractors. One way to avoid this is to introduce additional regulatory feedback mechanisms to increase the robustness of the attractors (Orrell and Bolouri 2004).

In Publication-II, the study of the dynamics of the two-gene toggle switch was conducted by means of simulations of a stochastic model of gene expression at the nucleotide and codon levels. Parameter values for the model were taken from measurements in *Escherichia coli* . It was shown that the coupling strength between the two genes of the switch, determined by the binding and unbinding rates of the repressors, affects the stability of the noisy attractors of the switch. The stability, in this context, is defined as the mean time that the switch is found in either of two regions of state space. One region is such that one protein exists in higher numbers than the other, and in the other region, the opposite occurs.

The stability of these noisy attractors was found to decrease with decreasing coupling strength, as the frequency of switching between the states increases. Within the bi-stability region, the temporal pairwise mutual information between the numbers of the two proteins was found to have a

maximum point, when the coupling strength is varied. This indicates that there exists a degree of coupling that is optimal for information propagation between the two genes. This property can be of importance given that the genetic switch was hypothesized to be used as a memory unit (Gardner et al. 2000).

The gene length of the component genes of the switch impacts on the number of RNA polymerases expected to be bound to the DNA template of the genes, at any given time. Consequently, we observed that the stability of the noisy attractors increased when the gene length is increased. These RNA polymerases act as a 'memory mechanism' of past states of the system and as the RNA and proteins are produced from their activity, they tend to reinforce these past states. This effect is strong enough to break the symmetry of the two states of the toggle switch. Namely, by setting one of the genes to have a longer length, it was found to be possible to bias the probability of occurrence of the two noisy attractors.

The study of the role of sequence dependent mechanisms in the dynamics of the genetic toggle switch was continued in Publication-III. In this publication, it was shown that events during transcription and translation, that cannot be captured by the simpler models, affect significantly the dynamics of the genetic switch, namely, the stability of its noisy attractors, the main macroscopic property of this network. The parameters defining the properties of each gene and those associated with the interaction between two genes were found to have different effects on the network's dynamics. The mutual pairwise information is maximized when varying parameters determining the coupling between the two genes. However, the same effect was not observed for a simpler model, without sequence-level transcription and translation elongation. Additionally, this effect cannot be attained when varying the parameters determining the coupling between transcription and translation within each gene.

The effect of transcriptional pauses was also considered. In this regard, it was shown that such pauses, as they enhance traffic and collisions between RNA polymerases and ribosomes, reduce the stability of the noisy attractors of the switch. Overall, the results indicate that sequence dependent events, captured by the detailed model of the genetic toggle switch, cannot be neglected when studying dynamical behaviors and mechanisms of regulation of this system.

Following this study, subsequently, Publication-V focuses on the effects of codon sequence on the expression of single genes and on the dynamics of small genetic networks, namely, switches and clocks. To execute this study, we implemented several models, differing in the codon sequence. These variations were applied randomly, according to the frequency of occurrence of each

codon in *E. coli* in 1000-nucleotide gene. It was observed that such variations suffice to cause changes in mean protein numbers near equilibrium, due to the mean codon translation efficiencies along the sequence. Additionally, not all regions of these sequences play the same role in determining the mean protein levels, as the ribosome binding site (RBS), which regulates the frequency of ribosome bindings to the RNA, was found to be most determinant.

These results were shown to be in close agreement with measurements of protein expression levels as a function of the codon sequence. In these experiments, the various genes contained the same promoter region and coded for the same protein, but differed in the usage of synonymous codons (Welch et al. 2009). We modeled these genes, simulated the kinetics of their expression and computed the correlation between the mean protein expression levels and those reported in (Welch et al. 2009). It was found a strong positive correlation for the subset of genes with slower average codon translation rates at the start of the sequence.

Also in this publication, the specific codon translation efficiency profiles along the sequence were examined. The short codon sequences at the start the gene with linearly increasing translation efficiencies, known as slow ramps (Tuller et al. 2010), were shown to affect the mean, but not the fluctuations, in protein numbers, by affecting the rate of translation initiation. The ramps with linearly decreasing translation efficiencies at the end of the gene do not have the same effect.

Finally, the slow ramps were shown to affect the dynamics of genetic switches and clocks. The start ramps cause a significant decrease in the stability of the two-gene toggle switch. The end ramps were not found to cause any effect, due to their weak effect on the mean protein numbers. These results indicate that the dynamics of genetic networks can be affected by ramps of slow codons and that the effect depends on the location of the ramp relative to the RBS. These effects can be enhanced if these ramps are not placed symmetrically on the two genes sequences. If only one of the genes has a start ramp, the behavior of the network is significantly biased, resulting in a significant increase in stability.

