University of Warsaw
Faculty of Mathematics, Informatics and Mechanics

Paulina Szymańska-Rożek

Space, noise, and information transmission in mathematical modelling of signalling pathways

PhD dissertation

Supervisors:

prof. dr hab. Jacek Miękisz
Institute of Applied Mathematics and Mechanics
University of Warsaw

prof. dr hab. Tomasz Lipniacki
Institute of Fundamental Technological Research
Polish Academy of Sciences

June 2018

Author's declaration:
aware of legal responsibility I hereby declare that I have written this dissertation myself and all the contents of the dissertation have been obtained by legal means.

June 8, 2018 .....

date *Paulina Szymańska-Rożek*

Supervisors' declaration:
the dissertation is ready to be reviewed

June 8, 2018 .....

date *prof. dr hab. Jacek Miękisz*

June 8, 2018 .....

date *prof. dr hab. Tomasz Lipniacki*

Abstract

This thesis covers three mathematical models of theoretical biology, more precisely of signalling pathways. Their common core is the mathematical tool applied to describe them: they are all modelled by a Markov Chain with continuous time. The three parts focus on a different feature of the signalling pathways - as the titles of the three chapters suggest it, I investigate: spatiality, inherent stochasticity, and the phenomenon of information transmission.

In the first part we investigate spatial effects of a phosphorylation–dephosphorylation cycle taking place on a biological membrane. The dynamics of such a network is governed by reaction rates, which are strongly influenced by diffusivity of reactants, their subcellular localization, and nonspecific molecular crowding. In particular, I analyse the dependence of effective macroscopic reaction rate coefficients on diffusion, and we found, among others, that effective macroscopic reaction rates depend in a nontrivial manner on the diffusion coefficient. Generally, they decrease with decreasing diffusion and their formulae contain a term linearly proportional to the diffusion coefficient. However, there is the spatial effect that complicats the analysis – in dense systems, especially at low enzyme concentration, enzymes become encircled by converted substrates, which lowers substantially the effective reaction rate. Additionally, steady states of (de)phosphorylated substrates are controlled by molecular crowders which, mostly by lowering the effective diffusion of reactants, favour the more abundant enzyme.

In the second part we highlight the importance of stochasticity underlying every process that involves small numbers of elements. Using stochastic analyses of a bistable genetic toggle switch, we developed a control strategy that maximizes the chances that a cell, chosen among identical cells, will express one phenotype, while the rest express another. We restrict to the strategies that irradiate all cells simultaneously with the same intensity, enhancing protein degradation in all cells identically. Control of individual cells is made possible only by monitoring stochastic protein fluctuations and applying UV control at favourable times and levels. Such control is theoretically impossible if restricted to a deterministic setting. Among other results, we showed that for two identical cells, our stochastic control law can drive protein expression of a chosen cell above its neighbour with a better than 99% success rate.

In the third part we model how information in a simple regulatory network is transmitted. The analyzed models are simple regulatory circuits comprising of two binary random variables. Mutual information measured between them is considered to quantify how much information is transmitted in the system. The aim of this research was to find optimal mutual information under constraints set on the energy (entropy production rate) available, which is a function of the parameters defining the system. We compared and classified the models without feedback and with feedback, starting in the steady state or out of the steady state, in terms of the amount of information transmitted and the energetic cost of this trans-

mission. It turns out that if the system starts at steady state, feedback is beneficial in terms of optimal information, but the costs of optimal information are the same in both model variants. In the case of the initial distribution subjected to optimization there is less difference in the optimal information, but the cost remains highly larger if there is no feedback.

Keywords: effective reaction rate, kinetic Monte Carlo, stochastic gene expression, genetic toggle switch, mutual information, entropy production rate

AMS Classification: 60J, 82D, 92C, 92C05, 92C40

Streszczenie

Niniejsza rozprawa dotyczy trzech modeli zaczerpniętych z dziedziny biologii teoretycznej, a konkrenie modelowania ścieżek sygnał owych. Ich wspólnym rdzeniem jest narzędzie matematyczne użyte do opisu, a mianowicie Łańcuch Markowa z czasem ciągłym. W trzech częściach pracy skupiam się na różnych aspektach modelowania: przestrzenności, stochastyczności i przekazywaniu informacji.

W pierwszej części badamy efekty przestrzenne w cyklach fosforylacji–defosforylacji na błonie biologicznej. Dynamika takich sieci zależy od stałych szybkości reakcji, które to z kolei silnie zależą od współczynników dyfuzji reagentów, ich lokalizacji w komórce i zagęszczenia molekularnego. W szczególności, przeanalizowaliśmy zależność efektywnych makroskopowych stałych reakcji od stałej dyfuzji i nasza analiza wykazała między innymi, że efektywne makroskopowe stałe reakcji zależą nietrywialnie od współczynnika dyfuzji. Ogólny wniosek jest taki, że maleją one wraz z malejącą dyfuzją, a wzory zawierają liniową zależność od współczynnika dyfuzji. Obliczenia komplikuje efekt przestrzenny – w zapełnionych sieciach (gęstych układach), szczególnie przy niskich stężeniach enzymów, te ostatnie zostają otoczone przez subtratraty, które już weszły z nimi w reakcję. To istotnie zmniejsza efektywne stałe szybkości reakcji. Ponadto stacjonarne frakcje ufosforylowanych i zdefosforyzowanych substratów zależą od zatłoczenia molekularnego, które poprzez zmniejszenie efektywnych stałych reakcji "sprzyjają" temu enzymowi, którego jest więcej.

W drugiej części badamy znaczenie stochastyczności wpisanej w każdy proces, w którym liczba rozważanych elementów (np. produkowanych cząsteczek białka) jest mała. Wykorzystując stochastyczną analizę bistabilnego przełącznika genetycznego, definiujemy strategię, która maksymalizuje prawdopodobieństwo, że jedna komórka, wybrana spośród identyczych komórek, będzie miała inny fenotyp niż reszta. Używaną strategią jest naświetlanie wszystkich komórek jednakowo promieniowaniem UV, degradującym białka w komórkach. Taka strategia działa tylko wtedy, gdy wykorzystamy stochastyczne fluktuacje i włączymy promieniowanie w odpowiednich momentach i z odpowiednim natężeniem. W przypadku, gdy ewolucja układu będzie wyłącznie deterministyczna, taka strategia nie zadziała. Jednym z kilku wyników jest pokazanie, że dla dwóch identycznych komórek, przy użyciu skonstruowanej strategii poziom białka w wybranej komórce będzie wyższy niż w drugiej z prawdopodobieństwem 0.99.

W trzeciej części modelujemy, jak przekazywana jest informacja w prostym układzie biologicznym, składającym się z dwóch binarnych zmiennych losowych. Wyliczona informacja wzajemna jest interpretowana jako ilość informacji przekazanej w systemie. Celem badań zawartych w trzecim rozdziale było znalezienie maksymalnej informacji wzajemnej z ograniczeniem położonym na dostępną energię - szybkość produkcji entropii, która to jest natomiast funkcją parametrów - szybkości przejść między stanami. Porównywaliśmy optymalną ilość przekazanej informacji i koszt energetyczny dla modeli bez sprzężenia i ze sprzężeniem zwrotnym, ze stacjonarnym lub niestacjonarnym warunkiem początkowym. Pokazal-

iśmy, że w przypadku stacjonarnego warunku początkowego, sprzężenie zwrotne pozwala na przekazanie większej ilości informacji, ale koszt energetyczy jest taki sam, jak dla modelu bez sprzężenia zwrotnego. Jeśli warunek początkowy jest również poddany optymalizacji, wówczas dla obu modeli (bez i ze sprzężeniem zwrotnym) jest mniejsza rżnica w ilości przekazanej informacji, ale koszt energetyczny jest dużo wywższy dla modelu bez sprzężenia zwrotnego.

Acknowledgements

First and foremost I thank my invaluable supervisors, Jacek Miękisz and Tomasz Lipniacki. They have been for me real mentors – helpful, available and patient. They guided me throughout my PhD studies and have become not only supervisors but also friends. I learnt a lot from them and probably I cannot express enough my gratitude.

I would also like to thank Aleksandra Walczak who hosted me for nearly a year at Ecole Normale Supérieure and devoted a lot of her time for our research. I am grateful for giving me an interesting topic to work on, and for all the discussions and encouragement I received from her.

Many thanks go to my colleagues with whom I collaborated, especially Marek Kochańczyk, a flag example of helpfulness and selflessness.

I thank my family who enabled me to pursue my studies and all my friends for their prayers and encouragement.

AMDG

Contents

Introduction

This thesis is dedicated to mathematical modelling of signalling pathways. Modelling as a useful concept to approximate the phenomena of the surrounding world has been used for quite a while now in all sciences. Even in humanities researchers has perceived the benefits of building models – even though the description of what is modelled might get simplified, it is the invaluable ability to start a formal analysis and deduction that renders modelling so attractive to scientists. And it is obviously not whichever modelling that enables precise reasoning and concluding – indeed, the Queen of Sciences provides the best tools for translating the reality into a robust enough and yet accurate mathematical model. Since mathematical modelling has become almost a separate branch of Mathematics, and it widens its scope all the time, this thesis comprises three, out of numerous, aspects of this modelling - space, noise, and information transmission. All three are analyzed on the canvas of a different model, yet they all belong to a wide domain of "signalling pathways" models. Signalling pathways are networks of interacting molecules governing probably all activities of a cell. A signalling pathway can be of any degree of complexity, and indeed, those found in nature are usually multi-component networks, enabling feedback, signal amplification, cascades of signals, etc.. Obviously, this is in the modeller's task to simplify such huge networks enough so that the simplified version exhibits the same behaviour, yet it is much simpler to handle mathematically and (or) numerically.

It is not a new trend to cross borders between disciplines that in school are separate subjects. We were not taught by Helmholtz, Maxwell, and Rayleigh, and perhaps that is why we do not see the complementarity of biology, chemistry, physics, and other sciences, because our teachers did not show us how closely related they are. Also, we are not used to think that phenomena can be translated into mathematical models and subjected to a rigorous mathematical analysis. Luckily, in graduate studies in Science it becomes obvious that the intersection of disciplines is nonempty and that an adequate mathematical model is required in order to formulate statements to validate heuristic reasoning and check the intuition. This thesis enjoys such interdisciplinarity, as all three families of the models are drawn from theoretical biology, physical chemistry, and biophysics.

Lastly, but perhaps most importantly, all models presented here have a common mathematical background - Markov Chains with continuous time, the *Markov Jump Process*. At

the beginning of every chapter I describe the state space and transition rates that define the underlying Markov Jump Process. Formal definition and mathematical construction of the state spaces, as well as some facts about Markov Jump Process are provided in the Appendix.

Construction of the thesis

As mentioned above, this thesis is divided into three parts (chapters); each one of them resulted from a different question asked about mathematical modelling of signalling pathways.

The first part is about spatiality. In particular, it focuses on the rates of coupled biochemical reactions, and as these take place mainly on the membrane, the presented model accounts for spatial effects such as the subcellular localization of the reactants, and non-specific molecular crowding. I chose to present space as the first of the three aspects of mathematical modelling of signalling pathways not only because chronologically it was the first research I undertook in my PhD studies, but also because it is still a growing topic in applied mathematics, biomathematics, and mathematical modelling. I remember one of the first conferences I attended and a memorable introductory slide on partial differential equations in cancer modelling. The witty professor showed a snapshot from "Star Trek" with the famous quotation: "space is the final frontier". This is not just a bon-mot from a good film, but a diagnose of the direction in which research in the mentioned fields has headed in the last years and will still be heading in the future. Maybe not surprisingly my supervisors were interested in taming space in modelling of even simple, but stochastic phenomena encountered in biology. At that time a co-student and colleague of mine, Marek Kochańczyk, developed a powerful software, SpatKin, that effectively performs spatial Kinetic Monte Carlo simulations. Our goal was to support numerical simulations by as accurate as possible approximations. The latter resulted from purely analytical considerations and were then "verified" by performed numerical simulations. We published our results in [50]. Later on another colleague, Paweł Nałęcz-Jawecki, continued and extended our cooperation, and published the results in [40]. We aimed at establishing the dependencies between diffusion and densities of substrates and enzymes in a model of a reversible reaction cycle (phosphorylation–dephosphorylation). We investigated effects that might play a role there, for example the molecular crowding or the size of the membrane, formation of transient enzyme-substrate complexes. We found analytical expressions for the effective macroscopic reaction rate constants (EMRRCs) in two limits of infinite and zero motility. For non-zero but small diffusion, we found numerically that the EMRRCs contain a term linearly proportional to the diffusion coefficient. EMRRCs decrease with decreasing diffusion and this dependence is stronger for the less abundant enzyme. Steady-state fraction of substrates in a given state (phosphorylated/dephosphorylated) can increase or decrease with diffusion,

depending on relative concentrations of enzymes changing the state of the substrate (kinases phosphorylate, phosphatases dephosphorylate). Molecular crowders favour the more abundant enzyme. Analytical reasoning comprising *mean-first passage time* considerations approximate very well spatial kinetic Monte Carlo simulations.

The subject of the second part of the thesis is stochasticity. In biophysics it has become an axiom that fluctuations are present in cellular processes [7] and scientists argue that noise in gene regulation for instance, enables adaptation and thus survival of cells [30, 56, 48]. But I will focus there on yet another feature of a "noisy" signalling pathway - its controllability. The study began as a project at the 2014 q-bio Summer School (qbSS, Albuquerque, New Mexico) and resulted in a publication in *Physical Biology* in 2015 [51]. In this work my colleagues, Johannes Keegstra, Nicola Gritti, Mohammad Soltani, and myself, under the supervision of Brian Munsky, built simplified mathematical and computational models to capture the dynamics of stochastic gene regulatory responses when subjected to temporary fluctuating, yet spatially homogenous, environmental conditions. We explored the theoretical possibility to select at random, from within a population, one cell, and then to use a *single, spatially uniform input* to control that particular cell, that is to drive it to achieve a different phenotype than the rest. It is not possible if the model is strictly deterministic, so we explored the extent to which it is possible in the setting that accounts for noise at the single-cell level. One of the most spectacular results is that the stochastic control law (UV radiation, described in details in the appropriate section), applied identically to the whole population of cells, can render the chosen cell phenotypically different with a very high precision - in the case of a two-cell population, the probability that the chosen cell performs better (synthesises more proteins) is 99% and for a population of 30 identical cells we were able to maintain the chosen cell within the top 20% cells. The "freshness" of our work for biophysicists resulted from the fact that cellular noise is typically said to impair the predictability of biological responses; we managed to show that it can improve their controllability. All figures presented in this part were published in [51], and I use them with their author's, Johannes Keegstra, kind permission.

Finally, the scope of the third chapter is information transmission in signalling pathways. The biological inspiration for this part was the recurring question about how biological circuits transmit signals and at what price. The analyzed models are relatively simple, consisting of two binary random variables, interacting with one another. The specific research aim was to compute optimal mutual information, interpreted as information transmitted in the system, under constraints set on entropy production rate, interpreted as energy dissipated during the process of transmitting information. Considered models (they differed in transition rules and initial conditions) were compared in terms of optimal mutual information and the energetic cost. The mathematical tools exploited in this part comprise

discrete information theory and constrained optimization. Although we have touched somehow the phenomenon of information transmission in the first chapter, where the products of chemical reactions were some sort of "information carrier" about the density of enzymes and substrates, the size of the reactor etc., the term "information" can be used there rather colloquially. In the last chapter, we will treat it mathematically, introducing two random variables between which mutual information will be measured. Mathematically speaking, we will optimize mutual information with a constraint set on entropy production rate, which is, as will be shown, a function of . Heuristically speaking, if energy is considered a cost, we will ask a simple and fashionable question: how costly information is? And how can we transmit as much information as possible having limited energetic resources? The results presented in this part were obtained in collaboration with Aleksandra Walczak during my one year internship at Ecole Normale Supérieure in Paris. Preliminary results presented in this part concerning information optimization case of no constraint set on "energy" were first obtained by Francesca Mancini in [33]. Some further results for the case of constrained energy, but only for when the system starts at the steady state were published in [32]. Within my research conducted in Paris I continued these investigations, extending the results to the out of steady state cases and I introduced and evaluated the "cost" of optimal information. I found, among others, that although it is usually more efficient for the amount of information transmitted if there is a feedback from the output variable to the input variable, the cost of transmission is greater for the no-feedback variant only if the systems are allowed to start out of the steady state. If they start at the steady state, the cost of information transmission for the no-feedback system and the system with feedback is the same.

Chapter 1

Space

> If you know where you are, you do not know how fast you are moving, and if you know how fast you are moving, you do not know where you are.
>
> ven. Fulton J. Sheen
> *Philosophy of Science*

1.1 Motivation

In this chapter I introduce and analyze a model of chemical reactions that take place in two-dimensional structures. The biological inspiration was drawn from studying plasma membranes, where signalling pathways enable communication between the cell and the outer world."Space" thus refers to a 2-D plane with periodic boundary conditions, so that it mimics the spherical character of the cell membrane - every lattice site has six neighbours and there is no "border" as such. The mathematical task was to embrace the spatial aspect of such phenomenon. One would naturally think to do it through partial differential equations, precisely reaction-diffusion equations. Our modelling approach was different - we discretized the continuous space into sites thus obtaining a (finite) lattice. However, the time in our model remained a continuous variable, so that the mathematical concept we obtained is the *Markov Jump Process* (see the Appendix).

The chemical canvas we work on is a cycle of coupled antagonistic reactions. There are two kinds of enzymes, kinases and phosphatases, that change the state of the substrate - it is either phosphorylated (by kinases) or dephosphorylated (by phosphatases). This phosphorylation–dephosphorylation motif is therefore a cycle of reversible, antagonistic reactions. I will not go into chemical details, as this is only a working example of any such motif where a molecule can be in either of two states and it needs the encounter with a

proper molecule type to change the state. It can be written down as:

$$\mathrm{K} + \mathrm{S_u} \xrightarrow{c} \mathrm{K} + \mathrm{S_p}, \tag{1.1a}$$

$$\mathrm{P} + \mathrm{S_p} \xrightarrow{d} \mathrm{P} + \mathrm{S_u}, \tag{1.1b}$$

where $\mathrm{S_u}$ and $\mathrm{S_p}$ stand for dephosphorylated and phosphorylated substrates, respectively, K represents the kinase, and P – the phosphatase. The letters above the arrows represent corresponding microscopic reaction rates.

Just as a reference to the above scheme, we also consider a variant in which dephosphorylation is a first-order reaction, i.e.,

$$\mathrm{S_p} \xrightarrow{d_0} \mathrm{S_u}, \tag{1.2}$$

whereas phosphorylation still occurs as previously.

The dynamics of such networks is governed by reaction rates, which are strongly influenced by diffusivity of reactants [12], their subcellular localization, and molecular crowding [15, 18, 24]. Therefore the overall aim of this part is to establish the dependencies between the densities of substrates in a given state (phosphorylated/dephosphorylated) and the densities of the two enzymes (kinases and phosphatases) and diffusion. However, we also discuss other possible effects that might play a role, for example the mentioned molecular crowding or the size of the membrane (number of lattice sites).

1.2 Mathematical description

Having introduced in the previous section our model, we can specify the state space and the transition rates that characterize the considered Markov Chain. The state space is a set of functions attributing to every lattice site a given molecule occupying it, or an empty element. It is given explicitly for every model considered, in the Appendix. Transitions are defined by microscopic reaction rates c and d, and the microscopic diffusion rate, motility, m.