Finally, the Fast Fourier Transform was used to assess the robustness of the oscillations in the protein numbers of the repressilator model with slow ramps. Since the main frequency of the oscillations corresponds to the main peak of the Fourier Transform, the robustness can be assessed by the height of the peak. The higher the peak, the more times is the oscillation duration equal to the given period. Relevantly, the more robust is the oscillator the more predictable is its behavior, i.e. with the period close to the one corresponding to the main peak of the transform. Interestingly, three differing models of the repressilator (with random sequences, start and end slow ramps

in each gene, respectively) exhibited the same frequency of oscillations, when assessed by the Fourier Transform. However, the robustness of the dynamics of these three models differed. Namely, the model with the slow ramp at the start has the strongest decrease in the transform peak's height, indicating the lowest robustness for the model. The effect of adding the slow ramp to the end of the three genes is also tangible, as opposed to the cases of the genetic toggle switch and of the model of a single gene considered above.

In conclusion, from these studies of the effects of sequences on the dynamics of gene expression and small genetic circuits, we find that the delayed stochastic models of bacterial gene expression at the nucleotide and codon levels are able to accurately capture measurements of protein mean levels, and also the networks' noisy attractors, suggesting that these models can be utilized in the future in combination with experimental data to better understand the behavior of these systems *in vivo*.

Further elaboration of these models will improve the quality of the modeling. For instance, by incorporating additionally regulatory mechanisms into the stochastic models, such as those acting following translation, or prior to transcription elongation. One example of the former mechanisms is phosphorylation, which was found to be a non-negligible mechanism for controlling the fine-tuning of the transcription factor activity (Ashcroft et al. 1999; Holmberg et al. 2002). Additionally, present models do not account for the effects that ribosomes may have on the kinetics of transcription elongation. Recent studies (Burmann et al. 2010) suggest that the ribosome's translation efficiency affects the speed of transcription elongation, since the ribosome bound to the growing mRNA may interact with the RNA polymerase transcribing the corresponding gene region. However, some indirect effects of the interaction between ribosomes and RNA polymerases are already present in the model at the nucleotide and codon levels (Mäkelä et al. 2011). Namely, the stalled polymerase, due to pause, arrest or misincorporation, causes the ribosome preceding it to pause. Also, the ribosome close to the transcribing polymerase prevents the transcriptional backtracking, because the RNA is not available for this event to be possible.

Along these lines, it is worthwhile to mention that these studies rely on models whose kinetics is driven by the SSA algorithm (Gillespie 1976, 1977), which assumes well-stirred systems with chemically interacting molecules and includes all possible random movements of the reacting species. However, for example, transcription and translation processes, carried out on the huge polymers and by large protein complexes and numerous co-factors that exist in small numbers in the cells (Alberts et al. 2002; Lewin 2008), may require another algorithm accounting for more realistic physical processes and, thus,

resulting in more exact modeling. In particular, one of the possible future directions for improving models of gene expression is likely the accounting for spatial features, such as molecular dimensions, among others. For example, the RNA polymerase moving along the DNA template cannot be modeled as a three dimensional process during which the RNA polymerase finds an appropriated nucleotide to collide, governed by the simple mechanistic laws of molecular motion. This process rather resembles one dimensional movement of the polymerase on the DNA strand with preferable sliding direction (the enzyme mainly moves from the transcription start site to the end of a gene). The fluctuations still exist in this representation, although they are formed by different causes. For instance, DNA template thermal vibrations and stochasticity of physical interactions between the enzyme and the template might cause the nonuniform movements of the polymerase during the transcription elongation. We believe the modeling relying on the biologically plausible physical interactions of the biomolecules would result in the noise level conformable with that in the real systems and lead to a better understanding of how large complexes of interacting molecules, like the RNA polymerase and DNA, appear to be capable to 'freeze fluctuations' in some spatial directions, resulting in the robustness of the cellular dynamical behaviors as well as their time hierarchy, observed in the live cell (Murray 2002; Winfree 2000).

The plethora of dynamical behaviors observed in living organisms reflects their immense regulatory capabilities to adapt to changing environments. Nevertheless, they are able to exhibit the most robust behaviors due to the stability of behaviors of their underlying systems.

The viability of an organism can be affected at different levels of organization: from the molecular interactions involved in key internal cellular processes to the mechanisms involved in cell-to-cell communication. The switching behavior, capable of decision making, and the oscillatory dynamics, capable of time tracking, seem to be elementary functions, necessary to determine the temporal behavior of an organism as a whole, which includes both continuous processes as well as choices between opposite processes.

In this thesis, we considered the effects of intercellular coupling, based on the diffusion of a signal molecule and sensitive to the population density, on the dynamics of protein and mRNA molecular species in a bacterial population of genetic clocks. We also have considered the effects of the gene sequence on the dynamics of small genetic networks possessing switching and oscillatory dynamical behaviors. Due to the importance of these two types of behavior, we believe that this study will assist in a better understanding of gene regulatory networks dynamics as well as of their evolutionary process.

Bibliography

Acar, M., J. Mettetal, and A. van Oudenaarden (2008). "Stochastic switching as a survival strategy in fluctuating environments". *Nat. Genet.* 40.4, pp. 471–5.

Alberts, B., A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter (2002). *Molecular Biology of the Cell.* 4th edition. New York: Garland Science.

Aldana, M. (2003). "Boolean dynamics of networks with scale-free topology". *Physica D* 185, pp. 45–66.