Since our aim is to find stationary mean densities of molecules, we will have to calculate the expected value of a binary random variable X_i:

$$X_i = \begin{cases} 1 & \text{if the } i\text{-th site is occupied by M} \\ 0 & \text{if the } i\text{-th site is not occupied by M,} \end{cases} \tag{1.3}$$

and sum it over all lattice sites i. M is the symbol of the substrate molecule, either phosphorylated or not, $M \in \{\mathrm{S_u}, \mathrm{S_p}\}$. We will not calculate the densities of enzymes, as their counts are fixed.

1.3 Mean field ansatz

Before surrendering to numerical methods and finding the expected value of X_i by simulations only, we tried to "reduce" the space to two non-spacial coefficients that would be used in the ordinary differential equation for the time evolution of substrate densities:

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho_{\mathrm{S_u}} = -c_{\mathrm{eff}}\rho_{\mathrm{K}}\rho_{\mathrm{S_u}} + d_{\mathrm{eff}}\rho_{\mathrm{P}}\rho_{\mathrm{S_p}}, \tag{1.4a}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho_{\mathrm{S_p}} = c_{\mathrm{eff}}\rho_{\mathrm{K}}\rho_{\mathrm{S_u}} - d_{\mathrm{eff}}\rho_{\mathrm{P}}\rho_{\mathrm{S_p}}. \tag{1.4b}$$

These two equations are complementary, since their solutions satisfy $\rho_{\mathrm{S_u}}(t) + \rho_{\mathrm{S_p}}(t) = \rho_{\mathrm{S}} =$ const (total number of substrate molecules is fixed). The steady-state solution of Eqs. (1.4) reads:

$$\rho_{\mathrm{S_u}} = \frac{d_{\mathrm{eff}}\rho_{\mathrm{P}}}{c_{\mathrm{eff}}\rho_{\mathrm{K}} + d_{\mathrm{eff}}\rho_{\mathrm{P}}}\rho_{\mathrm{S}}, \tag{1.5a}$$

$$\rho_{\mathrm{S_p}} = \frac{c_{\mathrm{eff}}\rho_{\mathrm{K}}}{c_{\mathrm{eff}}\rho_{\mathrm{K}} + d_{\mathrm{eff}}\rho_{\mathrm{P}}}\rho_{\mathrm{S}}. \tag{1.5b}$$

Equations (1.4) are linear and their stationary solution is globally asymptotically stable. We call c_{eff} the *effective macroscopic phosphorylation rate constant* and d_{eff} the *effective macroscopic dephosphorylation rate constant*. These coefficients are supposed to approximate as accurately as possible the joint effect of all molecules - this is the essence of the so-called *mean field approximation*.

If we were able to track the time evolution of our system and count the reactions that occurred until a given time, we would use the following definition of the *effective macroscopic reaction rates*, (EMRRCs):

$$c_{\mathrm{eff}} = \frac{n}{\rho_{\mathrm{S_u}}\rho_{\mathrm{K}}V\Delta t}, \tag{1.6a}$$

$$d_{\mathrm{eff}} = \frac{n}{\rho_{\mathrm{S_p}}\rho_{\mathrm{P}}V\Delta t} \tag{1.6b}$$

where n is the number of (de)phosphorylation reactions that fired during a time interval Δt and V is the lattice surface area (i.e., total number of lattice sites). The densities of kinases, phosphatases and substrates are denoted by ρ with a respective subscript: ρ_{K}, ρ_{P}, $\rho_{\mathrm{S_u}}$, and $\rho_{\mathrm{S_p}}$. But these can only be obtained by numerical methods, which I describe in the next section.

1.4 Numerical methods

The time evolution of chemical kinetics is simulated through the stochastic simulation algorithm, SSA, first proposed by Gillespie [17]. In its classical formulation:

- the initial state of the system is chosen;
- for a given configuration, the transition rates, λ_i, are calculated;
- time τ until the next reaction is drawn from the exponential distribution with parameter (mean) $\frac{1}{\sum_i \lambda_i} = 1/\lambda$;
- the transition type is chosen by drawing from the discrete distribution with probability $P(\text{transition} = i) = \lambda_i/\lambda$;
- the system is updated until a stopping time.

The above algorithm was extended to account for a different than usual"reaction" allowed - hopping of molecules between adjacent empty sites. This change was implemented by a colleague of mine, Marek Kochańczyk [27, 60, 26]. So first, molecules are placed on discrete sites of a 2-dimensional triangular lattice which forms a square domain with periodic boundary conditions. Possible transitions (hopping of molecules to adjacent empty lattice sites or reactions) are listed with their respective probabilities. Time-step is drawn at random from the exponential distribution with the rate parameter equal to the sum of the rates of all possible events. The rates of transitions are defined by microscopic rate constants, c and d, and motilities, m (the propensity of hopping to a neighboring empty site on the triangular lattice is $m/6$). Motilities are assumed to be equal for molecules of all types (unless otherwise specified). After every event, the list of all events is updated. However, since the change in the system configuration after every simulation step is local, only a partial update of the list is necessary.

Initial distribution of molecules on the lattice is uniformly random. The influence of the lattice size on the results is discussed in the "lattice size" subsection. The simulations were preceded by an equilibration phase and mean densities of substrates were estimated by averaging over a sufficiently long trajectory [26]. In order to calculate the EMRRCs, the number of reactions that fired in a time interval (after cutting off the equilibration phase) was recorded. The number of enzyme molecules was fixed. All other details are given in the captions of figures.

1.5 Basic Model

Recall that the state space is a set of functions ordering to each lattice site the molecule by which the site is occupied, or an empty element. Therefore, for this first model considered, governed by reactions (1.4), the function takes five values for every lattice site - either it is empty, or is occupied by one of four molecule types - a phosphorylated substrate molecule, a dephosphorylated substrate molecule, a kinase, or a phosphatase. The transitions between the states consist of phosphorylation and dephosphorylation reactions (these change the densities of phosphorylated and dephosphorylated substrates) and the diffusion events (these change only the distribution of molecules on the lattice).

Analytical results

Infinite-motility limit

We assume that in the *infinite-motility limit* the probability of finding a given molecule is uniform on the lattice. Thus, at any time the density of enzyme–substrate pairs is given by the product of densities multiplied by the number of potential neighbours, e.g. the kinase–dephosphorylated substrate pair density is equal to $6\rho_{\mathrm{K}}\rho_{\mathrm{S_u}}$. Therefore, the phosphorylation rate is equal to $6c\rho_{\mathrm{K}}\rho_{\mathrm{S_u}}$, which in light of Eq. (1.4) gives $c_{\mathrm{eff}}^{\infty} = 6c$. The limit of infinite motility is compared later with simulations performed for high motilities.

Zero-motility limit

The *zero-motility limit* is a singular limit, since without mixing the whole process is determined by initial positions of enzymes and substrates. For an arbitrarily small motility, however, the system relaxes after a sufficiently long time. The zero-motility limit approximates the behaviour of dense systems, in which diffusion is substantially reduced, but reactions still occur for substrates in a close vicinity of opposing enzymes. Increased density, together with reduced diffusion has been intensively modelled in recent years see [21, 52, 10, 19, 22]. Formation of dense ordered patterns of proteins and other molecules was for example considered in [42] (and references therein).

The analysis of this limit starts with the calculation of the steady-state densities of phosphorylated and dephosphorylated substrates, $\rho_{\mathrm{S_p}}$ and $\rho_{\mathrm{S_u}}$:

$$\rho_{\mathrm{S_p}} = p^{+} \cdot \rho_{\mathrm{S}}, \qquad \rho_{\mathrm{S_u}} = \rho_{\mathrm{S}} - \rho_{\mathrm{S_p}}, \tag{1.7}$$

where p^{+} is the probability that a substrate molecule is in the phosphorylated state.

When the motility is zero, the probability that a given substrate molecule is phosphorylated depends solely on the number of neighbouring kinases, i, and the number of neigh-

bouring phosphatases, j, and is equal to

$$p_{ij}^{+} = \frac{ic}{ic + jd}. \tag{1.8}$$

The probability of having exactly i kinase and j phosphatase neighbours is

$$p_{ij} = \binom{6}{i} \rho_{\mathrm{K}}^{i} \binom{6-i}{j} \rho_{\mathrm{P}}^{j} \left(1 - \rho_{\mathrm{K}} - \rho_{\mathrm{P}}\right)^{6-i-j}, \quad i, j \in \{0, 1, \ldots, 6\}, \quad 1 \leq i + j \leq 6. \tag{1.9}$$

Eq. (1.9) is exact only on infinite domains with infinite number of kinases and phosphatases, however it serves as a good approximation when the number of enzymes of each type is much larger than one.

The density of phosphorylated substrates, $\rho_{\mathrm{S_p}}$, can be calculated in the following way:

$$\begin{aligned}
\rho_{\mathrm{S_p}} =& \rho_{\mathrm{S_p}} p_{0,0} + \\
&+ \rho_{\mathrm{S}} \left(p_{1,0} + p_{1,1} \frac{c}{c+d} + \ldots + p_{1,5} \frac{c}{c+5d} + \ldots + p_{5,0} + p_{5,1} \frac{5c}{5c+d} + p_{6,0} \right) \\
=& \rho_{\mathrm{S_p}} p_{0,0} + \rho_{\mathrm{S}} \sum_{1 \leq i+j \leq 6} p_{ij} \frac{ic}{ic + jd} \\
=& \rho_{\mathrm{S_p}} p_{0,0} + \rho_{\mathrm{S}} \sum_{1 \leq i+j \leq 6} p_{ij} p_{ij}^{+} \\
&\Rightarrow \\
&\rho_{\mathrm{S_p}} = \rho_{\mathrm{S}} \frac{\sum\limits_{1 \leq i+j \leq 6} p_{ij} p_{ij}^{+}}{1 - p_{0,0}}.
\end{aligned} \tag{1.10}$$

In this way we obtained the formula for p^{+}, as it is equal to $\rho_{\mathrm{S_p}}/\rho_{\mathrm{S}}$ and $1 - p_{0,0} = \sum\limits_{1 \leq i+j \leq 6} p_{ij}$:

$$p^{+} = \sum_{1 \leq i+j \leq 6} p_{ij} p_{ij}^{+} \Big/ \sum_{1 \leq i+j \leq 6} p_{ij}. \tag{1.11}$$

Steady state EMRRCs can be now calculated. Note that in the zero-motility limit reactions occur only for the substrate molecules which have neighbours of different types (i.e., at least one kinase and one phosphatase). Again, the probability that the substrate which has i neighbouring kinases and j neighbouring phosphatases is dephosphorylated is $jd/(ic + jd)$. Thus the rate at which phosphorylation reaction occurs is $jd/(ic + jd) \cdot ic$. In the stationary state, as there is no change in the fraction of phosphorylated and dephosphorylated substrates, the rates at which phosphorylation and dephosphorylation reactions occur must be the same. Therefore, the number of phosphorylation/dephosphorylation reactions per reactor volume per time are equal to $\rho_{\mathrm{S}} \sum_{i,j \geq 1, i+j \leq 6} p_{ij} (ic \cdot jd)/(ic + jd)$ and, correspondingly,

the effective phosphorylation and dephosphorylation rate constants are equal to

$$c^0_{\text{eff}} = \frac{\rho_{\text{S}}}{\rho_{\text{K}}\rho_{\text{S}_\text{u}}} \sum_{\substack{i,j\geq 1 \\ i+j\leq 6}} p_{ij} \frac{ic \cdot jd}{ic + jd}, \tag{1.12a}$$

$$d^0_{\text{eff}} = \frac{\rho_{\text{S}}}{\rho_{\text{P}}\rho_{\text{S}_\text{p}}} \sum_{\substack{i,j\geq 1 \\ i+j\leq 6}} p_{ij} \frac{ic \cdot jd}{ic + jd}, \tag{1.12b}$$

where $\rho_{\text{S}}/\rho_{\text{S}_\text{u}} = 1/(1-p^+)$ and $\rho_{\text{S}}/\rho_{\text{S}_\text{p}} = 1/p^+$, with p^+ given by Eq. (1.11).

Finite, non-zero motility

Two extreme cases of zero and infinite motility have been analyzed. In the infinite motility limit, also known as the *reaction-controlled limit*, the effective macroscopic reaction rate constants are proportional to the microscopic reaction propensities (for molecules in contact). In this limit, since $m \gg c$ and $m \gg d$, the probability that an enzyme reacts with a substrate at a single encounter is negligibly small and proportional to the microscopic rates c and d.

The small motility limit arises when the microscopic reaction rates c and d are big compared to motility. Processes characterized by low motility and large reaction propensities are called *diffusion-limited*. For such processes the probability that an allowed reaction fires at every collision of molecules is close to 1. Therefore, for such processes, EMRRCs are proportional to the collision frequency. And the collision frequency is proportional to the motility m. Our situation is more complex, since even in the limit of zero motility the reaction rates are nonzero, as shown in the previous section. Accordingly, one could expect the following formula for macroscopic reaction rate:

$$\tfrac{\text{d}}{\text{d}t}\rho_{\text{S}_\text{p}} = (\lambda m + c^0_{\text{eff}})\,\rho_{\text{K}}\,\rho_{\text{S}_\text{u}} - (\lambda m + d^0_{\text{eff}})\,\rho_{\text{P}}\,\rho_{\text{S}_\text{p}}, \tag{1.13}$$

where λ is some coefficient. In fact, the considered case is even more complicated, since, especially at low enzyme densities, the spatial distribution of the phosphorylated and dephosphorylated substrates is nonuniform. That is, the phosphorylated substrate molecules are more likely to be present in the vicinity of a kinase, while the dephosphorylated substrate molecules - in the vicinity of a phosphatase. As a result, even in the symmetric case of $c = d$ and $\rho_{\text{K}} = \rho_{\text{P}}$, in which the overall probability that a substrate is phosphorylated is $\frac{1}{2}$, kinase molecules collide much more often with phosphorylated substrates, which reduces the effective phosphorylation rate. Intuitively, this effect increases with decreasing density of enzymes and causes that each phosphatase molecule is surrounded by a cloud of dephosphorylated substrates and each kinase molecule by a cloud of phosphorylated substrates.

Unfortunately, in the general case of finite motility, EMRRCs are controlled simultaneously by the motility, both contact reaction propensities, and densities of both enzymes and therefore analytical determination of these rates is a challenging problem. It is only for *Mul-*

tiple lattice occupancy model that we were able, after a small modification of the considered Markov Chain, to go further and provide some analytically derived approximations.

Numerical results

Steady state dependence on enzyme density and motility

In this section we analyze numerically the dependence of the steady-state density of phosphorylated and dephosporylated substrates and EMRRCs on motility and densities of the opposing enzymes.

Recall that in the infinite-motility limit the effective macroscopic phosphorylation and dephosphorylation rate constants are: $c_{\text{eff}}^{\infty} = 6c$ and $d_{\text{eff}}^{\infty} = 6d$, and correspondingly (due to Eq. (1.5)) the density of phosphorylated substrates is

$$\rho_{\text{S}_\text{p}} = \frac{c\rho_{\text{K}}}{c\rho_{\text{K}} + d\rho_{\text{P}}}\rho_{\text{S}}. \tag{1.14}$$

To keep the steady-state densities of phosphorylated and dephosphorylated substrates equal to $\frac{1}{2}$ in the limit of the infinite motility, we keep $c\rho_{\text{K}} = \text{const}$ and $d\rho_{\text{P}} = \text{const}$, that is, we set $c = 1/6\rho_{\text{K}}$ and $d = 1/6\rho_{\text{P}}$.

The first result is that for finite motilities the phosphorylated substrate fraction increases with $\rho_{\text{K}}/\rho_{\text{P}}$ (in the analysis we keep $\rho_{\text{K}} = 0.1$ and vary ρ_{P}), and we show that the smaller the motility is, the more pronounced this effect is, see Fig. 1.1(a).

The dashed line for $m = 0$ tends to 1 with $\rho_{\text{K}}/\rho_{\text{P}}$ tending to infinity. For low motility, $m = 1$, the numerically estimated ρ_{S_p} matches closely the *zero-motility limit*. Similarly, for large motilities, ρ_{S_p} is close to the *infinite-motility limit*. Because of the symmetry, for $\rho_{\text{K}} = \rho_{\text{P}}$ the phosphorylated substrate fraction is equal to $\frac{1}{2}$ for all motilities.

In Fig. 1.1(b) we show that when kinases are more abundant than phosphatases, but at the same time have much lower catalytic activity, the dependence of $\rho_{\text{S}_\text{p}}/\rho_{\text{S}}$ on motility is strongly pronounced. At low motilities, substrates remain mostly in the phosphorylated state, $\rho_{\text{S}_\text{p}}/\rho_{\text{S}} \approx 0.9$, while at high motilities they are mostly dephosphorylated, $\rho_{\text{S}_\text{p}}/\rho_{\text{S}} \approx 0.1$. The above shows that, generically, in the regime of low motilities (*diffusion-limited*) it is the density of enzymes that decides about the state of the system and for large motilities (*reaction-controlled* limit) crucial is the product of the microscopic reaction rates and densities.

In Fig. 1.1(c) we show that the density of phosphorylated substrate can either decrease or increase with motility depending on the enzyme densities ratio. For a fixed density of kinases ($\rho_{\text{K}} = 0.1$) we analyze the dependence of ρ_{S_p} on motility for four values of phosphatase densities, as well as for the FOD model. Since, as in Fig. 1.1(a), dephosphorylation microscopic rate is set $d = 1/6\rho_{\text{P}}$, for increasing motility, $\rho_{\text{S}_\text{p}}/\rho_{\text{S}}$ tends to $\frac{1}{2}$, regardless of the phosphatase density. However, for small motilities $\rho_{\text{S}_\text{p}}/\rho_{\text{S}}$ depends strongly on the phosphatase density, and in general is different than for the FOD model. Only for a very

high density ($\rho_P = 0.3$), the fraction ρ_{S_p}/ρ_S closely matches the FOD model prediction with $d_0 = 1$. This is due to the fact that for $\rho_P = 0.3$ the probability that a given substrate molecule is in contact with at least one phosphatase is high (equal to $1 - (0.7)^6 = 0.88$) and therefore the dephosphorylation is effectively of first order. This demonstrates that the FOD model cannot serve as a good approximation across a broad range of motilities.

A descriptive explanation of the results shown in Fig. 1.1 is as follows: for a decreased phosphatase density (compensated by a proportionally increased dephosphorylation rate d), phosphatases are surrounded by dephosphorylated substrates and therefore the effective dephosphorylation rate decreases. Intuitively, this effect becomes stronger for low motilities, for which substrates have a higher chance to be dephosphorylated after a single encounter with a phosphatase and vanishes in the limit of infinite motility, when the probability that a substrate molecule is in the phosphorylated state does not depend on its position.

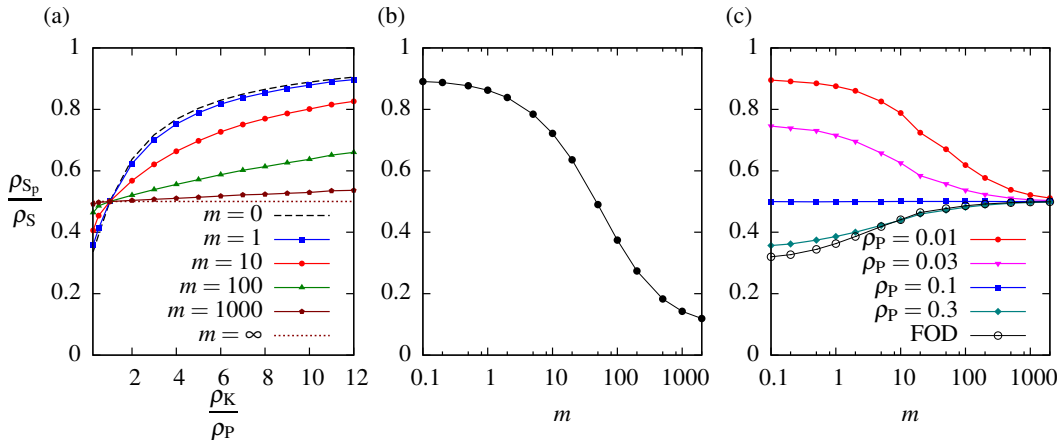

Figure 1.1: (a) Fractional density of phosphorylated substrates, ρ_{S_p}/ρ_S, as a function of the enzymes density ratio for different values of motility, m. Analytically computed limits of zero and infinite motility are marked with dashed and dotted lines. Parameters are: $\rho_S = 0.3$, $\rho_K = 0.1$, $c = 1/6\rho_K$, $d = 1/6\rho_P$. In this series of simulations, the density of kinases was kept constant, while the density of phosphatases was varied from $\rho_P = \rho_K/0.25 = 0.4$ to $\rho_P = \rho_K/12 \approx 0.008$. By setting $d = 1/6\rho_P$, the change of phosphatases density was compensated by the proportional change of the microscopic dephosphorylation rate. (b) Fractional density of phosphorylated substrates as a function of m, in the case when the more abundant enzyme (kinase) has much lower catalytic activity. Simulations were performed for $\rho_S = 0.3$, $\rho_K = 0.1$, $\rho_P = 0.01$, $c = 1$, $d = 100$. (c) Fractional density of phosphorylated substrate as a function of m for different values of phosphatase density ρ_P as well as for the*first-order dephosphorylation* model marked as FOD, with $d_0 = 1$. Simulations were performed for $\rho_S = 0.2$, $\rho_K = 0.1$, $c = 1/6\rho_K$, $d = 1/6\rho_P$.

Effective macroscopic reaction rate constants – EMRRCs

In here I provide some numerical results for EMRRCs estimate. As argued before, c_{eff} can be estimated according to Eq. (1.6). In Fig. 1.2 we show $c_{\text{eff}}/c_{\text{eff}}^{\infty}$ for three values of dephosphorylation rate d, as well as for the FOD model with $d_0 = 1$.

Effective macroscopic phosphorylation rate, c_{eff}, increases with reagents' motility and this effect is more visible for small dephosphorylation reaction rate d. This shows that the phosphorylation kinetics is strongly coupled with the dephosphorylation kinetics and therefore the effective macroscopic phosphorylation and dephosphorylation reaction rates cannot be estimated separately. Figure 1.2 shows that c_{eff} is a function of ρ_{K}, ρ_{P}, c, d, and m. The dependence of c_{eff} on motility is the strongest at the smallest considered enzyme densities, $\rho_{\text{K}} = \rho_{\text{P}} = 0.01$, see Fig. 1.2(c), and the weakest for the highest considered densities, $\rho_{\text{K}} = \rho_{\text{P}} = 0.2$, see Fig. 1.2(a), where $c_{\text{eff}}^{0}/c_{\text{eff}}^{\infty}$ is large. This, consistently with Fig. 1.1, is due to the fact that at high enzyme densities, substrates are constantly in contact with both kinases and phosphatases, and thus the phosphorylation and dephosphorylation reactions can occur almost independently of the diffusion. As shown for $\rho_{\text{K}} = \rho_{\text{P}} = 0.2$ and $\rho_{\text{K}} = \rho_{\text{P}} = 0.05$, Fig. 1.2(a,b), numerically estimated c_{eff} for $m = 0.1$ matches well the analytically calculated limit of c_{eff}^{0}; for $\rho_{\text{K}} = \rho_{\text{P}} = 0.01$, Fig. 1.2(c), the agreement is worse since the convergence of $c_{\text{eff}}(m)$ to c_{eff}^{0} is slower.

Lets analyze these effects in the limit when phosphorylation is a diffusion-driven process. As discussed above, such a limit can be achieved when diffusion-independent reactions are very infrequent compared to those driven by diffusion, i.e., when:

$$c_{\text{eff}}^{0} \ll \lambda m, \quad d_{\text{eff}}^{0} \ll \lambda m. \tag{1.15}$$

Simultaneously, the microscopic contact reaction propensities, c and d, should be much larger than motility, so that the probability of a reaction firing at a collision is close to 1,

$$c \gg m, \quad d \gg m. \tag{1.16}$$

These conditions are difficult to satisfy in numerical simulations, therefore to estimate the diffusion-limited contribution, λm, we subtracted the analytically calculated zero-motility rate constant c_{eff}^{0} from the numerically estimated c_{eff}. We assume here high reaction propensities, $c = d = 1000$, and consider motilities $m \in [0, 1000]$ and enzyme densities $\rho_{\text{E}} \in [0.0001, 0.1]$. The EMRRC is estimated, as previously, from long-run numerical simulations on the 100×100 lattice, based on Eq. (1.6).

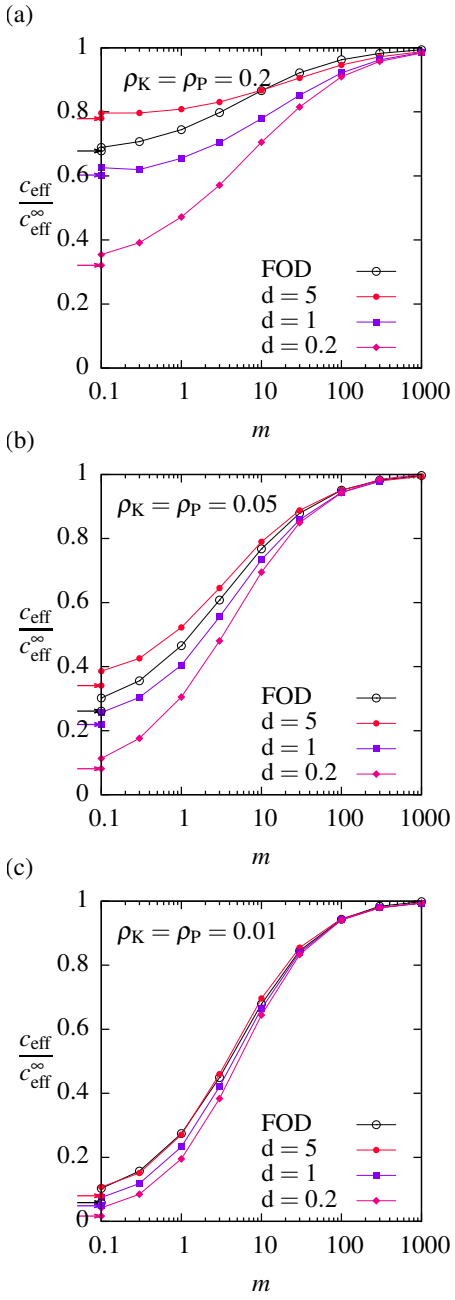

Figure 1.2: Scaled effective macroscopic phosphorylation rate constant $c_{\text{eff}}/c_{\text{eff}}^{\infty}$ as a function of motility m. Densities of enzymes are: $\rho_K = \rho_P = 0.2$ in (a), $\rho_K = \rho_P = 0.05$ in (b), and $\rho_K = \rho_P = 0.01$ in (c). *First order dephosphorylation* model marked as FOD, with $d_0 = 6\rho_P$, which corresponds to $d = 1$ in the basic model. Analytically calculated c_{eff}^{0} are marked by respective arrows next to the vertical axis. For all panels $\rho_S = 0.3$, $c = 1$.

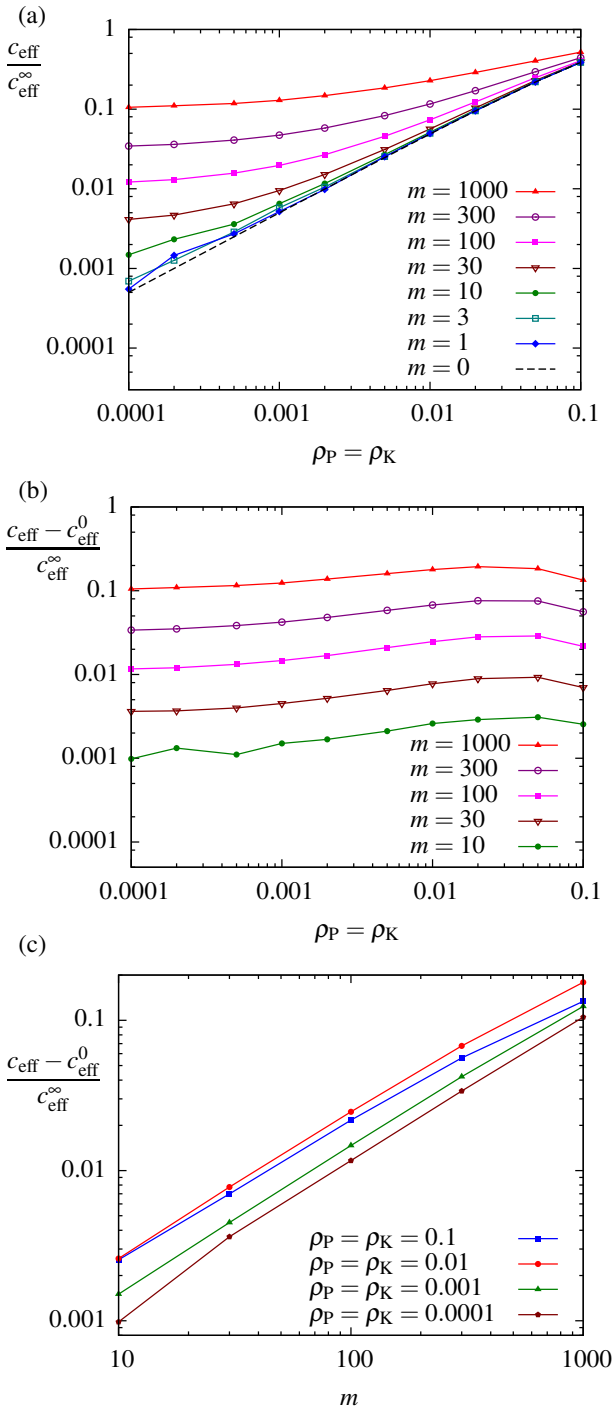

Figure 1.3: (a) Scaled effective macroscopic phosphorylation rate constant $c_{\text{eff}}/c_{\text{eff}}^{\infty}$ as a function of enzyme density $\rho_{\text{K}} = \rho_{\text{P}}$. (b) Scaled effective macroscopic phosphorylation rate constant with subtracted zero-motility contribution: $(c_{\text{eff}} - c_{\text{eff}}^{0})/c_{\text{eff}}^{\infty}$ with respect to enzyme density. (c) $(c_{\text{eff}} - c_{\text{eff}}^{0})/c_{\text{eff}}^{\infty}$ with respect to motility. For all panels $c = d = 1000$.

First we investigate the symmetric case of $\rho_K = \rho_P =: \rho_E$. In Fig. 1.3(a) we show the dependence of $c_{\text{eff}}/c_{\text{eff}}^{\infty}$ on enzyme densities in a log–log scale for seven values of motility. The numerical predictions for small motilities, $m = 1$ and $m = 3$, lie close to the theoretical prediction of the *zero-motility limit* (dashed line). It shows that for relatively small motilities and large enzyme densities the zero-motility contribution is a substantial part of the overall effective rate. The theoretically predicted c_{eff}^{0} is the lower bound for the effective rate coefficient. The zero-motility contribution is proportional to the enzyme density and thus for intermediate motilities, $m \in \{10, 30\}$, it becomes dominant as enzyme density increases.

In order to eliminate the zero-motility contribution from the effective rate coefficient, we show $(c_{\text{eff}} - c_{\text{eff}}^{0})/c_{\text{eff}}^{\infty}$ with respect to enzyme densities (Fig. 1.3(b)) and with respect to motility (Fig. 1.3(c)). In light of Eq. (1.13) we would expect $c_{\text{eff}} - c_{\text{eff}}^{0} = \lambda m$ and therefore $(c_{\text{eff}} - c_{\text{eff}}^{0})/c_{\text{eff}}^{\infty}$ to be proportional to m for fixed densities of enzymes, which is confirmed in Fig. 1.3(c). The average of gradients of lines on the log–log plot is equal to 0.99. We therefore numerically confirmed our heuristic prediction that in the small motility limit:

$$c_{\text{eff}} = c_{\text{eff}}^{0} + \lambda(\rho_E)m. \tag{1.17}$$

Figure 1.3(b) confirms that the coefficient λ decreases (weakly) with decreasing enzyme density. This dependence follows from the fact that at low enzyme densities, enzymes are surrounded by clouds of converted substrates.

Lattice size

So far my "taking into account space" in the modelling of the presented reaction network was focused on considering diffusivity and its interplay with the reaction rates. One could also ask a very natural question about the impact of, simply, the size of the reactor, in which the molecules react. We did perform an analysis of the influence of the lattice size on the estimated EMRRCs (see Fig. 1.4). The simulations were performed on lattices of sizes 300×300, 100×100, 30×30, and 10×10. For each lattice size and each parameter set (corresponding to the parameters chosen for Fig. 1.2(b), we performed 10 independent simulations with simulation times $t = 10^3$, $t = 9 \times 10^3$, $t = 100 \times 10^3$, $t = 900 \times 10^3$, i.e., inversely proportional to the lattice size, which assured that more than 5×10^4 reactions fired in each simulation. Each simulation was preceded by an equilibration phase lasting for 1000. We calculated the scaled effective macroscopic phosphorylation rate constant $c_{\text{eff}}/c_{\text{eff}}^{\infty}$ independently for each simulation, and then, based on the set of ten simulations (for each lattice size and each parameter set), we calculated the mean value of $c_{\text{eff}}/c_{\text{eff}}^{\infty}$ and the error of the mean. In each case the error of the mean was found smaller than 10^{-3}. In conclusion, we found that for assumed densities of molecules the differences between the 10×10 lattice and the remaining lattices are significant, while the differences between larger latices are of the order of the statistical error. One could expect that the dependence of EMRRCs on the

lattice size can be stronger for systems of smaller molecule densities. In the analyzed system there are 45 phosphatases, 45 kinases, and 300 substrates on the 30×30 lattice.

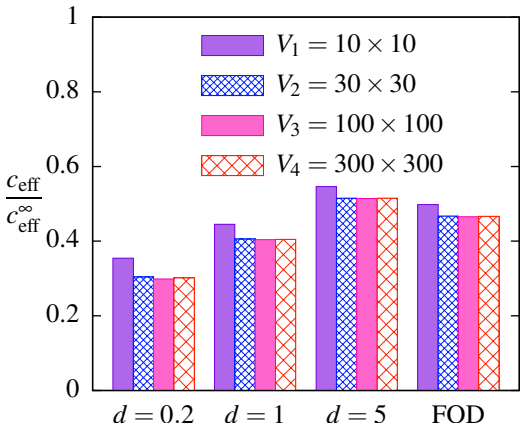

Figure 1.4: Scaled effective macroscopic phosphorylation rate constant $c_{\text{eff}}/c_{\text{eff}}^{\infty}$, estimated in simulations performed on lattices of different sizes. For all simulations $\rho_{\text{S}} = 0.3$, $c = 1$, $m = 1$, $\rho_{\text{K}} = \rho_{\text{P}} = 0.05$. In the *first-order dephosphorylation*, model marked as FOD, $d_0 = 6\rho_{\text{P}}$, which corresponds to $d = 1$ in the basic model. The difference between the 10×10 lattice and the remaining lattices is statistically significant, the differences between larger lattices are of order of the statistical error.

1.6 Molecular crowding model

Another natural question would be about the effect of the presence of additional molecules, so-called crowding agents, which do not react but occupy space and diffuse with a given motility m_{C} (not necessarily equal to substrate and enzyme motility, m). We also looked into such model variant and analyzed how the densities of active substrates in the stationary state change due to the presence of crowding agents.

In this model the state space of the Markov Chain is slightly different than previously - there is an additional type of molecule that can occupy the lattice site and since the crowding molecules are not reagent, there is just one additional element in the set of transitions - the diffusion of crowders.

Some analytical considerations

As could be suspected, the presence of crowding agents leads to a decrease of effective motility of reacting molecules. We tried to quantify this effect and performed the following reasoning: the macroscopic diffusion coefficient, D, of a single tracer molecule having motility m depends on the total density of the crowding molecules ρ_{C} (i.e., the fraction of lattice

sites occupied by molecules), their motility $m_{\mathrm{C}} = m/\gamma$, and the lattice constant ℓ:

$$D = f(\rho_{\mathrm{C}}, \gamma)(1-\rho_{\mathrm{C}})\ell^2 m/4, \tag{1.18}$$

where f is the correlation function that can be approximated by the following formula [54, 2]:

$$f(\rho_{\mathrm{C}}, \gamma) = \\ = \frac{\{[(1-\gamma)(1-\rho_{\mathrm{C}})f_0 + \rho_{\mathrm{C}}]^2 + 4\gamma(1-\rho_{\mathrm{C}})f_0^2\}^{1/2} - [(1-\gamma)(1-\rho_{\mathrm{C}})f_0 + \rho_{\mathrm{C}}]}{2\gamma(1-\rho_{\mathrm{C}})f_0}, \tag{1.19}$$

where

$$f_0 = (1-\alpha)/[1+\alpha(2\gamma-1)]. \tag{1.20}$$

The coefficient α depends on the lattice type; for triangular lattice (considered here) $\alpha = 0.282$, for square lattice $\alpha = 1 - 2/\pi$ and for honeycomb (or hexagonal) lattice $\alpha = 1/2$ [13]. The parameter

$$m_{\mathrm{eff}} = f(\rho_{\mathrm{C}}, \gamma)(1-\rho_{\mathrm{C}})m \tag{1.21}$$

will be considered as the effective motility of the tracer molecule in the presence of crowding molecules of density ρ_{C} and motility m_{C}.