Andersen, J., C. Sternberg, L. Poulsen, S. Bjorn, M. Givskov, and S. Molin (1998). "New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria". *Appl. Environ. Microbiol.* 64.6, pp. 2240–6.

Andrianantoandro, E., S. Basu, D. Karig, and R. Weiss (2006). "Synthetic biology: new engineering rules for an emerging discipline". *Mol. Syst. Biol.* 2.

Arkin, A., J. Ross, and H. McAdams (1998). "Stochastic kinetic analysis of developmental pathway bifurcation in phage λ-infected *Escherichia coli* cells". *Genetics* 149.4, pp. 1633–48.

Ashcroft, M., M. Kubbutat, and K. Vousden (1999). "Regulation of p53 function and stability by phosphorylation". *Mol. Cell. Biol.* 19.3, pp. 1751–8.

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

Balázsi, G., A. Cornell-Bell, A. Neiman, and F. Moss (2001). "Synchronization of hyperexcitable systems with phase-repulsive coupling". *Phys. Rev. E* 64, p. 041912.

Benner, S. and A. Sismour (2005). "Synthetic biology". *Nat. Rev. Genet.* 6.7, pp. 533–43.

Boiteux, A., A. Goldbeter, and B. Hess (1975). "Control of oscillating glycolysis of yeast by stochastic, periodic, and steady source of substrate:

a model and experimental study". *Proc. Natl. Acad. Sci. USA* 72.10, pp. 3829–33.

Bratsun, D., D. Volfson, L. Tsimring, and J. Hasty (2005). "Delay-induced stochastic oscillations in gene regulation." *Proc. Natl. Acad. Sci. USA* 102.41, pp. 14593–8.

Buck, J. (1988). "Synchronous rhythmic flashing of fireflies. II". *Quater. Rev. Biol.* 63.3, pp. 265–89.

Burmann, B., K. Schweimer, X. Luo, M. Wahl, B. Stitt, M. Gottesman, and P. Rösch (2010). "A NusE:NusG complex links transcription and translation". *Science* 328.5977, pp. 501–4.

Butcher, J. (2008). *Numerical Methods for Ordinary Differential Equations.* 2nd edition. Wiley.

Callen, H. and T. Welton (1951). "Irreversibility and generalized noise". *Phys. Rev.* 83 (1), pp. 34–40.

Chance, B., R. Estabrook, and A. Ghosh (1964). "Damped sinusoidal oscillations of cytoplasmic reduced pyridine nucleotide in yeast cells". *Proc. Nat. Acad. Sci. USA* 51.6, pp. 1244–51.

Chang, H., P. Oh, D. Ingber, and S. Huang (2006). "Multistable and multistep dynamics in neutrophil differentiation". *BMC Cell Biol.* 7.1, p. 11.

Chang, H., M. Hemberg, M. Barahona, D. Ingber, and S. Huang (2008). "Transcriptome-wide noise controls lineage choice in mammalian progenitor cells". *Nature* 453.7194, pp. 544–7.

Cheng, C., KK. Yan, W. Hwang, J. Qian, N. Bhardwaj, J. Rozowsky, ZJ. Lu, W. Niu, P. Alves, M. Kato, M. Snyder, and M. Gerstein (2011). "Construction and analysis of an integrated regulatory network derived from high-throughput sequencing data". *PLoS Comput. Biol.* 7.11, e1002190.

Chua, G., Q. Morris, R. Sopko, M. Robinson, O. Ryan, E. Chan, B. Frey, B. Andrews, C. Boone, and T. Hughes (2006). "Identifying transcription factor functions and targets by phenotypic activation". *Proc. Natl. Acad. Sci. USA* 103.32, pp. 12045–50.

Collins, J. and I. Stewart (1993). "Coupled nonlinear oscillators and the symmetries of animal gaits". *J. Nonlin. Sci.* 3, pp. 349–92.

Cormack, B., R. Valdivia, and S. Falkow (1996). "FACS-optimized mutants of the green fluorescent protein (GFP)". *Gene* 173, pp. 33–8.

Crick, F. (1970). "Central dogma of molecular biology". *Nature* 227.5258, pp. 561–3.

Danino, T., O. Mondragón-Palomino, L. Tsimring, and J. Hasty (2010). "A synchronized quorum of genetic clocks". *Nature* 463, p. 326.

De la Fuente, A., P. Brazhnik, and P. Mendes (2002). "Linking the genes: inferring quantitative gene networks from microarray data". *Trends in Genetics* 18.8, pp. 395–8.

Dhooge, A., W. Govaerts, and Yu. Kuznetsov (2003). "MATCONT: a MATLAB package for numerical bifurcation analysis of ODEs". *ACM Transactions on Mathematical Software* 29.2, pp. 141–64.

Dockery, J. and J. Keener (2001). "A mathematical model for quorum sensing in *Pseudomonas aeruginosa*". *Bull. Math. Biol.* 63, pp. 95–116.

Doedel, E. (1981). "AUTO: A program for the automatic bifurcation analysis of autonomous systems". *Cong. Num.* Vol. 30. Univ. of Manitoba. Winnipeg, Canada., pp. 265–84.