The correlation function f satisfies $0 < f < 1$ for $0 < \gamma < \infty$. In the limit of $\gamma \to 0$, i.e., when the crowding molecules move infinitely fast and a tracer molecule does not sense their positions, $f \to 1$; in the limit of $\gamma = \infty$, i.e., when crowding molecules do not move, the expression for f reads:

$$f(\rho_{\mathrm{C}}) = \max\left\{0, \frac{(1-\alpha) - \rho_{\mathrm{C}}(1+\alpha)}{(1-\rho_{\mathrm{C}})(1-\alpha)}\right\}. \tag{1.22}$$

According to the equation above, the diffusion coefficient of a tracer molecule drops to zero when the fractional density of immobile obstacles equals $\rho_{\mathrm{crit}} = (1-\alpha)/(1+\alpha) = 0.56$, which agrees reasonably well with the percolation threshold of $1/2$ for the triangular lattice. In the case most interesting to us, i.e., when all molecules have the same motility ($\gamma = 1$), Eq. (1.18) simplifies to

$$D(\rho_{\mathrm{C}}, 1) = \frac{\sqrt{\rho_{\mathrm{C}}^2 + 4(1-\rho_{\mathrm{C}})\left(\frac{1-\alpha}{1+\alpha}\right)^2} - \rho_{\mathrm{C}}}{2\left(\frac{1-\alpha}{1+\alpha}\right)}\ell^2 m/4. \tag{1.23}$$

Numerical results

The approximate Eq. (1.18) agrees well with our simulation results presented in Fig. 1.5(a). In these simulations we estimated the effective motility of the tracer molecule $m_{\text{eff}} := \langle \text{Dist}^2 \rangle / \Delta t$, based on the mean square distance $\langle \text{Dist}^2 \rangle$ covered by the tracer molecule in time Δt. To obtain reasonable statistics at a modest computational cost we performed simulations in which the number of tracer molecules was larger than one, but always smaller than 1% of the number of crowding molecules. Finally, in order to analyze the influence of crowding molecules with a given motility on the effective motility of reacting molecules, we performed simulations in which the density of reacting molecules was 30%, while different densities and motilities of crowding molecules were considered, see Fig. 1.5(b). These results are used to interpret the effect of molecular crowding on the steady state of the reacting system: the reduction of the effective substrate motility either increases the fraction of phosphorylated substrates in the stationary state, provided that $\rho_{\text{K}} > \rho_{\text{P}}$, or, because of the symmetry of the model, decreases this fraction for $\rho_{\text{P}} > \rho_{\text{K}}$. As shown in Fig. 1.6(b), the effect of crowding agents can be almost fully reproduced by the appropriate scaling of reagents motility:

$$\tilde{m}_{\text{R}} := m_{\text{R}} \frac{m_{\text{eff}}(\rho_{\text{R}}, \rho_{\text{C}}, m_{\text{R}}, m_{\text{C}})}{m_{\text{eff}}(\rho_{\text{R}}, m_{\text{R}})}, \tag{1.24}$$

where $\rho_{\text{R}} = \rho_{\text{S}} + \rho_{\text{K}} + \rho_{\text{P}}$ is the fractional density of all reacting molecules assumed to have the same motility m_{R}. In the numerator of Eq. (1.24) there is the effective motility of reacting molecules of density ρ_{R} and motility m_{R} in the presence of crowding agents of density ρ_{C} and motility m_{C}, estimated in numerical simulations and given in Fig. 1.5(b). In the denominator of Eq. (1.24) there is the effective motility of reacting molecules of density ρ_{R} and motility m_{R} in the absence of additional molecules, given by the approximate Eq. (1.21).

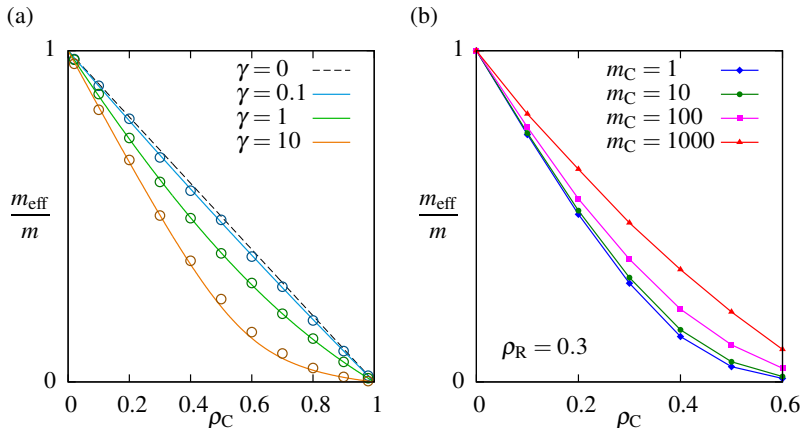

Figure 1.5: (plotted by Marek Kochańczyk) Scaled effective motility $m_{\rm eff}/m$ as a function of density of crowding molecules $\rho_{\rm C}$, and motility $m_{\rm C} = m/\gamma$. (a) Effective motility of a tracer molecule in the presence of crowding molecules. Lines correspond to the theoretical result given by Eq. (1.18), circles mark results of the corresponding numerical simulations. (b) Scaled effective motility $m_{\rm eff}/m$ of reacting molecules with fractional density $\rho_{\rm R} = 0.3$ and motility $m = 1000$ in the presence of crowders. This result is used in simulations shown in Fig. 1.6(b).

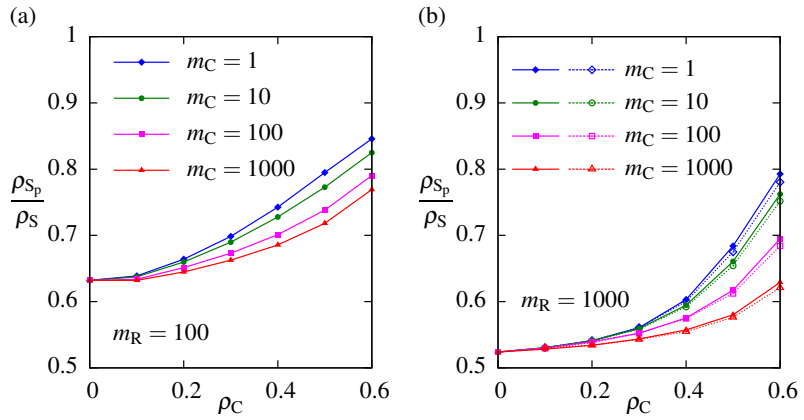

Figure 1.6: Phosphorylated substrate fractional density with respect to the density of crowding agents $\rho_{\rm C}$. Reagents motility $m_{\rm R} = 100$ in (a), $m_{\rm R} = 1000$ in (b), for four values of crowding agents motility $m_{\rm C}$. Dashed lines refer to the simulations without crowding agents, with scaled reagents motility $\tilde{m}$, see Eq. (1.24). Other parameters are $\rho_{\rm S} = 0.2$, $\rho_{\rm K} = 0.09$, $\rho_{\rm P} = 0.01$, $c = 1/6\rho_{\rm K}$ and $d = 1/6\rho_{\rm P}$.

1.7 Transient enzyme–substrate complexes model

The two models presented thus far do not account explicitly for the formation of the enzymatic encounter complex: both the phosphorylation and dephosphorylation are considered to be single-step reactions. In reality, these reactions are multi-step processes (enzyme–substrate binding, reaction, and enzyme–product dissociation). This simplification does not significantly affect our key findings, at least when the enzyme sequestration is weak [50]. However, we decided to show this part of the analysis, because in reality such processes as the phosphorylation–dephosphorylation cycle involve at least three steps and require formation of a transient enzyme–substrate complex. It was therefore important to verify whether the analyzed effects are preserved when a more accurate description is executed. In the more detailed model, reactions (1.1) are replaced by

$$\mathrm{K} + \mathrm{S_u} \underset{c_2}{\overset{c_1}{\rightleftharpoons}} [\mathrm{K}\cdot\mathrm{S_u}] \quad \xrightarrow{c_3} \quad [\mathrm{K}\cdot\mathrm{S_p}] \xrightarrow{c_4} \mathrm{K} + \mathrm{S_p}, \tag{1.25a}$$

$$\mathrm{P} + \mathrm{S_p} \underset{d_2}{\overset{d_1}{\rightleftharpoons}} [\mathrm{P}\cdot\mathrm{S_p}] \quad \xrightarrow{d_3} \quad [\mathrm{P}\cdot\mathrm{S_u}] \xrightarrow{d_4} \mathrm{P} + \mathrm{S_u}. \tag{1.25b}$$

The state space of the Markov Chain handled here is a bit different than in the Basic Model - we now distinguish between a state in which two molecules occupying adjacent lattice sites are forming a complex (then a reaction might occur or they can dissociate) or not (then they can associate or drift apart). Accordingly, the set of possible transitions between the states is larger - there are additional reactions that occur - association and dissociation of two types of complexes. This accounts for four new reactions.
We will consider two sets of reaction rate coefficients corresponding to the short or longer enzyme–substrate binding, implying respectively either weak or stronger, but still moderate enzyme sequestration. The coefficients for the two cases are:

weak sequestration:

$$c_1 = 2c, \quad c_2 = 10c, \quad c_3 = 10c, \quad c_4 = 100c, \tag{1.26a}$$

$$d_1 = 2d, \quad d_2 = 10d, \quad d_3 = 10d, \quad d_4 = 100d; \tag{1.26b}$$

moderate sequestration:

$$c_1 = 10c, \quad c_2 = 9c, \quad c_3 = c, \quad c_4 = 100c, \tag{1.27a}$$

$$d_1 = 10d, \quad d_2 = 9d, \quad d_3 = d, \quad d_4 = 100d. \tag{1.27b}$$

Numerical results

For these two sets of coefficients a substrate being initially in contact with an enzyme molecule is modified with almost the same probability as in the original model. For this model variant we performed an analysis analogous to that shown in Fig. 1.1 (see Fig. 1.7). In the case of weak sequestration, we obtained the quantitatively similar dependence of fraction of phosphorylated substrate on enzyme density and on motility (Fig. 1.7(a,b)) as in the original model. For stronger sequestration, for which the fraction of sequestered kinases exceeds 60% (for large motilities), the agreement with the original model (Fig. 1.7(b)) is only qualitative. Importantly, the fraction of sequestered enzymes and substrates significantly grows with motility. This is due to the fact that the increase of motility implies more enzyme–substrate encounters, and therefore increases their binding rate, not influencing the dissociation rate.

Overall, with the analysis of the above model variant we verified that the reported dependence of steady state on motility is independent of the details of the phosphorylation and dephosphorylation processes, as long as the fractions of sequestered enzymes and substrate are small, and results from the presence of opposing enzymes in the reaction network. However, for stronger enzyme–substrate binding, the fraction of sequestered reactants is higher (and dependent on their motility), and therefore the quantitative dependence of the phosphorylated substrate fraction on motility can be different.

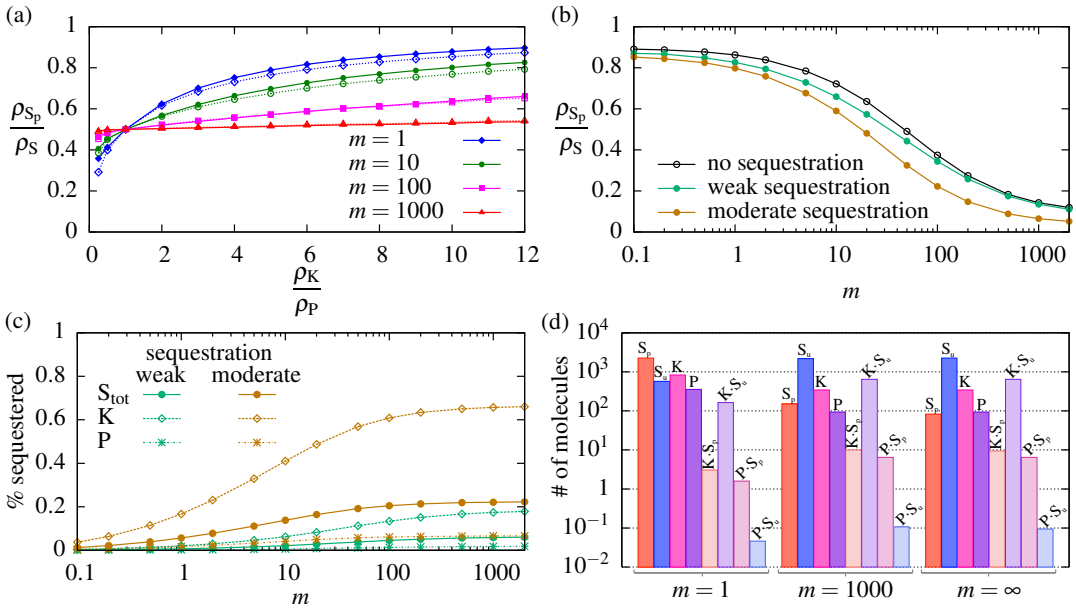

Figure 1.7: (plotted by Marek Kochanczyk) (a) Fractional density of phosphorylated substrates as a function of the enzyme ratio for different values of motility m. We compare the model variant in which the formation of a transient enzyme–substrate complex is explicitly included (the case of weak enzyme sequestration, Eqs. (1.26); dotted lines) with the original model prediction shown in Fig. 1.1(a) (solid lines). The parameters used in the simulations of the basic (original) model: $\rho_S = 0.3$, $\rho_K = 0.1$, $c = 1/6\rho_K$, $d = 1/6\rho_P$; the parameters for the model variant considered are defined by Eqs. (1.26). In the calculation of the phosphorylated substrate fraction only free (unbound) substrates are considered. (b) Fractional density of phosphorylated substrates as a function of m. We compare the original model prediction shown in Fig. 1.1(b) (black line) with the model variant in which the formation of a transient enzyme–substrate complex is explicitly included; two cases are considered: weak sequestration, Eqs. (1.26), and moderate enzyme sequestration, Eqs. (1.27). The parameters used in simulations are: $\rho_S = 0.3$, $\rho_K = 0.1$, $\rho_P = 0.01$, $c = 1$, $d = 100$. (c) Fraction of sequestered reactants for the weak and moderate sequestration cases as a function of m in simulations performed for (b). (d) Steady-state densities of all reactants and complexes in the case of moderate sequestration, Eqs. (1.27), for three motilities: $m = 1$, $m = 1000$ and $m = \infty$. Values for finite motility come from simulations performed for (b). Values for infinite motility are given by the steady state of the corresponding system of ODEs.

1.8 *Multiple lattice occupancy* model

The reason why we ventured on yet another model, at the first glance quite similar to the previous ones, is that we were unsatisfied with the amount of analytical results obtained - only the limits of zero and infinite motility were properly embraced by rigorous analytical reasoning. The non-zero, finite motility case was mainly approached by computational methods, but we felt that more analytical approach, which leads to quite precise predictions, is still possible in this regime. However, there had to be a small modification introduced, both in the construction of the Markov Chain and consequently in the numerical implementation - here **we will allow a substrate molecule and an enzyme molecule to enter the same lattice site** (but not two substrate molecules nor two enzyme molecules) and we will require for the reaction to occur that the substrate (in the proper state) and the enzyme are in the same lattice site. This changes the state space of the Markov Chain - the function mapping lattice sites on the set of elements that can enter the lattice site now takes 9 values - empty, four single molecules and four enzyme-substrate pairs. The set of possible transitions remains the same, as there are no additional types of reactions.

Analytical results

The new assumptions imposed on the model change slightly the two analytical limits we computed previously.

Inifinite-motility limit

The concentration of enzyme–substrate pairs is now given by the product of their concentrations: $\rho_{\mathrm{K}}\rho_{\mathrm{S_u}}$, and the phosphatase–phosphorylated substrate pair concentration equals $\rho_{\mathrm{P}}\rho_{\mathrm{S_p}}$. Recall that before in both these quantities there was a prefactor equal to the number of neighbouring sites : 6. Analogously the numbers of phosphorylation and dephosphorylation reactions that fired during a time interval Δt in a reactor of volume V are $c\,\rho_{\mathrm{K}}\rho_{\mathrm{S_u}}V\Delta t$ and $d\,\rho_{\mathrm{P}}\rho_{\mathrm{S_p}}V\Delta t$. Consequently, the EMRRCs in the infinite-motility limit are equal to:

$$c_{\mathrm{eff}}^{\infty} = c, \qquad d_{\mathrm{eff}}^{\infty} = d. \tag{1.28}$$

Zero motility limit

As the substrate and enzyme molecules react only when present in the same lattice site, the reactions cease in this limit completely. Recall that in the previous model, the substrate molecules having both a kinase molecule and a phosphatase molecule at adjacent sites were repeatedly converted between the phosphorylated and the unphosphorylated state, which resulted in (sometimes significant) zero-diffusion contribution to the macroscopic reaction rates coefficients. This is not the case here.

Finite, non-zero motility

We want to derive, as before, formulae for the effective macroscopic reaction rate coefficients, EMRRCs, c_{eff} and d_{eff}, as functions of microscopic reaction rates c and d and the remaining parameters of the model. Again, we restrict our study to the steady-state values. First, we express the steady-state EMRRCs via the *mean first-passage time*, MFPT, in which a substrate molecule after changing its state upon the reaction with a given enzyme reaches an opposite enzyme molecule. Then we estimate this MFPT by the average number of steps, $w(\rho)$, until trapping a random walker in the system of randomly distributed traps with a given concentration, ρ. These are needed for final formulae.

The whole derivation starts with a simple observation that the steady-state fractions of unphosphorylated and phosphorylated substrate, $\rho_{\text{S}_\text{u}}/\rho_{\text{S}}$ and $\rho_{\text{S}_\text{p}}/\rho_{\text{S}}$, can be expressed in terms of the average time intervals during which a substrate molecule remains unphosphorylated, τ_{u}, and phosphorylated, τ_{p}:

$$\frac{\rho_{\text{S}_\text{u}}}{\rho_{\text{S}}} = \frac{\tau_{\text{u}}}{\tau_{\text{u}} + \tau_{\text{p}}}, \qquad \frac{\rho_{\text{S}_\text{p}}}{\rho_{\text{S}}} = \frac{\tau_{\text{p}}}{\tau_{\text{u}} + \tau_{\text{p}}}. \tag{1.29}$$

Now we express c_{eff} and d_{eff} through τ_{u} and τ_{p}:

$$c_{\text{eff}} = \frac{1}{\tau_{\text{u}}\rho_{\text{K}}}, \qquad d_{\text{eff}} = \frac{1}{\tau_{\text{p}}\rho_{\text{P}}}. \tag{1.30}$$

To calculate time intervals τ_{u} and τ_{p} we split them into:

$$\tau_{\text{u}} = \tau_{\text{u}_1} + \tau_{\text{u}_2}, \qquad \tau_{\text{p}} = \tau_{\text{p}_1} + \tau_{\text{p}_2}, \tag{1.31}$$

where τ_{u_1} (τ_{p_1}) is the MFPT, in which a substrate molecule after being modified by a phosphatase (kinase) molecule meets a kinase (phosphatase) molecule for the first time, and τ_{u_2} (τ_{p_2}) is the average time after which a substrate molecule occupying initially the same lattice site as a kinase (phosphatase) molecule becomes phosphorylated (unphosphorylated).

Time intervals τ_{u} and τ_{p} depend on the effective motilities of enzyme and substrate molecules, $\tilde{m}_{\text{E}}$ and $\tilde{m}_{\text{S}}$. As discussed previously, the effective motilities are lower than the nominal motility of all molecules, m, due to molecular crowding, and when $\rho_{\text{E}} \neq \rho_{\text{S}}$, then $\tilde{m}_{\text{E}}$ and $\tilde{m}_{\text{S}}$ differ because enzyme and substrate molecules are crowding agents only for themselves. The effective relative motility of enzyme and substrate molecules is $\tilde{M} = \tilde{m}_{\text{E}} + \tilde{m}_{\text{S}}$. The time between encounters of enzyme and substrate molecules scales inversely with $\tilde{M}$.

Again, because all molecules have the same motility, the expression for $\tilde{m}$ is simply (see Eq. (1.21))

$$\tilde{m}(m,\rho_{\mathrm{C}}) = m\frac{\sqrt{\rho_{\mathrm{C}}^2 + 4(1-\rho_{\mathrm{C}})\left(\frac{1-a}{1+a}\right)^2} - \rho_{\mathrm{C}}}{2\left(\frac{1-a}{1+a}\right)}. \tag{1.32}$$

The modification introduced in this model allows two molecules to enter the same lattice site only if they are an enzyme-substrate pair. So neither two enzyme molecules nor two substrate molecules can enter the same lattice site. This means that enzyme as well as substrate molecules play the role of crowding agents only for themselves. Accordingly, $\tilde{m}_X$ for $X \in \{\mathrm{S},\mathrm{K},\mathrm{P}\}$ is given by:

$$\tilde{m}_X = \begin{cases} \tilde{m}(m_X,\rho_{\mathrm{S}}) & \text{for} \quad X = \mathrm{S}, \\ \tilde{m}(m_X,\rho_{\mathrm{K}}+\rho_{\mathrm{P}}) & \text{for} \quad X \in \{\mathrm{K},\mathrm{P}\}. \end{cases} \tag{1.33}$$

Having this scaling at hand, we start the calculation of time intervals τ_{k_i}, $k \in \mathrm{u},\mathrm{p}$, $i \in \{1,2\}$ by the simplest, τ_{u_2}. When an unphosphorylated substrate molecule and a kinase molecule meet in the same lattice site, two exclusive events are possible: either the substrate molecule gets phosphorylated, or the molecules move apart before the reaction fires. The expected time for which an unphosphorylated substrate molecule and a kinase molecule remain in the same lattice site, τ_{short}, is inversely proportional to the sum of rates of these two events, $\tau_{\mathrm{short}} = 1/(c+\tilde{M})$. With the probability of the phosphorylation event, which is $c/(c+\tilde{M})$, τ_{u_2} will be equal to τ_{short}, and with the probability of the separation event, which is $\tilde{M}/(c+\tilde{M})$, τ_{u_2} will be equal to τ_{long}, which is the expected time for substrate molecule phosphorylation in the case when it moves away from the kinase molecule. Taken together, τ_{u_2} may be expressed as:

$$\tau_{\mathrm{u}_2} = \frac{c}{c+\tilde{M}}\tau_{\mathrm{short}} + \frac{\tilde{M}}{c+\tilde{M}}\tau_{\mathrm{long}}, \tag{1.34}$$

where

$$\tau_{\mathrm{long}} = \tau_{\mathrm{find}} + \tau_{\mathrm{short}} + \tau_{\mathrm{u}_2}. \tag{1.35}$$

Here, τ_{find} is the average time for the substrate molecule to meet a kinase molecule (the same or another) under the condition that it is in a site adjacent to a site occupied by a kinase molecule. When the substrate molecule meets a kinase molecule, the initially considered situation reoccurs and therefore the third term is τ_{u_2}.

To calculate τ_{find} we notice that since the fraction of lattice sites occupied by kinase molecules is equal to ρ_{K}, on average every $1/\rho_{\mathrm{K}}$ step the substrate molecule meets a kinase molecule. This is, when a substrate molecule and a kinase molecule occupy the same lattice site, the expected number of steps after which the substrate molecule meets the same or another kinase molecule is $1/\rho_{\mathrm{K}}$. Therefore, if these two molecules are in adjacent lattice

sites, i.e., when one step toward next meeting has already been done, the expected number of steps is $1/\rho_{\mathrm{K}} - 1$. Thus

$$\tau_{\mathrm{find}} = \frac{1/\rho_{\mathrm{K}} - 1}{\tilde{M}}. \tag{1.36}$$

Finally, Eq. (1.34), Eq. (1.35), and Eq. (1.36) together yield

$$\tau_{\mathrm{u}_2} = \frac{c}{c+\tilde{M}}\frac{1}{c+\tilde{M}} + \frac{\tilde{M}}{c+\tilde{M}}\left(\frac{1/\rho_{\mathrm{K}} - 1}{\tilde{M}} + \frac{1}{c+\tilde{M}} + \tau_{\mathrm{u}_2}\right), \tag{1.37}$$

from which we obtain a simple expression for τ_{u_2} and an analogous expression for τ_{p_2}:

$$\tau_{\mathrm{u}_2} = \frac{1}{c\rho_{\mathrm{K}}}, \qquad \tau_{\mathrm{p}_2} = \frac{1}{d\rho_{\mathrm{P}}}. \tag{1.38}$$

Now the only remaining calculation is that of τ_{u} and τ_{p}. For these we need to estimate τ_{u_1} and τ_{p_1}. These two MFPTs can be expressed as:

$$\tau_{\mathrm{u}_1} = \frac{w(\rho_{\mathrm{P}}, \rho_{\mathrm{K}})}{\tilde{M}}, \qquad \tau_{\mathrm{p}_1} = \frac{w(\rho_{\mathrm{K}}, \rho_{\mathrm{P}})}{\tilde{M}}, \tag{1.39}$$

where $w(\rho_{\mathrm{P}}, \rho_{\mathrm{K}})$ and $w(\rho_{\mathrm{K}}, \rho_{\mathrm{P}})$ are the expected numbers of steps needed for a substrate molecule to reach a kinase and a phosphatase molecule, respectively, after being converted by a phosphatase (kinase) molecule. Eventually, we arrive at the following formulae:

$$c_{\mathrm{eff}} = \frac{1}{(\tau_{\mathrm{u}_1} + \tau_{\mathrm{u}_2})\rho_{\mathrm{K}}} = \left(\frac{1}{c} + \frac{\rho_{\mathrm{K}}\, w(\rho_{\mathrm{P}}, \rho_{\mathrm{K}})}{\tilde{M}}\right)^{-1}, \tag{1.40a}$$

$$d_{\mathrm{eff}} = \frac{1}{(\tau_{\mathrm{p}_1} + \tau_{\mathrm{p}_2})\rho_{\mathrm{P}}} = \left(\frac{1}{d} + \frac{\rho_{\mathrm{P}}\, w(\rho_{\mathrm{K}}, \rho_{\mathrm{P}})}{\tilde{M}}\right)^{-1}. \tag{1.40b}$$

The reader is now ready for further analytical consideration leading to formulae for $w(\rho_{\mathrm{K}}, \rho_{\mathrm{P}})$. The following paragraphs reproduce the reasoning presented in our paper [50] and conducted mostly by Paweł Nałęcz-Jawecki.

Under the assumption that the search for enzyme molecules of an appropriate type starts from a random position, functions $w(\rho_{\mathrm{P}}, \rho_{\mathrm{K}})$ and $w(\rho_{\mathrm{K}}, \rho_{\mathrm{P}})$ can be simplified to

$$w(\rho_{\mathrm{P}}, \rho_{\mathrm{K}}) = w(\rho_{\mathrm{K}}), \qquad w(\rho_{\mathrm{K}}, \rho_{\mathrm{P}}) = w(\rho_{\mathrm{P}}). \tag{1.41}$$

To understand when the above simplifying assumption is valid, let us consider the case when on the lattice there is only one kinase molecule and a large number of phosphatase molecules. In such a case, a substrate molecule phosphorylated by the kinase molecule will be dephosphorylated in its vicinity by one of numerous phosphatase molecules, and therefore the next search for the single kinase molecule will start not from a random position with respect to the kinase molecule but more likely from its vicinity. Thus, in the considered

case, the assumption is not valid for the phosphorylation reaction; however, since there is only one kinase molecule and thus the expected time to phosphorylation is relatively long, the abundant phosphatase molecules change significantly their positions between two dephosphorylation reactions, so that one can assume that the search for a phosphatase molecule starts from a random position with respect to positions of phosphatase molecules.

Lets consider now the system of N different enzyme molecules, E_i, $i = 1, ..., N$, and assume that each enzyme molecule E_i converts substrate molecules to a distinct state S_i with reaction rate q. Let us assume that $N \gg 1$ and let ρ_{E} denote the total concentration of all enzyme molecules. In light of the observation made in the previous paragraph, substrate molecules converted by E_i (i.e., in state S_i) will start their search for the remaining $N-1$ enzyme molecules at a position that can be considered random (with respect to remaining enzyme molecules). Thus, the average time τ for which the substrate molecules will remain in each of the states S_i is (by analogy to Eqs. (1.31), with Eqs. (1.38) and Eqs. (1.39), and since the concentration of $N-1$ enzyme molecules is $\approx \rho_{\mathrm{E}}$)

$$\tau = \frac{1}{q\rho_{\mathrm{E}}} + \frac{w(\rho_{\mathrm{E}})}{\tilde{M}}. \tag{1.42}$$

The number of reactions per substrate molecule per time is equal to $r = 1/\tau$. Let us assume that one part of these enzyme molecules are kinase molecules and the rest are phosphatase molecules, so that $\rho_{\mathrm{K}} + \rho_{\mathrm{P}} = \rho_{\mathrm{E}}$. Therefore, the probability that an unphosphorylated substrate molecule will be converted in the next reaction to the phosphorylated state is $\rho_{\mathrm{K}}/\rho_{\mathrm{E}}$, while with probability $\rho_{\mathrm{P}}/\rho_{\mathrm{E}}$ it will be converted to the other (unphosphorylated) state (such pseudo-conversions are possible because we assumed that each enzyme molecule converts the substrate to a distinct state). The number of real phosphorylation reactions (i.e., conversions from the unphosphorylated to the phosphorylated state) per unphosphorylated substrate molecule is $r_p = r \cdot \rho_{\mathrm{K}}/\rho_{\mathrm{E}}$, and therefore the average time spent by a substrate molecule in the unphosphorylated state is $\tau_{\mathrm{u}} = 1/r_p = \tau \cdot \rho_{\mathrm{E}}/\rho_{\mathrm{K}}$.

From $c_{\mathrm{eff}} = 1/(\tau_{\mathrm{u}}\rho_{\mathrm{K}})$, Eqs. (1.30), we obtain

$$c_{\mathrm{eff}} = \left(\frac{1}{q} + \frac{1}{\tilde{M}}\rho_{\mathrm{E}}\, w(\rho_{\mathrm{E}})\right)^{-1}. \tag{1.43}$$

To derive the above equation we had to assume that all substrate states, S_i, are equiprobable, which requires $c = d = q$. In the case when $c \neq d$ we propose to replace q by c or d, appropriately, which leads to the following approximations for EMRRCs:

$$c_{\mathrm{eff}} = \left[\frac{1}{c} + \frac{1}{\tilde{M}}(\rho_{\mathrm{K}} + \rho_{\mathrm{P}})\, w(\rho_{\mathrm{K}} + \rho_{\mathrm{P}})\right]^{-1}, \tag{1.44a}$$

$$d_{\mathrm{eff}} = \left[\frac{1}{d} + \frac{1}{\tilde{M}}(\rho_{\mathrm{K}} + \rho_{\mathrm{P}})\, w(\rho_{\mathrm{K}} + \rho_{\mathrm{P}})\right]^{-1}, \tag{1.44b}$$

where, recall, $w(\rho)$ is the average number of steps until trapping a random walker in a system of randomly distributed traps of concentration ρ.

The very last thing is now to calculate $w(\rho)$. Here I evoke an old result by Montroll [35], who obtained an analytical asymptotic formula for the average number of steps of a random walker, (for walks on lattices with periodic distributions of traps, of concentration ρ, or, equivalently, on finite lattices of volume $V = 1/\rho$ with periodic boundary conditions, containing a single trap). The approximate formula reads:

$$w_{\mathrm{P}}(1/V) = \alpha V \log V + \beta V + \gamma + \mathcal{O}(1/V), \tag{1.45}$$

where α is constant for a particular lattice structure ($\alpha = 1/\pi$ for a square lattice, $\alpha = \sqrt{3}/(2\pi)$ for a triangular lattice), whereas β and γ depend also on the shape of the reactor. For a triangular lattice and a square-shaped reactor their values are $\beta \approx 0.235$ and $\gamma \approx -0.251$. We can restrict ourselves to two first terms of the right-hand side of Eq. (1.45). However, having found this result, we faced the problem of how to adapt it to the case when traps are distributed randomly. In here it was Paweł's excellent coding skills that saved our analytical effort. He established, by fitting Montroll's formula, coefficients α' and β' in

$$w_{\mathrm{R}}(\rho) = \alpha' \rho^{-1} \log \rho^{-1} + \beta' \rho^{-1}, \tag{1.46}$$

to be $\alpha' = \alpha$ and $\beta' = 1.00$.

At this point we could plug every element of the puzzle (some of the elements are exact, some are approximations) into the formula for EMRRCs. After setting $1/V = \rho_{\mathrm{K}} + \rho_{\mathrm{P}}$ from Eqs. (1.44) we obtained the compact form:

$$c_{\mathrm{eff}} = \left[\frac{1}{c} + \frac{1}{\tilde{M}} \left(\alpha' \log \frac{1}{\rho_{\mathrm{K}} + \rho_{\mathrm{P}}} + \beta' \right) \right]^{-1}, \tag{1.47a}$$

$$d_{\mathrm{eff}} = \left[\frac{1}{d} + \frac{1}{\tilde{M}} \left(\alpha' \log \frac{1}{\rho_{\mathrm{K}} + \rho_{\mathrm{P}}} + \beta' \right) \right]^{-1}, \tag{1.47b}$$

with $\alpha' = \alpha = \sqrt{3}/(2\pi)$, $\beta' = 1.00$.

Some numerical results

As I already mentioned before, as far as analytical results are concerned, they were produced in cooperation with my colleague Paweł Nałęcz-Jawecki. However, he did all the numerical work, as well as plots, so I do not dare to present it here, but I am most willing, with Paweł's consent, to present for an avid reader just one plot of comparison how well these analytical reasoning works, or at least how better it is, compared to the previous study. The assumption in the former model [40] implies that there are on average six times more enzyme–substrate pairs than in the current model (wherefrom in the infinite-diffusion limit,

the EMRRC was equal $c_{\mathrm{eff}}^{\infty} = 6c$ and not c as in the current model). Therefore, to compare EMRRCs between the Basic Model and the *Mutliple lattice occupancy* model, (see Fig. 1.8) we used a respective $c_{\mathrm{eff}}^{\infty}$ value to normalize c_{eff} for both models. The "effective" distances between enzymes were in the previous approach "shorter", so that for finite motilities the EMRRCs were greater compared to those calculated in the current model, see Fig. 1.8 (a) and (b). As mentioned before, in the zero-motility limit, the substrate molecules could have two antagonistic enzymes in their reaction volumes (consisting of six neighbouring sites), this led to nonzero c_{eff}, particularly when the probability of having two antagonistic enzymes is large, i.e., for dense systems, Fig. 1.8 (a). Fig. 1.8 (c) shows the discrepancy between steady-state phosphorylated substrate fraction predicted by the two models, which arises when $c \neq d$ (for $c = d$, both models predicts that $\rho_{\mathrm{S_p}}/\rho_{\mathrm{S}} = \rho_{\mathrm{K}}/(\rho_{\mathrm{K}} + \rho_{\mathrm{P}})$ independently of motility).

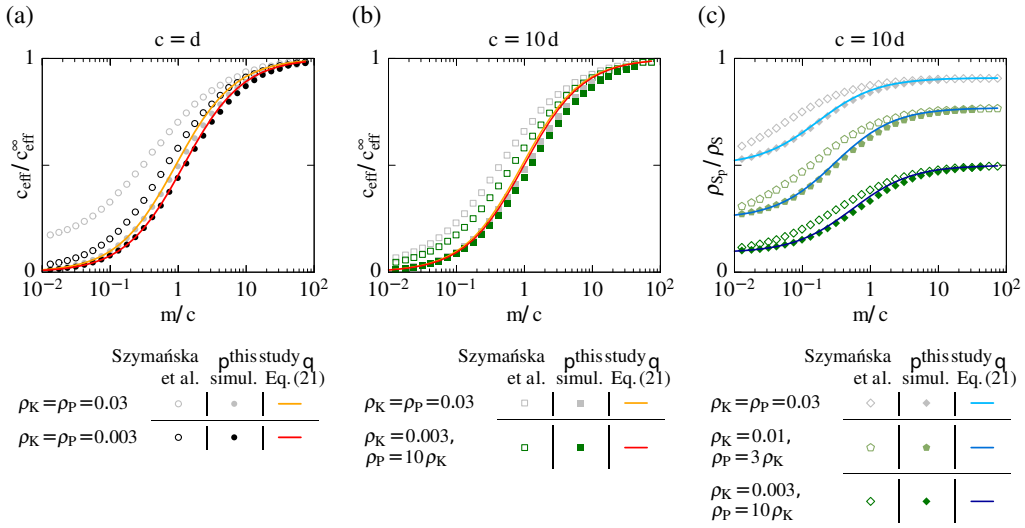

Figure 1.8: Comparison between the *Mutliple lattice occupancy* model, which approximates the effective reaction rate constants with Mean First Passage Time considerations, and the Basic Model presented in an earlier section. Solid lines correspond to theoretical predictions of the new model. (a) Normalized effective phosphorylation rate constant, $c_{\mathrm{eff}}/c_{\mathrm{eff}}^{\infty}$, in the fully symmetric case: $c = d$, $\rho_{\mathrm{K}} = \rho_{\mathrm{P}}$ with ρ_{K} equal 0.03 or 0.003. (b) Normalized effective phosphorylation rate constant, $c_{\mathrm{eff}}/c_{\mathrm{eff}}^{\infty}$, in the asymmetric case $c = 10d$, $\rho_{\mathrm{P}} = 0.03$, with ρ_{K} equal 0.03 or 0.003. (c) Fraction of the phosphorylated substrate, $\rho_{\mathrm{S_p}}/\rho_{\mathrm{S}}$, for $c = 10d$, $\rho_{\mathrm{P}} = 0.03$, and ρ_{K} equal 0.03, 0.01 or 0.003.

1.9 Short conclusions

In this first part of the dissertation I reproduced the results I obtained while investigating the kinetics of the phosphorylation–dephosphorylation cycle on a 2D lattice. Presented results were published in [50] and [40]. In the first one we defined and established the dependence of effective macroscopic reaction rate coefficients, as well as the steady-state phosphorylated substrate fraction, on the diffusion coefficient and concentrations of opposing enzymes: kinases and phosphatases. Analytical expressions were found for two limits of infinite and zero-motility and numerical simulations agree with these predictions. In the regime of non-zero but small diffusion, a contribution linearly proportional to the diffusion coefficient appears in the reaction rate. In this regime, the presence of opposing enzymes creates inhomogeneities in the (de)phosphorylated substrate distributions: enzymes are surrounded by clouds of "converted" substrates. This effect becomes important at low enzyme concentrations, substantially lowering effective reaction rates. Effective reaction rates decrease with decreasing diffusion and this dependence is more pronounced for the less abundant enzyme. Consequently, the steady-state fraction of phosphorylated substrates can increase or decrease with diffusion, depending on relative concentrations of both enzymes. Additionally, steady states are controlled by molecular crowders which, mostly by lowering the effective diffusion of reactants, favour the more abundant enzyme.

In the second paper we ventured on analytically treating the regime on finite motility, and, after modifying one assumption in the model (two appropriate molecules are allowed to enter the same lattice site), we were able to conduct analytical derivations, in just few places supported by approximations, and obtain compact, almost exact expressions for EMRRCs. They were first expressed as the average time, τ_{u} (or τ_{p}), a substrate molecules spends between opposing reactions. This time, in turn, is the sum of the time to find the opposing enzyme molecule, τ_{u_1} (or τ_{p_1}), and the time to react after the first encounter with the enzyme molecule, τ_{u_2} (or τ_{p_2}). As the time τ_{u_2} (or τ_{p_2}) was found to be simply $\tau_{\mathrm{u}_2} = 1/(c\,\rho_{\mathrm{K}})$ (or $\tau_{\mathrm{p}_2} = 1/(d\,\rho_{\mathrm{P}})$), the main difficulty was in calculating the time to find the opposing enzyme molecule, τ_{u_1} (or τ_{p_1}). The final formulae show how the EMRRCs depend on reaction propensities, motilities, and densities.