Dunlap, J. (1999). "Molecular bases for circadian clocks". *Cell* 96, pp. 271–90.

Edelstein-Keshet, L. (1988). *Mathematical Models in Biology.* New York: McGraw-Hill.

Elowitz, M. and S. Leibler (2000). "A synthetic oscillatory network of transcriptional regulators". *Nature* 403, p. 335.

Engebrecht, J. and M. Silverman (1984). "Identification of genes and gene products necessary for bacterial bioluminescence". *Proc. Natl. Acad. Sci. USA* 81.13, pp. 4154–8.

Erie, D., O. Hajiseyedjavadi, M. Young, and P. von Hippel (1993). "Multiple RNA polymerase conformations and GreA: control of the fidelity of transcription". *Science* 262.5135, pp. 867–73.

Ermentrout, B. (2002). *Simulating, Analyzing and Animating Dynamical Systems: a Guide to XPPAUT for Researchers and Students.* 1st edition. SIAM.

Feigenbaum, M. (1978). "Quantitative universality for a class of nonlinear transformations". *J. Stat. Phys.* 19.1, pp. 25–52.

García-Ojalvo, J., M. Elowitz, and S. Strogatz (2004). "Modeling a synthetic multicellular clock: Repressilators coupled by quorum sensing". *Proc. Natl. Acad. Sci. USA* 101, p. 10955.

Gardner, T., C. Cantor, and J. Collins (2000). "Construction of a genetic toggle switch in *Escherichia coli*". *Nature* 403.6767, pp. 339–42.

Gekakis, N., D. Staknis, H. Nguyen, F. Davis, L. Wilsbacher, D. King, J. Takahashi, and C. Weitz (1998). "Role of the CLOCK protein in the mammalian circadian mechanism". *Science* 280.5369, pp. 1564–9.

Ghosh, A. and B. Chance (1964). "Oscillations of glycolytic intermediates in yeast cells". *Bioch. Biophys. Res. Commun.* 16.2, pp. 174–81.

Gibson, M. and J. Bruck (2000). "Efficient exact stochastic simulation of chemical systems with many species and many channels". *J. Phys. Chem. A* 104.9, pp. 1876–89.

Gillespie, D. (1976). "A general method for numerically simulating the stochastic time evolution of coupled chemical reactions." *J. Comput. Phys.* 22, pp. 403–34.

Gillespie, D. (1977). "Exact stochastic simulation of coupled chemical reactions." *J. Phys. Chem.* 81, pp. 2340–61.

Gillespie, D. (2007). "Stochastic simulation of chemical kinetics". *Ann. Rev. Phys. Chem.* 58, pp. 35–55.

Glansdorff, P. and I. Prigogine (1971). *Thermodynamic Theory of Structure, Stability and Fluctuations*. New York: John Wiley & Sons Ltd.

Glass, L. and M. Mackey (1988). *From Clocks to Chaos: The Rhythms of Life*. Princeton, NJ: Princeton University Press.

Goldbeter, A. (1996). *Biochemical Oscillations and Cellular Rhythms. The molecular bases of periodic and chaotic behaviour.* Cambridge: Cambridge University Press.

Goryachev, A. (2011). "Understanding bacterial cell-cell communication with computational modeling". *Chem. Rev.* 111.1, pp. 238–50.

Grams, T. and S. Thiel (2002). "High lightinduced switch from C3-photosynthesis to Crassulacean acid metabolism is mediated by UV-A/blue light". *J. Exper. Botany* 53.373, pp. 1475–83.

Greive, S. and P. von Hippel (2005). "Thinking quantitatively about transcriptional regulation". *Nat. Rev. Mol. Cell Biol.* 6, pp. 221–32.

Greive, S., S. Weitzel, J. Goodarzi, L. Main, Z. Pasman, and P. von Hippel (2008). "Monitoring RNA transcription in real time by using surface plasmon resonance". *Proc. Natl. Acad. Sci. USA* 105.9, pp. 3315–20.

Gupta, S. (2002). "A decision between life and death during TNF-α-induced signaling". *J. Clin. Immun.* 22 (4), pp. 185–94.

Han, S., C. Kurrer, and Y. Kuramoto (1995). "Dephasing and bursting in coupled neural oscillators". *Phys. Rev. Lett.* 75, pp. 3190–3.

Hänggi, P. and H. Thomas (1982). "Stochastic processes: time evolution, symmetries and linear response". *Physics Reports* 88.4, pp. 207 –319.

Hanson, F. (1978). "Comparative studies of firefly pacemakers". *Fed. Proc.* 37, pp. 2158–64.

Heinemann, M. and S. Panke (2006). "Synthetic biology — putting engineering into biology". *Bioinformatics* 22.22, pp. 2790–9.

Herbert, K., A. La Porta, B. Wong, R. Mooney, K. Neuman, R. Landick, and S. Block (2006). "Sequence-resolved detection of pausing by single RNA polymerase molecules". *Cell* 125.6, pp. 1083–94.