In summary, our works were a step towards the determination of effective macroscopic reaction rate constants and steady states for ubiquitous cycles of opposing reactions with respect to the motility of substrates and enzymes, and their densities.

Chapter 2

Noise

Le hasard ne favorise que les esprits préparés.
-Chance favors only the prepared mind.

Louis Pasteur

2.1 Motivation

The existence of uncertainties and variability (i.e., "*noise*") is one of the main impediments to development of precise and effective biomedical control strategies. Heterogeneities arise even at the level of single, genetically identical cells, which can exhibit diverse responses when exposed to identical environmental conditions or drug regimens [8, 44, 14, 43, 38, 23, 47]. This phenotypic variation is due in large part to the intrinsic stochasticity of gene expression and associated biochemical reactions within individual cells, and it has been shown that heterogeneity can have tremendous consequences on levels ranging well beyond single cells and up to the entire population [31, 56, 4, 58, 59, 20]. In environments subjected to quick and potentially deadly changes, cells may exploit intrinsic variability and switch early into phenotypes that are more likely to survive catastrophic environmental changes [14, 56, 28, 5]. Such bet-hedging strategies enable persistent bacteria to survive severe antibiotic treatments, or persistent cancer cells to resist chemotherapy [56].

On the other hand, the goal for much of the biomedical sciences is to control biological processes to achieve a desired outcome (e.g., apply or maintain a sufficient dosage of chemicals to eliminate abnormal, cancerous cells) while minimizing deleterious effects (e.g., reduce toxicity, expense and inconvenience to the patient). In an ideal world, one would fix a treatment regimen that always achieves the same desired result without the need for intermediate changes or decisions. Real world is more complex though and every tissue, organism or person is different. We thus need to continually monitor and update the control strate-

gies, giving rise for instance to *personalized medicine*, where treatments are continuously adjusted in time.

In this chapter I will explore the potential to selectively control the phenotypes of individual cells within a larger population of genetically identical cells, all subjected to the same environmental conditions. I will show an effective control strategy, which exploits the random nature of protein production process, to attain this goal.

2.2 Self-activating gene model

The first model analyzed here is the model of self-activating gene, depicted schematically in Fig. 2.1(a). One kind of protein is synthesised from a gene whose state - active or inactive - depends on whether the protein is bound to a specific site on the gene, called the promoter site. If the promoter site is free, the gene is in its inactive state and the production of proteins happens only at some basal, low rate. If the protein is bound to the promoter, the gene is active, thus enhancing further production of protein molecules. Since the production cannot increase unlimitedly, its rate is modelled by a function that plateaus when the concentration of protein molecules increases. For this purpose, it is common to chose Hill's function. The details of the ordinary differential equation governing the time evolution of the concentration of protein are given below and the derivation of the Hill function in the Appendix to this chapter.

2.2.1 Deterministic description

In the model of self-activating gene, the protein is produced at a basal rate, k_0, plus an induced rate via positive feedback, modelled by a Hill function of the form: $k_1 x^m/(x^m+\beta^m)$, where x is the concentration of proteins, $k_0 + k_1$ is the saturated production rate and β is the concentration of protein at half maximum induction. Proteins are subject to natural degradation, proportional to their concentration. The differential equation for the time evolution of x in each cell thus reads:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = k(x) - R(x) = k_0 + k_1 \frac{x^m}{x^m + \beta^m} - \gamma x. \tag{2.1}$$

The number of stationary points (i.e., such that satisfy $\frac{\mathrm{d}x}{\mathrm{d}t} = 0$) of the above ordinary differential equation depends of course on m. If we set $m = 2$, equating $\frac{\mathrm{d}x}{\mathrm{d}t}$ to 0 returns a polynomial of degree 3, which has either one solution or three. If $\frac{\mathrm{d}x}{\mathrm{d}t} = 0$ has one solution, then the differential equation (2.1) has one stationary point, and it is stable. If $\frac{\mathrm{d}x}{\mathrm{d}t} = 0$ has three solutions, then (2.1) has two stable stationary points and one unstable. This qualitative behaviour is presented in the cartoon Figs. 2.1(c) and 2.2(b-d). However, for control strategy purposes, m was set equal to 4 (see quantitative Figs. 2.3, 2.4, 2.5, 2.6). This could potentially result in 5 stationary solutions to (2.1). We made sure other parameters

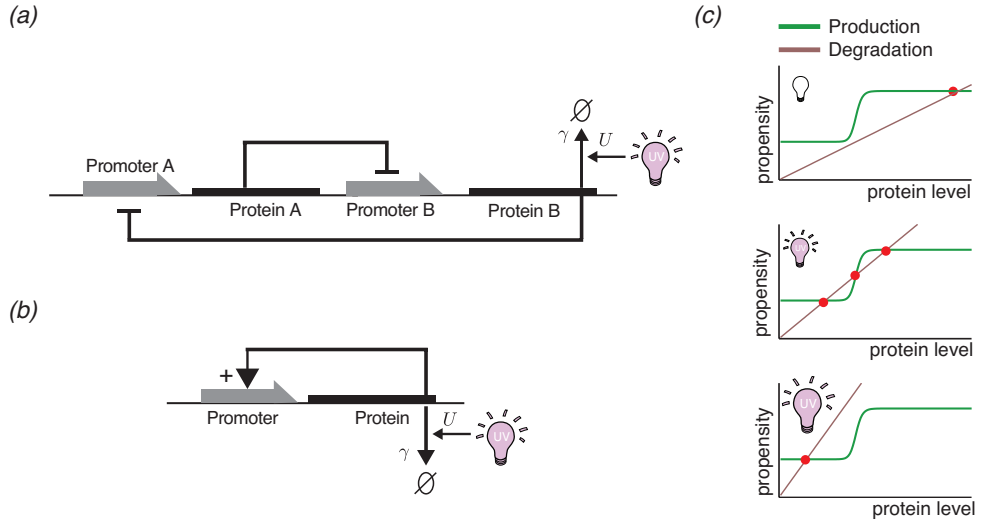

Figure 2.1: (*a*) A classical representation of a genetic toggle switch [16]. Protein A binds to Promoter B disabling synthesis of the opposing Protein B, which in turn might bind to Promoter A inhibiting production of Protein A. The degradation rate of one protein is enhanced by UV radiation. (*b*) Scheme of the self-activating gene model. This is the simplest model that exhibits bistability. Protein production is usually modelled as a two-step process - in the first one, transcription, messenger RNA molecules are "produced" from the DNA; they are then translated into protein molecules in the process of translation. Here we merge these two steps into one, neglecting mRNA production. We thus get rid of the majority of noise in this process, as mRNA production is proven to generate most of the variability of protein synthesis. Protein degradation can be enhanced by increasing levels of UV radiation. (*c*) Production and degradation rates versus the concentration of proteins, x, in the self-activating gene model. The production rate, $k(x) = k_0 + k_1 x^2/(x^2 + \beta^2)$, is plotted in green, and the degradation rate, $R(x) = (\gamma + u)x$, is plotted in brown. In a deterministic representation, protein concentrations evolve according to $\dot{x} = k(x) - R(x)$. Intersections where $k(x) = R(x)$ provide the stationary points. The three panels correspond to cases where UV is low (top, a single high stationary point), moderate (middle, two stable and one unstable stationary points) or high (bottom, a single low stationary point).

were chosen so that the polynomial resulting from $\frac{\mathrm{d}x}{\mathrm{d}t} = 0$ has at most 3, not 5 solutions and $m = 4$ case is reduced qualitatively to the $m = 2$ case.

Control strategy

The whole population of identical cells, in which the concentration of proteins being governed by the introduced ordinary differential equation is now subjected to UV radiation, which enhances the degradation rate of proteins. This means that from now on $\gamma = \gamma_0 + u(x)$. Under this control law, and with $m = 2$, the differential equation for the protein concentration

becomes:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = k_0 + k_1 \frac{x^2}{x^2+\beta^2} - (u(x)+\gamma_0)x, \qquad u \geq 0. \tag{2.2}$$

We assume that all other parameters controlling protein production (m, k_0, k_1, β) and basal degradation (γ_0) are fixed and independent of UV.

It is only by adjusting UV strength that we can control the population. Note however that UV depends solely on the protein concentration and not explicitly on time. The UV signal, u, will determine the number and location of all stationary points, as given by the zeroes of:

$$k_0 + k_1 \frac{x^2}{x^2+\beta^2} = (u(x)+\gamma_0)x.$$

We depicted in Fig. 2.1(c) three possible scenarios to compare the production rate, $k(x)$, and the degradation rate $R(x)$, where UV radiation is low (top), moderate (middle) or high (bottom). For low and high UV, there is a single stationary point at high or low values of x, respectively. For moderate UV, there are two stable and one unstable stationary points.

We introduced a model of a single cell that can be switched via UV control from a single stable high stationary point, to a pair of high and low stable stationary points, and then to a single low stationary point. Lets now examine the effects of UV on multiple cells simultaneously. For this, lets first consider a population of two independent cells, given by a two dimensional vector $\mathbf{x} = [x_1, x_2]$, where both x_1 and x_2 evolve according to the same ODE as before (equation (2.2)). Once again the number and location of stationary points depend upon the level of UV radiation. For low or high UV signals, x_1 and x_2 will converge to a stationary point that is high or low, respectively. For intermediate UV levels, a total of four stable stationary points will be possible: low protein expression in both cells, high protein expression in both cells, low protein in cell 1 and high protein expression in cell 2, or vice versa. Figures 2.2 (b-d) illustrate schematically how low, moderate or high UV radiations affect the phase diagrams of the protein trajectories for the two cells.

The deterministic model consisting of two identical cells evolves according to the following set of two differential equations:

$$\begin{cases} \frac{\mathrm{d}x_1}{\mathrm{d}t} = k_0 + k_1 \frac{x_1^2}{x_1^2+\beta^2} - (u(x_1,x_2)+\gamma_0)x_1, \\ \frac{\mathrm{d}x_2}{\mathrm{d}t} = k_0 + k_1 \frac{x_2^2}{x_2^2+\beta^2} - (u(x_1,x_2)+\gamma_0)x_2. \end{cases} \tag{2.3}$$

The set is coupled by the function $u(x)$ - UV radiation level, the control strategy that depends on protein levels of both cells. If we knew the initial condition to the above set, we would be able to predict the exact protein concentration at every time t. As a result, it would be impossible to modulate the ranking of gene expression levels in different cells using only a single control input that operates on the whole population. Luckily, when stochasticity of the biochemical reactions is taken into account, this is no longer true -

a given cell will occasionally, due to randomness, have higher protein expression than its immediate neighbour, and this could allow to systematically alter the UV strength in order to maximize or minimize this difference.

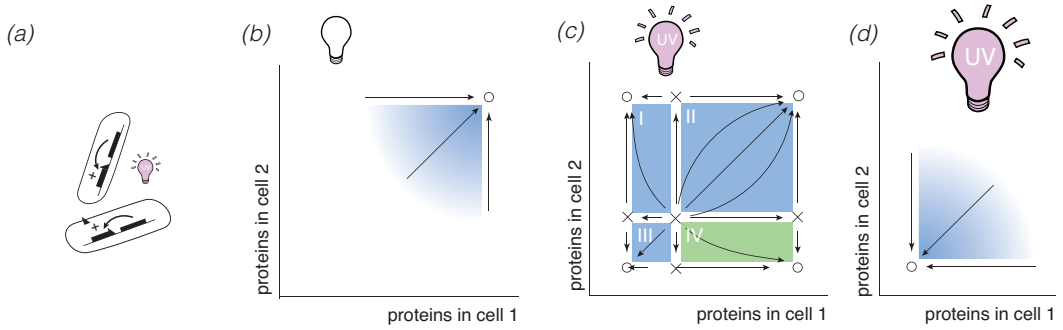

Figure 2.2: (a) A cartoon of the model considered: two identical cells each governed by a toggle switch are subjected to the same UV radiation. (b-d) A schematic phase diagram of the system with two cells, each governed by the same self-activating gene mechanism (see Fig. 2.1(b)) at three different UV radiation values: (b) low UV; (c) moderate UV; and (d) high UV. Axes represent number of proteins in cell 1 and cell 2, circles and crosses mark the stable and unstable stationary points, respectively.

2.2.2 Stochastic description

The Markov Jump Process underlying the model is defined by the state space and the transition rates between these states. If the population contains N cells, $N \geq 2$, the state space consists of integer vectors $\mathbf{x} = [x_1, \ldots, x_N]$, where x_h denotes the number of proteins in the h^{th} cell. Lets index these vectors by i, so that x_h^i denotes the number of proteins in the h^{th} cell for the i^{th} state. Accordingly, let $P_i(t)$ denote the probability that the system is in the i^{th} state at time t, and let $\mathbf{P}(t) = [P_1(t), P_2(t), \ldots]$ be the vector of probabilities for all states $\{\mathbf{x}^1, \mathbf{x}^2, \ldots\}$. Let me recall that $P(t)$ satisfies the master equation: $P'(t) = P(t)Q$. Matrix Q is given by:

$$q_{ij} = \begin{cases} k_0 + k_1(x_h^i)^2/(\beta^2 + (x_h^i)^2) & \text{for } \mathbf{x}^i = \mathbf{x}^j - \mathbf{e}^h \\ (\gamma + u(\mathbf{x}^i))x_h^i & \text{for } \mathbf{x}^i = \mathbf{x}^j + \mathbf{e}^h \\ -\sum_{h=1}^{N} \left(k_0 + k_1(x_h^i)^2/(\beta^2 + (x_h^i)^2) + (\gamma + u(\mathbf{x}^i))x_h^i\right) & \text{for } j = i \\ 0 & \text{otherwise} \end{cases} \tag{2.4}$$

where $\mathbf{e}^h$ is a N-component vector, it is zero except for the h^{th} entry, where it is unity.

Here again we will be aiming at the stationary probability distribution, i.e., such a probability vector P^{st} that $P^{\text{st}}Q = 0$. Since the potential number of proteins in each cell can be any

integer number, the number of equations in the master equation set is infinite. In such a case one normally performs the Stochastic Simulation Algorithm (SSA, [17]). This is, recall, the method used in the previous chapter, except that before we took into account spatiality. In this chapter we also perform the "conventional" SSA simulations – each time a reaction fires, new stochastic reaction rates, $k(x^i)$ and $R(x^i)$, are computed for each cell, and the UV level is based on the current number of proteins in each cell, $u(x_1, x_2, \ldots)$. Reactions change the number of proteins (by $+1$ if production occurred or -1 if degradation occurred) in one cell exactly in each time step.

Besides performing the SSA, we also resorted to a different, computationally much more efficient method called the *finite state projection*. This approach, described in details in [36], consists of reducing the infinite state space to a finite set. This means, we set a limit on the maximum number of proteins allowed in each cell. All states corresponding to the number of proteins greater or equal to the chosen limit are shrunk in one state and a reflective boundary condition is imposed there. Accordingly, the transition rate matrix is redefined, becoming a proper, finite matrix. Its entries for q_{ij}, if x^i and x^j have components smaller than the chosen limit of proteins, are given as in definition (2.4); other entries, i.e., those for which the states were aggregated in a single "boundary" state, are recomputed. In the infinite case, it is possible to estimate the error incurred through the described truncation [36].

2.2.3 Results for the self-activating gene

Effectiveness of the control strategy

Before describing any control law, we should introduce a quantity that would measure its effectiveness. For this purpose, we integrated numerically the master equation and estimated the stationary marginal and joint distributions of the protein concentration in the cells (two or more, depending on the population count). For a two-cell population, the effectiveness of the control strategy is given by the stationary probability distribution that the protein level in cell 1 exceeds that of cell 2: $P(x_1 > x_2)$, where x_1 is the protein level in the chosen cell and x_2 the protein level in the other cell. For a population containing more than two cells, we will compute the time-averaged rank of the chosen cell, i.e., all cells are first ranked with respect to the protein level they express and then an average rank, weighted by the length of the time interval, is calculated for the chosen cell.

Two cell population

We start with the small ($N = 2$) population of cells governed by the self-activating gene scheme. Our task is to construct such a control law that it favourably drives the chosen cell to express high level of proteins, while the other (genetically identical) cell is driven simultaneously to express low level of proteins.

The control law we built is depicted in Fig. 2.3 and it was built intuitively as follows: first, if the number of proteins in both cells is low (bottom left part of the plane), low UV is applied, allowing the cells to produce protein freely. Because of the symmetry between the cells, the trajectories of the protein expressions in cells 1 and 2 will move roughly parallel to the $x_1 = x_2$ line. However, because of the stochastic nature of the process, the system will fluctuate into the regions where the protein expression is higher in one cell than the other. Second, if the cell population randomly moves to the region where protein expression in cell 1 is high, and the protein expression in cell 2 is low (bottom right part of the plane), moderate UV is applied to stabilize the system in this favourable condition. Third, when expression in cell 2 becomes elevated, in order to knock it down, higher level of UV is applied. Of course this reduces expression in both of the cells, but briefly and only until cell 2 falls below the threshold, thereby allowing cell 1 a greater chance to remain closer to the high stable state at moderate UV levels. This strategy for how the UV level is chosen depending on the level of protein in the chosen cell (x_1) and in the competitor (x_2) is depicted in Fig. 2.3. Notice that the control law is as simple as possible, admitting only three values: u_1, u_2 or u_3. In our particular case, for a choice of gene regulatory parameters: $k_0 = 5$, $k_1 = 50$, $\beta = 20$, $m = 4$ and $\gamma = 0.5$, the control function $u(x)$ is given by:

$$u(x) = \begin{cases} u_1 = 0 & \text{if } x_1 < 40 \quad \text{and} \quad x_2 < 12 \\ u_2 = 0.75 & \text{if } x_1 \geq 40 \quad \text{and} \quad x_2 < 12 \\ u_3 = 1.75 & \text{if } x_2 \geq 12 \end{cases} \tag{2.5}$$

A natural question arises if with this control scheme the chosen cell really expresses higher protein concentration than the competitor cell. Figure 2.4 illustrates the results of this control law when applied to the trajectories of proteins versus time for a population of three cells, but only two of them being considered as the population (N is still equal 2). These cells correspond to one cell intentionally driven to have high expression (red), one cell intentionally driven to have low expression (blue), and a third cell (black) that is subjected to the changing UV level but upon which the control law does not depend. Marginal distributions computed using a long time simulated trajectory are shown for each cell to the right of the trajectories. From the figure it is clear that the UV control law successfully maintains the first cell in the high state (red, $\langle x_1 \rangle = 44.3 \pm 5.8$) and the second cell in the low state (blue, $\langle x_2 \rangle = 4.7 \pm 2.5$). Meanwhile, the cell that is not specifically controlled (black, $\langle x_3 \rangle = 23 \pm 21$) is free to fluctuate between high and low states.