Hess, B. (1979). "The glycolytic oscillator". *J. Exper. Biol.* 81.1, pp. 7–14.

Hodgkin, A. and A. Huxley (1952). "A quantitative description of membrane current and its application to conduction and excitation in nerve". 117.4, pp. 500–44.

Holmberg, C., S. Tran, J. Eriksson, and L. Sistonen (2002). "Multisite phosphorylation provides sophisticated regulation of transcription factors". *Trends Biochem. Sci.* 27.12, pp. 619–27.

Honerkamp, J. (1983). "The heart as a system of coupled nonlinear oscillators". *J. Math. Biol.* 18, pp. 69–88.

Horsthemke, W. and R. Lefever (1983). *Noise-Induced Transitions: Theory and Applications in Physics, Chemistry, and Biology*. Vol. 15. Springer Series in Synergetics. Berlin: Springer.

Izhikevich, E. (2007). *Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting*. The MIT Press.

Jacob, F. and J. Monod (1961). "On the regulation of gene activity". *Cold Spring Harbor Symposia on Quantitative Biology* 26, pp. 193–211.

Jørgensen, F. and C. Kurland (1990). "Processivity errors of gene expression in *Escherichia coli*". *J. Mol. Biol.* 215.4, pp. 511–21.

Kaplan, D. and L. Glass (1995). *Understanding Nonlinear Dynamics*. New York: Springer.

Kaplan, H. and E. Greenberg (1985). "Diffusion of autoinducer is involved in regulation of the *Vibrio fischeri* luminescence system". 163.3, pp. 1210–4.

Karlebach, G. and R. Shamir (2008). "Modelling and analysis of gene regulatory networks". *Nat. Rev. Mol. Cell Biol.* 9.10, pp. 770–80.

Kauffman, S. (1969). "Metabolic stability and epigenesis in randomly constructed genetic nets". *J. Theor. Biol.* 22.3, pp. 437–67.

Kauffman, S. (2004). "A proposal for using the ensemble approach to understand genetic regulatory networks". *J. Theor. Biol.* 230, pp. 581–90.

Keener, J. and J. Sneyd (1998). *Mathematical Physiology*. New York-Berlin-Heidelberg: Springer-Verlag.

Keiler, K. (2008). "Biology of trans-translation". *Annu Rev Microbiol* 62, pp. 133–51.

Kubo, R., M. Toda, and N. Hashitsume (1985). *Statistical Physics II*. Vol. 31. Springer Series in Solid-State Science. Berlin,Heidelberg: Springer.

Kuznetsov, A., M. Kærn, and N. Kopell (2004). "Synchrony in a population of hysteresis-based genetic oscillators". *SIAM J. Appl. Math.* 65, p. 392.

Kuznetsov, Y. (2004). *Elements Of Applied Bifurcation Theory*. 3rd edition. New York: Springer.

Laje, R. and G. Mindlin (2002). "Diversity within a birdsong". *Phys. Rev. Lett.* 89, p. 288102.

Lee, H., M. Molla, C. Cantor, and J. Collins (2010). "Bacterial charity work leads to population-wide resistance". *Nature* 467.7311, pp. 82–5.

Lemons, D. and A. Gythiel (1997). "Paul Langevin's 1908 paper "On the Theory of Brownian Motion" ["Sur la théorie du mouvement brownien", C. R. Acad. Sci. (Paris) **146**, 530–533 (1908)]". *Amer. J. Physics* 65.11, pp. 1079–81.

Levskaya, A., A. Chevalier, J. Tabor, Z. Simpson, L. Lavery, M. Levy, E. Davidson, A. Scouras, A. Ellington, E. Marcotte, and C. Voigt (2005).

"Synthetic biology: engineering *Escherichia coli* to see light". *Nature* 438.7067, pp. 441–2.

Lewin, B. (2008). *Genes IX*. USA: Jones and Bartlett Publishers.

Lindner, B., J. García-Ojalvo, A. Neiman, and L. Schimansky-Geier (2004). "Effects of noise in excitable systems". *Physics Reports* 392.6, pp. 321 –424.

Lipshtat, A., A. Loinger, N. Balaban, and O. Biham (2006). "Genetic toggle switch without cooperative binding". *Phys. Rev. Lett.* 96.18, p. 188101.

Liu, C., X. Fu, L. Liu, X. Ren, C. Chau, S. Li, L. Xiang, H. Zeng, G. Chen, L. Tang, P. Lenz, X. Cui, W. Huang, T. Hwa, and J. Huang (2011). "Sequential establishment of stripe patterns in an expanding cell population". *Science* 334.6053, pp. 238–41.

Loinger, A. and O. Biham (2007). "Stochastic simulations of the repressilator circuit". *Phys. Rev. E* 76, p. 051917.

Lotka, A. (1909). "Contribution to the theory of periodic reactions". *J. Phys. Chem.* 14.3, pp. 271–4.

Mäkelä, J., J. Lloyd-Price, O. Yli-Harja, and A. Ribeiro (2011). "Stochastic sequence-level model of coupled transcription and translation in prokaryotes". *BMC Bioinformatics* 12.1, p. 121.