We depicted the effectiveness of the above control strategy by comparing it to the constant (low, moderate, and high) UV radiation scheme in Fig. 2.5. In each case the joint distribution is shown on the left and the marginal distributions are shown on the right. UV levels that do not depend upon the concentrations in the two cells, shown in Figs. 2.5(a)-(c), do not break the symmetry between the cells, and the UV-dependent marginal distributions

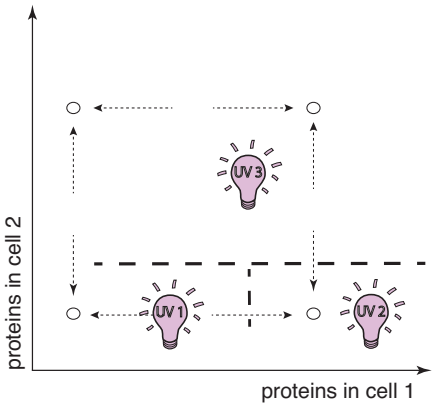

Figure 2.3: Preliminary control law for the application of UV radiation based upon comparison of the protein content for two cells. When both cells contain little protein (bottom left corner), no UV is applied ($u_1 = 0$). When protein content is high in cell 1 and low in cell 2 (bottom right corner) moderate UV is applied ($u_2 = 0.75$). When protein content in cell 2 is above a threshold ($= 12$) (upper part), high UV is applied ($u_3 = 1.75$). All parameters for the gene regulatory circuit ($k_0 = 5$, $k_1 = 50$, $\beta = 20$, $m = 4$ and $\gamma = 0.5$) are fixed and identical for both cells. Degradation rate parameters (i.e., γ and u_i) have arbitrary units of inverse time, and production rates (i.e., k_0 and k_1) have units of molecules per unit time.

are identical for the two cells. In contrast, Fig. 2.5(d) shows that the chosen UV control law is successful to make it highly probable that cell 1 has high protein expression and cell 2 has low protein expression. The stationary probability distribution that the protein level in cell 1 exceeds that of cell 2 is $P(x_1 > x_2) = 0.998$ with the specified control law. This number is obtained by numerically solving the master equation. To confirm the consistency between the stochastic simulations and the direct solutions, Fig. 2.5(d) plots the marginal distributions for both approaches.

Three or more cell population

We next extended the control law found for the two-cell population to the case of a N-cell population, $N \geq 3$. To keep the control law as simple as possible, we maintained the same three regions of different UV values shown in Fig. 2.3, as well as the same three values for the UV radiation $\{u_1, u_2, u_3\}$. The choice of the UV level is now based on the comparison of the number of proteins in cell 1, x_1, and the **maximum number of proteins in all other cells**, $\max_{i=2}^{N}(x_i)$. As above, we simulated the trajectories of all N cells simultaneously under the controlled UV radiation. Foreseeably, the effectiveness of the control law decreased when a larger population was considered. This effect is quantified in Fig. 2.6(a), (top, black diamonds), which shows the time averaged rank of the chosen cell versus the total number of cells in the population.

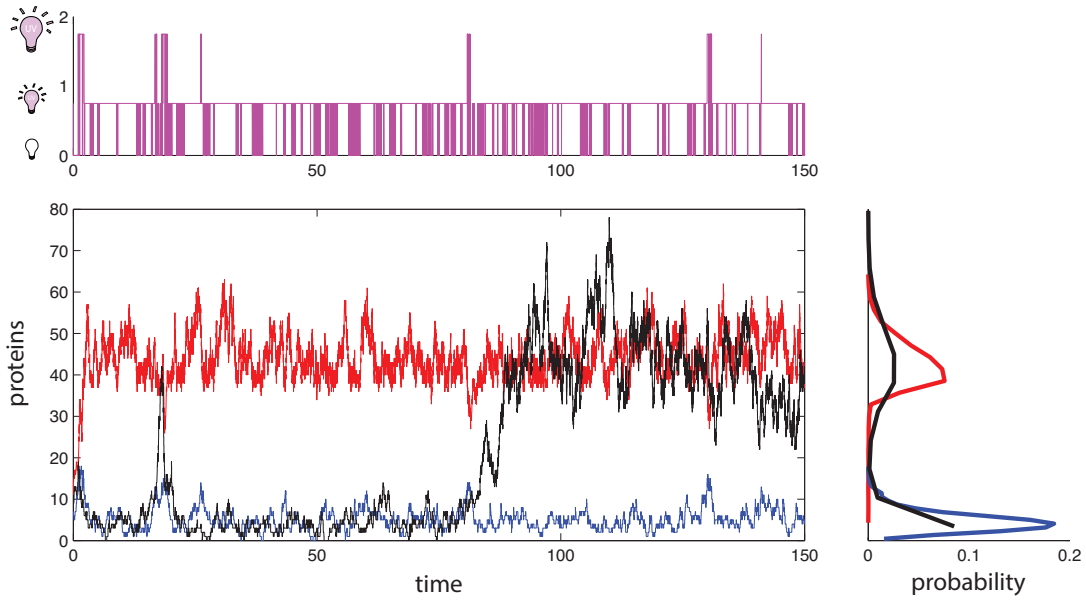

Figure 2.4: Example stochastic trajectories [17] for the number of proteins versus time for three cells, all subjected to the same UV signal over time. Trajectories for cells whose expressions are controlled to be high and low are shown in red and blue, respectively. A trajectory for a third cell that experiences the same UV signal but is not included in the control decisions, is shown in black. The top plot represents the UV radiation applied to all cells over time. The curves to the right show the marginal probabilities for the number of proteins in each cell, as determined form a single stochastic trajectory.

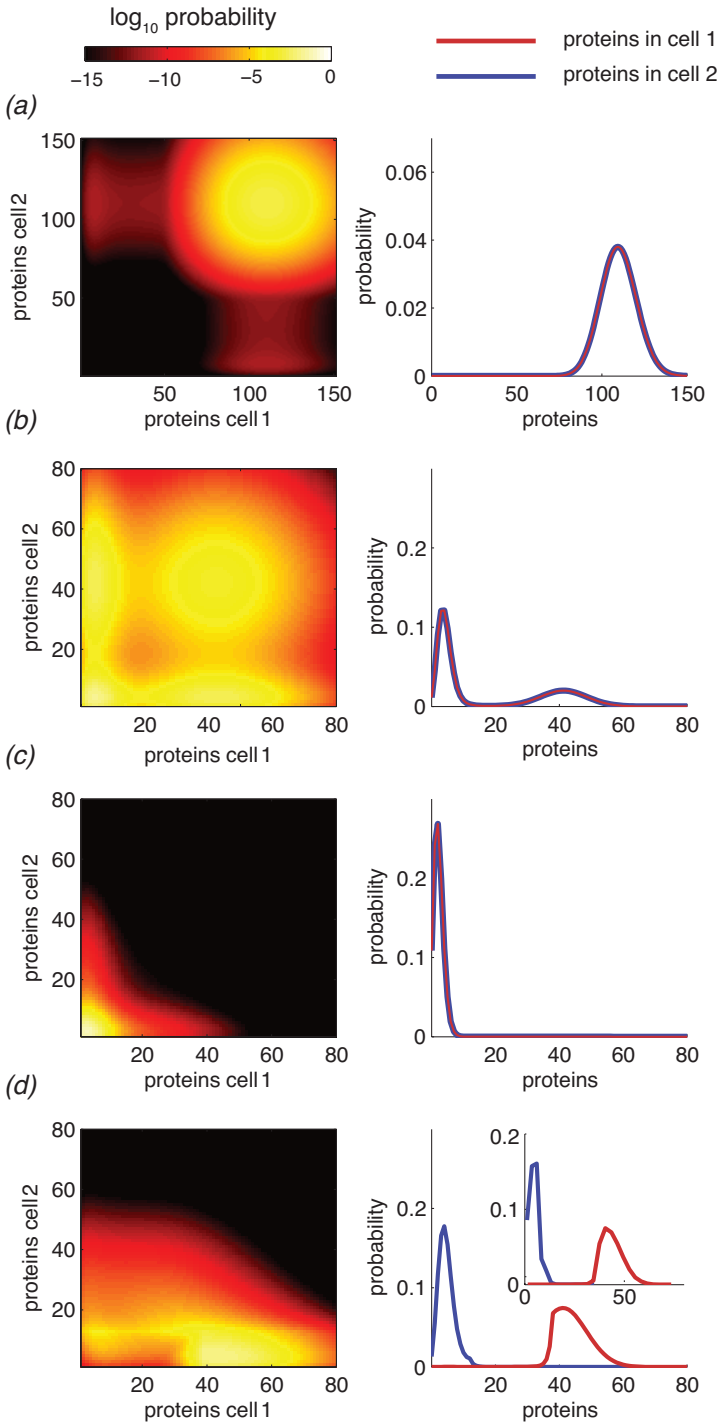

Figure 2.5: Stationary distributions for the joint probabilities (left) and marginal probabilities (right) for the protein content in two cells under four different UV control laws: (a) low UV; (b) with moderate UV; (c) with high UV; and (d) with the UV control law given in Fig. 2.3. On the right, marginal distributions for cell 1 (red) and cell 2 (blue), are compared. The inset in panel (d) shows the marginal probability distributions computed from a single trajectory obtained using the SSA [17].

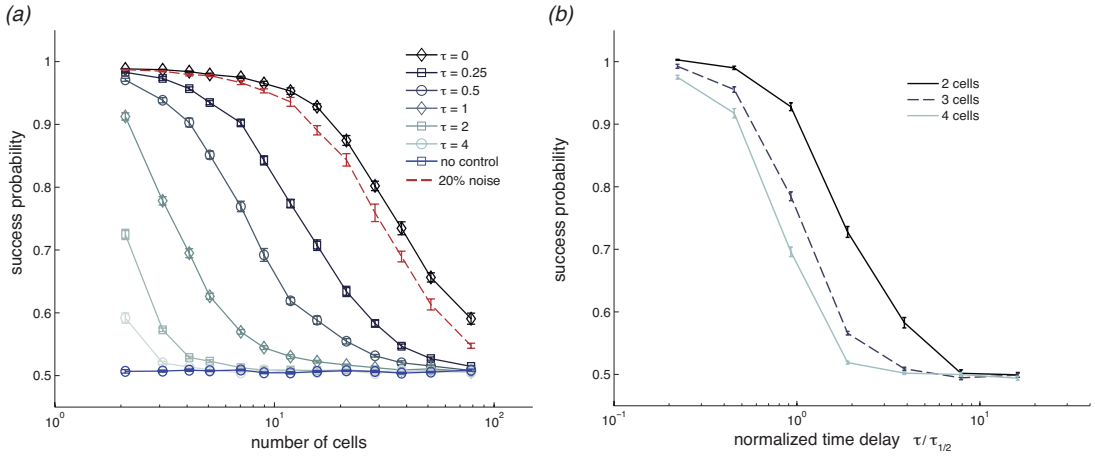

Figure 2.6: Time-averaged rank of the chosen cell versus the number of competing cells in the population for measurement noise and time delays. The UV is chosen as discussed in the main text. (*a*) Effectiveness of the control strategy versus the number of cells in the population. Solid lines correspond to different measurement time delays, τ, using the same time units as in Fig. 2.3. Dashed line corresponds to no time delay but with 20% standard deviation uncorrelated white noise added to all measurements. The blue line corresponds to the success rate in the absence of any control law. (*b*) Effectiveness of the control strategy versus the time delay normalized by the protein half-life at the lowest UV setting $\tau/\tau_{1/2} = \tau/(\log 2/\gamma)$ for 2, 3 and 4 competing cells.

2.3 Extensions to the self-activating model

2.3.1 Time delays

The first extension considered was to introduce time delays that are present in all cellular processes [1, 49], especially in the production and maturation of the fluorescent proteins that are frequently used to quantify gene regulatory responses [46]. To introduce the effects that such time delay would have on the time-evolution of protein concentrations in single cells, UV control law will now depend not upon the state at the current time, $\mathbf{x}(t)$, but upon the state at an earlier time $\mathbf{x}(t-\tau)$. Figures 2.6(a) and 2.6(b) show how the effectiveness of the control law decreases as the time delay increases (e.g., due to longer fluorescent maturation times or delays in the activation of the UV-induced SOS pathway). While short delays may be tolerated, once the time delays reach or exceed the half-life of the controlled protein, the effectiveness of the control law rapidly diminishes. Maturation times for fluorescent proteins vary considerably from about two minutes for the fastest variant of yellow fluorescent protein to thirty or more minutes for common fast folding green fluorescent proteins [39, 11]. For bacterial division times of about 30 minutes, this result emphasizes the substantial impact that the choice of protein reporter could have on our ability to actively monitor and control cellular responses.

2.3.2 Measurement errors

Lets now assume that the measurements of the protein levels were corrupted by uncorrelated white noise with 20% standard deviation. One can explore the effect that these errors would have on the success rate of the control law. Since the control law only needs to examine two cells (i.e., the chosen cell and the maximum of its competitors) and assign them to one of the three coarse regions (see Fig. 2.3), the success rate is highly insensitive to errors in the measurement of the absolute protein numbers (see Fig. 2.6(a), dashed line).

2.3.3 Extrinsic variability

In addition to the effects of time delays and measurement errors, there is yet another natural source of noise - extrinsic variability in protein synthesis rate. This quantity can vary in a reasonable range around the mean. In a series of simulations, variability in the synthesis rate was introduced by scaling both k_0 and k_1 by a common factor chosen from a normal distribution with a 10% standard deviation at the beginning of each independent simulation. The production rate of the chosen cell was set to be equal to the mean, the mean plus one standard deviation or the mean minus one standard deviation of the whole population. Figure 2.7(a) shows the average effectiveness of the control law based on the maximum protein expression in other cells. From the figure one can see that extrinsic variability can have a substantial effect on the average cell ranking. If the chosen cell has weaker

production strength than its neighbours, then its uncontrolled ranking will be lower than the 50^{th} percentile, but the addition of control can significantly increase this ranking (compare solid and dashed purple lines in Fig. 2.7(a)). If the chosen cell has stronger production rate than its competitors, then the control can further increase that cell's advantage, at least for small populations size of about 70 cells or less (compare solid and dashed orange lines in Fig. 2.7(a)). Note that the extrinsic variability plays an important role in the effectiveness of the control strategy from one random population to the next. To illustrate this concern, Fig. 2.7(b) plots the success probability for two cells with different relative production rates. When the chosen cell has an elevated production rate, control becomes very easy, but if the chosen cell's production rate falls below the threshold needed to maintain bistability (i.e., to the left of the red bar in Fig. 2.7(b)), successful control becomes almost impossible. For an exhaustive analysis, 200 populations of 30 cells apiece were randomly generated. The chosen cell's transcription rate is the mean of transcription rates in other cells. Figure 2.7(c) plots the distribution of the success probability for the controlled (green) and uncontrolled (blue) system. From the figure, it is clear that the control law substantially improves the cell rankings overall, but for populations of heterogeneous cells, the specific success rate depends heavily on the particular population.

2.3.4 Different control laws

The control law used thus far was always based on the maximum protein expression in other cells – the decision about the level of UV applied was made after assessing the maximum number of proteins in all other cells. One could wonder how the effectiveness of control changes if the decision about the UV level is based on different statistics of the protein expression levels, for example the mean, median, percentiles. Figure 2.8 shows the success rates versus the number of cells for additional control laws based upon the mean, the median, and the 75th percentile of the neighbouring cells. We also considered an independent control law where the UV depends only upon the level of protein in the chosen cell compared to a null cell assumed to have no expression (i.e., lying on the x-axis of Fig. 2.3). The success probabilities plotted in Fig. 2.8 show that our preliminary control law based on the maximum works the best for small populations. As the number of cells increases, the law based on the 75th percentile and then the median begin to outperform the control based upon the maximum. The success probability for the independent control law does not vary with the number of cells, suggesting that for very large populations, such a controller may be optimal. Interestingly, the success probability for the mean is not monotonic. This is likely due to fact that the average of a finite population fluctuating within a bimodal probability landscape can be very different from the instantaneous state of any individual within that population.

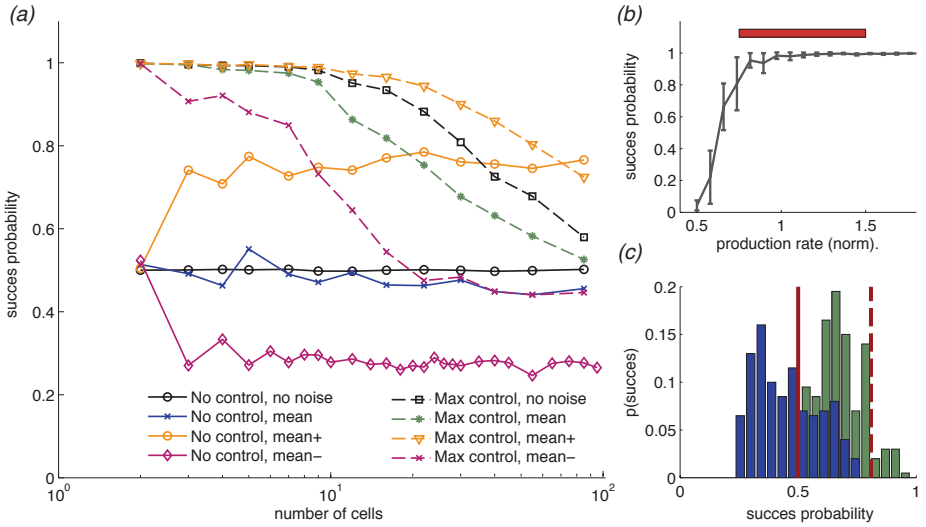

Figure 2.7: Influence of extrinsic variability in protein synthesis rates on the control law effectiveness. (*a*) Success probability as a function of the number of cells for different cases of variability of the synthesis rates. All cells are perturbed by 10% variability in their synthesis rate, and the synthesis rate of the chosen cell is fixed as the mean of the population (mean) plus or minus one standard deviation (mean+ and mean-, respectively). Results are shown for the uncontrolled system (No control), and for the system subjected to the control law based on the maximum protein expression in other cells (Max control). For comparison, previous cases without extrinsic parameter variation are replotted (No Noise). (*b*) Success probability for two cells versus the protein synthesis rate of the chosen cell divided by the synthesis rate of the other cell. The red bar on top indicates the regime for which a cell still exhibits bistability. Error bars indicate standard deviation computed over 20 simulations. (*c*) Distribution of the success probability for 200 simulations of different 30-cell populations with 10% random deviations in synthesis rates for the uncontrolled system (blue) and with the maximum control law (green). In all situations, the transcription rate of the chosen cell is equal to the population mean. We also plot reference lines for the system without protein synthesis noise (red lines, solid for the uncontrolled system, dashed for the controlled system).

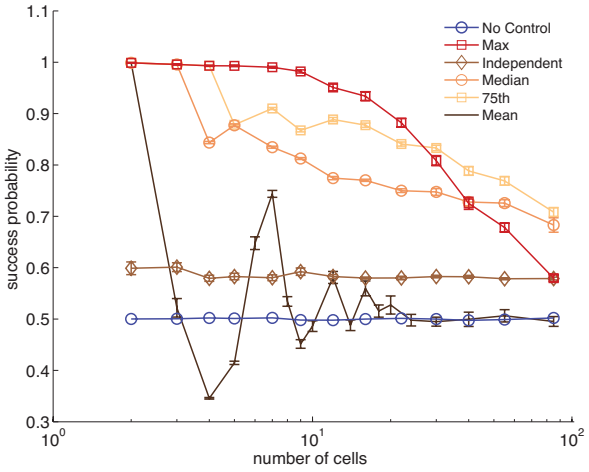

Figure 2.8: Comparison of the effectiveness of different control laws. Shown is the success probability as a function of the number of cells for different control strategies: the control law based on the maximum expression of other cells (red squares), the control law based on the mean expression, (black diamonds), the median (orange circle), the 75th percentile (gold squares) and the control law that only considers the chosen cell (brown diamonds). We plot for reference the uncontrolled system, for which the UV level is constant (blue circles). Error bars indicate the standard error of the mean (SEM) for 12 independent repeats.