Marsden, J. and M. McCracken (1976). *The Hopf Bifurcation and Its Applications*. New York: Springer-Verlag.

Matsuzawa, A. and H. Ichijo (2001). "Molecular mechanisms of the decision between life and death: regulation of apoptosis by apoptosis signal-regulating kinase 1". *J. Biochem.* 130.1, pp. 1–8.

McAdams, H. and A. Arkin (1997). "Stochastic mechanisms in gene expression". *Proc. Natl. Acad. Sci. USA* 94, pp. 814–9.

McClure, W. (1980). "Rate-limiting steps in RNA chain initiation." *Proc. Natl. Acad. Sci. USA* 77, pp. 5634–8.

McMillen, D., N. Kopell, J. Hasty, and J. Collins (2002). "Synchronizing genetic relaxation oscillators by intercell signalling". *Proc. Natl. Acad. Sci. USA* 99, p. 679.

Meinhardt, H. (1982). *Models of Biological Pattern Formation*. New York: Academic Press Inc.

Metzner, W. (1993). "The jamming avoidance response in *Eigenmannia* is controlled by two separate motor pathways". *J. Neurosci.* 13.5, pp. 1862–78.

Miller, M. and B. Bassler (2001). "Quorum sensing in bacteria." *Annu. Rev. Microbiol.* 55, pp. 165–99.

Mitarai, N., K. Sneppen, and S. Pedersen (2008). "Ribosome collisions and translation efficiency: optimization by codon usage and mRNA destabilization". *J. Mol. Biol.* 382, pp. 236–45.

Mondragón-Palomino, O., T. Danino, J. Selimkhanov, L. Tsimring, and J. Hasty (2011). "Entrainment of a population of synthetic genetic oscillators". *Science* 333.6047, pp. 1315–9.

Moore, S. and R. Sauer (2005). "Ribosome rescue: tmRNA tagging activity and capacity in *Escherichia coli*". *Mol Microbiol* 58.2, pp. 456–66.

Müller, S., J. Hofbauer, L. Endler, C. Flamm, S. Widder, and P. Schuster (2006). "A generalized model of the repressilator". *J. Math. Biol.* 53, pp. 905–37.

Murray, J. (2002). *Mathematical Biology. I. An Introduction.* 3rd edition. Springer.

Nealson, K. and Hastings J. (1979). "Bacterial bioluminescence: its control and ecological significance". *Microbiol. Rev.* 43.4, pp. 496–518.

Njus, D., D. van Gooch, and J. Hastings (1981). "Precision of the *Gonyaulax* circadian clock". *Cell Biochem. Biophys.* 3 (3), pp. 223–31.

Nurse, P. (2000). "A long twentieth century of the cell cycle and beyond". *Cell* 100.1, pp. 71–8.

Nyquist, H. (1928). "Thermal agitation of electric charge in conductors". *Phys. Rev.* 32 (1), pp. 110–3.

Orrell, D. and H. Bolouri (2004). "Control of internal and external noise in genetic regulatory networks". *J. Theor. Biol.* 230.3, pp. 301–12.

Ozbudak, E., M. Thattai, I. Kurtser, A. Grossman, and A. van Oudenaarden (2002). "Regulation of noise in the expression of a single gene". *Nat. Genet.* 31.1, pp. 69–73.

Park, P. (2009). "ChIP-seq: advantages and challenges of a maturing technology". *Nat. Rev. Genet.* 10.10, pp. 669–80.

Paulsson, J. (2005). "Models of stochastic gene expression". *Phys. Life Rev.* 2.2, pp. 157–75.

Pedraza, J. and J. Paulsson (2008). "Effects of molecular memory and bursting on fluctuations in gene expression". *Science* 319.5861, pp. 339–43.

Perrin, B., L. Ralaivola, A. Mazurie, S. Bottani, J. Mallet, and F. d'Alché-Buc (2003). "Gene networks inference using dynamic Bayesian networks". *Bioinformatics* 19.suppl 2, pp. ii138–ii148.

Prigogine, I. and R. Lefever (1968). "Symmetry breaking instabilities in dissipative systems. II". *J. Chem. Phys.* 48.4, pp. 1695–1700.

Proshkin, S., A. Rahmouni, A. Mironov, and E. Nudler (2010). "Cooperation Between Translating Ribosomes and RNA Polymerase in Transcription Elongation". *Science* 328.5977, pp. 504–508.

Purnick, P. and R. Weiss (2009). "The second wave of synthetic biology: from modules to systems". *Nat. Rev. Mol. Cell Biol.* 10.6, pp. 410–22.

Pye, K. and B. Chance (1966). "Sustained sinusoidal oscillations of reduced pyridine nucleotide in a cell-free extract of *Saccharomyces carlsbergensis*". *Proc. Nat. Acad. Sci. USA* 55.4, pp. 888–94.

Rajala, T., A. Häkkinen, S. Healy, O. Yli-Harja, and A. Ribeiro (2010). "Effects of transcriptional pausing on gene expression dynamics". *PLoS Comput. Biol.* 6, e1000704.

Ribeiro, A. (2008). "Dynamics and evolution of stochastic bistable gene networks with sensing in fluctuating environments". *Phys. Rev. E* 78, p. 061902.

Ribeiro, A. (2010). "Stochastic and delayed stochastic models of gene expression and regulation". *Math. Biosci.* 223.1, pp. 1–11.

Ribeiro, A. and S. Kauffman (2007). "Noisy attractors and ergodic sets in models of gene regulatory networks". *J. Theor. Biol.* 247.4, pp. 743–55.

Ribeiro, A., R. Zhu, and S. Kauffman (2006). "A general modeling strategy for gene regulatory networks with stochastic dynamics". *J. Comput. Biol.* 13.9, pp. 1630–9.

Ribeiro, A., O. Smolander, T. Rajala, A. Häkkinen, and O. Yli-Harja (2009a). "Delayed stochastic model of transcription at the single nucleotide level". *J. Comp. Biol.* 16.4, pp. 539–53.

Ribeiro, A., X. Dai, and O. Yli-Harja (2009b). "Variability of the distribution of differentiation pathway choices regulated by a multipotent delayed stochastic switch". *J. Theor. Biol.* 260.1, pp. 66–76.

Rinzel, J. (1981). "Nonlinear Phenomena in Physics and Biology". Ed. by R.H. Enns, B.L. Jones, R.M. Miura, and S.S. Rangnekar. New York: Plenum Press. Chap. Models in neurobiology, pp. 345–367.

Roussel, M. and R. Zhu (2006). "Validation of an algorithm for delay stochastic simulation of transcription and translation in prokaryotic gene expression." *Phys. Biol.* 3, pp. 274–84.

Rubinow, S. (1975). *Introduction to Mathematical Biology*. New York: Wiley.

Schrödinger, E. (1926). "An undulatory theory of the mechanics of atoms and molecules". *Phys. Rev.* 28 (6), pp. 1049–70.

Shoji, S., S. Walker, and K. Fredrick (2009). "Ribosomal translocation: one step closer to the molecular mechanism". *ACS Chem. Biol.* 4.2, pp. 93–107.

Sørensen, M. and S. Pedersen (1991). "Absolute in vivo translation rates of individual codons in *Escherichia coli* : The two glutamic acid codons GAA and GAG are translated with a threefold difference in rate". *J. Mol. Biol.* 222.2, pp. 265–80.

Stevens, A., K. Dolan, and E. Greenberg (1994). "Synergistic binding of the *Vibrio fischeri* LuxR transcriptional activator domain and RNA polymerase to the *lux* promoter region". 91.26, pp. 12619–23.

Stricker, J., S. Cookson, M. Bennett, W. Mather, L. Tsimring, and J. Hasty (2008). "A fast, robust and tunable synthetic gene oscillator". *Nature* 456, pp. 516–9.

Strogatz, S. (2001). *Nonlinear Dynamics And Chaos: With Applications To Physics, Biology, Chemistry, And Engineering*. Westview Press.

Strogatz, S., R. Mirollo, and P. Matthews (1992). "Coupled nonlinear oscillators below the synchronization threshold: relaxation by generalized Landau damping". *Phys. Rev. Lett.* 68 (18), pp. 2730–3.

Süel, G., J. García-Ojalvo, L. Liberman, and M. Elowitz (2006). "An excitable gene regulatory circuit induces transient cellular differentiation". *Nature* 440.7083, pp. 545–50.

Swain, P., M. Elowitz, and E. Siggia (2002). "Intrinsic and extrinsic contributions to stochasticity in gene expression". *Proc. Natl. Acad. Sci. USA* 99.20, pp. 12795–800.

Thieffry, D., A. Huerta, E. Pérez-Rueda, and J. Collado-Vides (1998). "From specific gene regulation to genomic networks: a global analysis of transcriptional regulation in *Escherichia coli*". *BioEssays* 20.5, pp. 433–40.

Tigges, M., T. Marquez-Lago, J. Stelling, and M. Fussenegger (2009). "A tunable synthetic mammalian oscillator". *Nature* 457, pp. 309–12.

Tuller, T., A. Carmi, K. Vestsigian, S. Navon, Y. Dorfan, J. Zaborske, T. Pan, O. Dahan, I. Furman, and Y. Pilpel (2010). "An evolutionarily conserved mechanism for controlling the efficiency of protein translation". *Cell* 141.2, pp. 344–54.

Tuttle, L., H. Salis, J. Tomshine, and Y. Kaznessis (2005). "Model-driven designs of an oscillating gene network". *Biophys. J.* 89.6, pp. 3873–83.

Tyson, J. and S. Kauffman (1975). "Control of mitosis by a continuous biochemical oscillation: synchronization; spatially inhomogeneous oscillations". *J. Math. Biol.* 1, pp. 289–310.

Ullner, E., A. Zaikin, E. Volkov, and J. García-Ojalvo (2007). "Multistability and clustering in a population of synthetic genetic oscillators via phase-repulsive cell-to-cell communication". *Phys. Rev. Lett.* 99.14, p. 148103.

Ullner, E., A. Koseska, J. Kurths, E. Volkov, H. Kantz, and J. García-Ojalvo (2008). "Multistability of synthetic genetic networks with repressive cell-to-cell communication". *Phys. Rev. E* 78, p. 031904.

Van Kampen, N. (2007). *Stochastic Processes in Physics and Chemistry*. 3rd edition. Amsterdam: North Holland.

Volkov, E. and V. Romanov (1995). "Bifurcations in the system of two identical diffusively coupled Brusselators". *Phys. Scr.* 51.1, p. 19.

Volkov, E. and M. Stolyarov (1991). "Birhythmicity in a system of two coupled identical oscillators". *Phys. Lett. A* 159, pp. 61–6.

Waters, C. and B. Bassler (2005). "Quorum sensing: cell-to-cell communication in bacteria". *Ann. Rev. Cell Dev. Biol.* 21, pp. 319–46.

Welch, M., S. Govindarajan, J. Ness, A. Villalobos, A. Gurney, J. Minshull, and C. Gustafsson (2009). "Design parameters to control synthetic gene expression in *Escherichia coli*". *PLoS ONE* 4.9, e7002.

Wen, J., L. Lancaster, C. Hodges, A. Zeri, S. Yoshimura, H. Noller, C. Bustamante, and I. Tinoco (2008). "Following translation by single ribosomes one codon at a time". *Nature* 452.7187, pp. 598–603.

Winfree, A. (1967). "Biological rhythms and the behavior of populations of coupled oscillators". *J. Theor. Biol.* 16.1, pp. 15–42.

Winfree, A. (1987). *The Timing of Biological Clocks*. New York: Scientific American Books, Inc.

Winfree, A. (2000). *The Geometry of Biological Time*. 2nd edition. Berlin-Heidelberg-New York: Springer-Verlag.

Yagil, E. and G. Yagil (1971). "On the relation between effector concentration and the rate of induced enzyme synthesis". *Biophys. J.* 11, pp. 11–27.

Yu, J., J. Xiao, X. Ren, K. Lao, and S. Xie (2006). "Probing gene expression in live cells, one protein molecule at a time." *Science* 311, pp. 1600–3.

Zhabotinsky, A. (1991). "A history of chemical oscillations and waves". *Chaos* 1.4, pp. 379–86.

Zhu, R. and D. Salahub (2008). "Delay stochastic simulation of single-gene expression reveals a detailed relationship between protein noise and mean abundance". *FEBS Lett.* 582.19, pp. 2905–2910.

Zhu, R., A. Ribeiro, D. Salahub, and S. Kauffman (2007). "Studying genetic regulatory networks at the molecular level: delayed reaction stochastic models". *J. Theor. Biol.* 246.4, pp. 725–45.

Errata for the publications

Publication-III

- In Sec. II, paragraph 4, the sentence "The model accounts for the nucleotides...", the "(Δ_{P})" should be read as "($2\Delta_{\mathrm{P}}+1$)"
- In the Table I caption, "Δ_{P}, range of nucleotides that Rp occupies, $\Delta_{\mathrm{P}}=25$. Δ_{R}, range of ribonucleotides that ribosome occupies, $\Delta_{\mathrm{R}}=31$" should be read as "($2\Delta_{\mathrm{P}}+1$), range of nucleotides that Rp occupies, $\Delta_{\mathrm{P}}=12$. ($2\Delta_{\mathrm{R}}+1$), range of ribonucleotides that ribosome occupies, $\Delta_{\mathrm{R}}=15$".
- In the Table I, reaction **22** (22nd row of the table), the delay time for protein folding should be $\tau_{\mathrm{fold}}=420\pm100$.

The results and conclusions are not affected by these changes.

Publications

Publication I

I. Potapov, E. Volkov, A. Kuznetsov, *Dynamics of coupled repressilators: the role of mRNA kinetics and transcription cooperativity*, Phys. Rev. E **83**, 031901, 2011.

Publication II

I. Potapov, J. Lloyd-Price, O. Yli-Harja and A.S. Ribeiro, *Dynamics of a genetic switch at the nucleotide and codon levels*, 8^{th} International Workshop on Computational Systems Biology (WCSB), Switzerland, Zurich, 2011.

Publication III

I. Potapov, J. Lloyd-Price, O. Yli-Harja and A.S. Ribeiro, *Dynamics of a genetic toggle switch at the nucleotide and codon levels*, Phys. Rev. E **84**, 031903, 2011.

Publication IV

I. Potapov, B. Zhurov, E. Volkov, *"Quorum sensing" generated multistability and chaos in a synthetic genetic oscillator*, Chaos **22**, 023117, 2012.

Publication V

I. Potapov, J. Mäkelä, O. Yli-Harja and A.S. Ribeiro, *Effects of codon sequence on the dynamics of genetic networks*, J. Theor. Biol. **315**, 17–25, 2012.

Tampereen teknillinen yliopisto
PL 527
33101 Tampere

Tampere University of Technology
P.O.B. 527
FI-33101 Tampere, Finland

ISBN 978-952-15-2995-5
ISSN 1459-2045