2.4 Toggle switch model

In this section I present results for the original, classical model of the toggle switch. It was originally described in [16] and ever since it has then been considered in the literature many times [37, 57, 29]. Gardner *et al.* designed the real toggle switch loop that describes the expression of two mutually repressive proteins. The two mutually inhibiting proteins are LacI and λcI [16]. A schematic for the standard toggle switch model is shown in Fig. 2.1(a). The ordinary differential equations for the concentration of LacI (x) and λcI (y) would read:

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}x &= \frac{k^{(m)}_{\lambda\mathrm{cI}}}{k^{(0)}_{\lambda\mathrm{cI}}+y^m} - \gamma_{\lambda\mathrm{cI}}x \\ \frac{\mathrm{d}}{\mathrm{d}t}y &= \frac{k^{(m)}_{\mathrm{LacI}}}{k^{(m)}_{\mathrm{LacI}}+x^m} - \gamma_{\mathrm{LacI}}y. \end{cases} \tag{2.6}$$

As redesigned by Kobayashi *et al.* [25], the degradation rate of λcI is controlled via UV radiation, such that one can optogenetically push the system from a state of high λcI expression to a state of high LacI expression. We will later use this feature to specify a UV radiation control law such that the expression of LacI is made high in one pre-specified cell and small in all of the others. Using parameters originally fit to the original UV dependent

response distributions [25, 37], the rates for the production of λcI and LacI are given by:

$$k_{\lambda \mathrm{cI}} = k_{\lambda \mathrm{cI}}^{(0)} + \frac{k_{\lambda \mathrm{cI}}^{(1)}}{1+\alpha_{\mathrm{LacI}}[\mathrm{LacI}]^{2.1}} = \left(6.8 \times 10^{-5} + \frac{0.016}{1+6.1\times 10^{-3}[\mathrm{LacI}]^{3}}\right) s^{-1} \tag{2.7}$$

$$k_{\mathrm{LacI}} = k_{\mathrm{LacI}}^{(0)} + \frac{k_{\mathrm{LacI}}^{(1)}}{1+\alpha_{\lambda \mathrm{cI}}[\lambda \mathrm{cI}]^{3}} = \left(2.2 \times 10^{-3} + \frac{0.017}{1+2.6\times 10^{-3}[\lambda \mathrm{cI}]^{2.1}}\right) s^{-1}. \tag{2.8}$$

The degradation rate for LacI is the dilution rate due to division ($\gamma_{\mathrm{LacI}} = 3.8 \times 10^{-4} s^{-1}$), and the degradation rate for λ_{cI} depends upon the UV levels as:

$$\gamma_{\lambda \mathrm{cI}} = \begin{cases} 3.8 \times 10^{-4} s^{-1} & \text{at } UV = 0 \text{ J/m}^2 \\ 6.8 \times 10^{-4} s^{-1} & \text{at } UV = 6 \text{ J/m}^2 \\ 2 \times 10^{-3} s^{-1} & \text{at } UV = 12 \text{ J/m}^2 \end{cases} \tag{2.9}$$

For the stochastic simulation of the process in N genetically identical cells, the reactions are similar to those described above for the toy model, but the expanded state vector is now given by $\mathbf{z} = [x_1, y_1, \ldots, x_N, y_N]^T \in \mathbb{N}^{2N}$.

Now again the aim is to find a UV radiation control law such that the expression of LacI is high in one cell and small in all others. As before, the control law will be UV radiation applied at three levels only and the choice depends on the current state of the system. However, in this case, the control law is based upon a limited amount of directly observable information and directly controllable dynamics. It is assumed that only the level of LacI can be measured experimentally through the use of a fluorescent protein reporter, and only the degradation rate for λcI can be controlled via the application of UV radiation [37].

Taking into account all that, the control law in this model should be designed essentially in a reverse manner to the previous one (compare to Fig. 2.3): high UV is applied when the system is in a state of low LacI expression in all cells, moderate UV is applied if the system has high LacI expression in the chosen cell and low expression in other cells (desired states), and low UV is applied if the system has high LacI expression in other cells, regardless of its expression in the chosen cell. Here the system is steered to the desired states indirectly, since the UV does not affect LacI expression directly and only the inhibitory interactions of LacI and λcI enable tuning of the LacI expression using UV radiation.

Given these additional difficulties one would expect less satisfactory results. However, Fig. 2.9 shows optimistic results of the control law described here, applied to a population of ten cells. In the chosen cell (red curve) the expression of LacI is driven to be high (panel (a)), while the levels of λcI are driven to be low (panel (b)). For the other nine cells in the population, cells are supposed to express the opposite behavior (grey curves represent individual trajectories and their mean is plotted in blue). To illustrate how dramatically different the controlled cell's behavior is in comparison to the others, panel (c) illustrates the same trajectories for the ten different cells, all on the LacI-λcI plane.

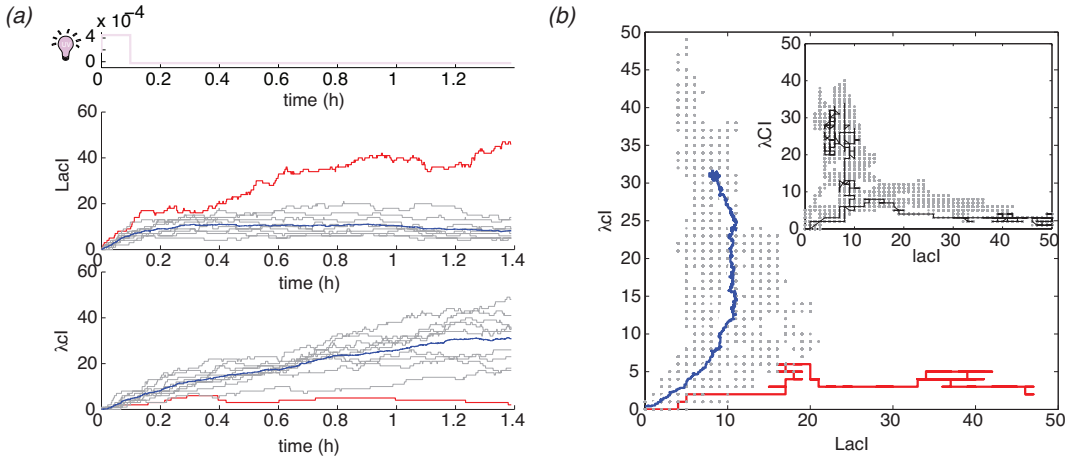

Figure 2.9: (*a*) Time evolution of the number of LacI proteins in ten cells expressing LacI and λcI proteins, subjected to UV radiation control law discussed in the main text. The trajectory for the chosen cell, in which we want to enhance LacI expression, is plotted in red. The trajectories for the nine other cells, in which we want to damp LacI expression, are plotted in grey and their mean is plotted in blue. (*b*) Number of LacI (x-axis) and λcI (y-axis) proteins in ten cells, as in panel (*a*). Parameters of the gene regulatory circuit (see main text) are fixed and identical for all cells. The inset in panel (*b*) shows two example trajectories of LacI and λcI content in two cells subjected to constant UV, under which the cell exhibits either high LacI or high λcI.

2.5 Short conclusions

Biomolecular processes that involve small numbers of molecules are by nature stochastic. This feature creates fluctuations that restrict the ability to make precise and accurate predictions for the behaviour of individual cells. Conversely, for deterministic processes that lack these intrinsic uncertainties, knowledge of the initial conditions and reaction dynamics provide accurate and complete predictions for the system evolution in time. In principle, noise is said to reduce the predictability of any system's behaviour. In particular, in gene regulation, the stochasticity is an obstacle to controlling individual cells without direct intervention in each individual cell. In this part of the thesis, where noise was the principal actor, I chose a very concrete case to show how it can actually enable precise control of many individual cells using a single input. Although cells in the presented study were identical and subjected to identical conditions, I showed that it is possible to exploit stochastic fluctuations to drive these individual cells each into a pre-specified, desired state. This capability was illustrated using two models of cell populations, where dynamics are governed by a simple self-activating gene model and the full toggle switch model. For a deterministically varying process, the phenotype of each cell depends entirely on its initial condition, and therefore it would be impossible to individually control cell phenotypes without first specifying these initial conditions. Exploration of the stochastic framework showed that fluctuations render identical cells to become distinct and susceptible to control inputs. In turn, this fluctuating susceptibility allowed for formulation of control laws that depend on the observed and continually perturbed states of the system.

2.6 Appendix

I show in here how the Hill function, used widely to model feedbacks, is derived from the law of mass-action. Consider the following binding reaction:

$$\mathrm{p} + \mathrm{DNA} \underset{u}{\overset{b}{\rightleftharpoons}} \mathrm{pDNA}, \tag{2.10}$$

where p is the protein molecule that binds to the free promoter site on the DNA, and they form a complex pDNA. We can write the differential equation for the time evolution of the complex concentration, [pDNA], in terms of the concentrations of protein, [p], and free DNA, [DNA]. For the case when there is only one copy of the gene, there can only be either 0 or 1 complexes pDNA, therefore we can also consider [pDNA] and [DNA] as the relative times that the DNA is bound, and free, respectively. The differential equation for [pDNA] reads:

$$\frac{\mathrm{d}[\mathrm{pDNA}]}{\mathrm{d}t} = b[\mathrm{p}][\mathrm{DNA}] - u[\mathrm{pDNA}]. \tag{2.11}$$

The Hill function we are looking for will model protein production. This process usually takes hours, whereas binding of protein to the DNA is thought to take only seconds, [3]. Therefore, we consider the adiabatic regime, (see [34]), in which protein production happens when the binding process has reached its steady state, i.e., we look at $\frac{\mathrm{d}[\mathrm{pDNA}]}{\mathrm{d}t} = 0$. This implies $[\mathrm{DNA}] = \frac{u}{b}\frac{[\mathrm{pDNA}]}{[\mathrm{p}]}$. Lets consider the fraction of bound DNA, i.e.: $\frac{[\mathrm{pDNA}]}{[DNA]+[\mathrm{pDNA}]}$:

$$\frac{[\mathrm{pDNA}]}{[\mathrm{DNA}]+[\mathrm{pDNA}]} = \frac{[\mathrm{pDNA}]}{\frac{u[\mathrm{pDNA}]}{b[\mathrm{p}]}+[\mathrm{pDNA}]} = \frac{1}{\frac{u}{b}\frac{1}{[\mathrm{p}]}+1} = \frac{[\mathrm{p}]}{\frac{u}{b}+[\mathrm{p}]} \overset{\beta:=\frac{u}{b}}{=} \frac{[\mathrm{p}]}{\beta+[\mathrm{p}]}. \tag{2.12}$$

If the transcription factor needs first to form a polymer of n molecules before binding to the DNA, the reaction is:

$$n\cdot \mathrm{p} + \mathrm{DNA} \underset{u}{\overset{b}{\rightleftharpoons}} (n\mathrm{p})\mathrm{DNA}, \tag{2.13}$$

and the resulting differential equation for the concentration of the whole complex (DNA and the polymer $(n\mathrm{p})$), $[(n\mathrm{p})\mathrm{DNA}]$ reads:

$$\frac{\mathrm{d}[(n\mathrm{p})\mathrm{DNA}]}{\mathrm{d}t} = b[\mathrm{p}]^n[\mathrm{DNA}] - u[(n\mathrm{p})\mathrm{DNA}]. \tag{2.14}$$

This gives the expression for the fraction of bound DNA:

$$\frac{[(n\mathrm{p})\mathrm{DNA}]}{[\mathrm{DNA}]+[(n\mathrm{p})\mathrm{DNA}]} = \frac{[(n\mathrm{p})\mathrm{DNA}]}{\frac{u}{b}\frac{[(n\mathrm{p})\mathrm{DNA}]}{[\mathrm{p}]^n}+[(n\mathrm{p})\mathrm{DNA}]} = \frac{1}{\frac{u}{b}\frac{1}{[\mathrm{p}]^n}+1} = \frac{[\mathrm{p}]^n}{\frac{u}{b}+[\mathrm{p}]^n} = \frac{[\mathrm{p}]^n}{\beta+[\mathrm{p}]^n}. \tag{2.15}$$

If we want to model self-activation of a protein, we set the production term in the ordinary differential equation to be proportional, with some positive constant, to the fraction of bound DNA, i.e., the expression computed above.

Chapter 3

Information

> It is a very sad thing that nowadays
> there is so little useless information.
>
> Oscar Wilde

3.1 Motivation

All living organisms, even the most simple ones, in order to adapt to the environment, must read and process information. This biological fact has been known for long; in the case of cells, transmitting information means sensing through receptors chemical stimuli and activating biochemical pathways in response. Such reading and transmitting signals comes at a price - it consumes energy. There are plenty of possible topologies of these regulatory circuits, yet not all of them are found in nature. The question arises why some network architectures are frequent and others nonexisting [3]. One way to approach such a question is to optimize a (specific) function by a circuit topology - it could be for example noise (minimization), time-delay of response (minimization) or information transmitted between the input and output (maximization).

A regulatory network is a collection of linked elements, which through direct and indirect interactions influence the state of each other and oneself. Here I analyze a simple regulatory model with binary input and output variables and a delayed information measured between them. The processes modelled here are irreversible, thus the whole system consumes energy. Irreversible reactions (for example biochemical cascades), come at a cost, energetic cost. When one hears "energy" in this context, one thinks about energy needed for protein production. Here, we rather look at the energy that is dissipated (used) in a given network topology, and by "network topology" I understand its architecture, for example the existence of feedbacks. Two different circuits can produce the same amount of proteins, but the energy dissipated in them is different. In other words, we assume that there is some amount of

energy needed for protein synthesis, and we think of it as a hardware of the network. This cannot be modified a lot. Instead, we want to find the best regulatory logic - software - we can implement, given a certain set of hardware.

In order to concentrate on this specific problem of dissipation coming from regulatory architecture, we choose to study a simplified model with two binary elements: a receptor and a product protein. Each element can be in one of two states: active or inactive, and its state regulates the state of the other element. We will consider a simple model, with no feedback from the output protein to the receptor, and a model with feedback.

3.2 Models

Similarly to the previous parts of this dissertation, we will be again investigating Markov Chains with continuous time. The considered system consists of two discrete random variables x_t and z_t, evolving in time, which we will interpret as the output and input, respectively. We assume that x and z can take only two values: $+$ (active state) and $-$ (inactive state). Thus the state space of the Markov Chain consists of four elements: $(-,-)$, $(-,+)$, $(+,-)$, and $(+,+)$. The transition rates will be given explicitly with a transition rate matrix.

The two models that will be investigated are: a simple model of regulation (input variable z regulates output variable x), and a model of regulation with feedback (input variable z regulates output variable x and output variable x regulates input variable z). These two models become slightly simplified if we do not distinguish between "pure" states, i.e., $(-,-)$ and $(+,+)$, and the "mixed" states, i.e., $(-,+)$ and $(+,-)$. We will be thus speaking rather about aligning/antialigning than activation/deactivation. This symmetry reduces the number of parameters and thus simplifies the computations.

Lets recall that the master equation of the Markov Chain is $P'(t) = P(t)Q$, where $P(t)$ is now the probability distribution of the model: $P(x_t, z_t)$ and Q depends on whether there is feedback from the output to the input. For the simple model, without feedback, $Q := Q_{\mathrm{S}}$:

$$Q_{\mathrm{S}} = \begin{pmatrix} -(u+s) & u & s & 0 \\ u & -(u+r) & 0 & r \\ r & 0 & -(u+r) & u \\ 0 & s & u & -(u+s) \end{pmatrix}, \tag{3.1}$$

and for the model with feedback, $Q := Q_{\mathrm{F}}$:

$$Q_{\mathrm{F}} = \begin{pmatrix} -(s+\alpha) & \alpha & s & 0 \\ y & -(r+y) & 0 & r \\ r & 0 & -(r+y) & y \\ 0 & s & \alpha & -(s+\alpha) \end{pmatrix}. \tag{3.2}$$

The transition rates are also given in Fig. 3.1 and Fig. 3.2.

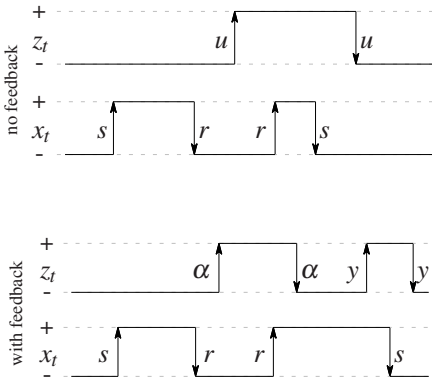

Figure 3.1: Representation of a possible time evolution of the system. Two variables flip between active (+) or inactive (-) states with respective rates. In the model without feedback (upper figure) the output variable depends on the input variable (aligns to it with rate r or antialigns, with rate s), the input variable z flips freely between its active and inactive state, regardless of the state of the output. In the model with feedback (lower figure), there is a difference in rates of flipping of the input that depends on the state of the output.

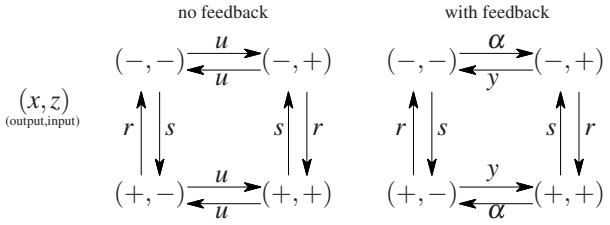

Figure 3.2: Scheme of the possible states and transitions between them for both models: without feedback (left figure), and with feedback (right figure). Since there are two binary variables there are four states; transition rates are marked next to respective arrows. Note the symmetry between the "pure" $((-,-)$ and $(+,+))$ states and the "mixed" states $((-,+)$ and $(+,-))$ in both models.

We will be interested in the joint probability $P(x_t, z_0)$, that is, we will look at the output variable x at time t and the initial state of the input variable z. This probability is needed in the computation of the central quantity we optimize - the mutual information between the initial state of the input and the state of the output at time t. After summing up over the possible states of z_0 we will obtain $P(x_t) = \sum_{z_0} P(x_t, z_0)$, which in turn is indispensable for calculating the entropy production rate of the system. These two basic quantities of interest - Mutual Information and Entropy production rate - are defined and discussed in the next section.

As mentioned before, there is a very handy symmetry introduced into both models - we do not distinguish between $(-,-)$ and $(+,+)$ states, as well as between $(-,+)$ and $(+,-)$ states. We can thus parametrize the probability distribution p by a single parameter μ_t. We write the joint distribution $P(x_t, z_0)$ in a general form:

$$P(x_t, z_0) = (a + bx_t + cz_0 + \mu_t x_t z_0)/4. \tag{3.3}$$

Since the probabilities must sum up to 1 we get that $a = 1$:

$$\begin{aligned}
&P(x_t = -1, z_0 = -1) + P(x_t = 1, z_0 = -1) + P(x_t = -1, z_0 = 1) + P(x_t = 1, z_0 = 1) = 1 \\
&\frac{1}{4}(a - b - c + \mu_t + a + b - c - \mu_t + a - b + c - \mu_t + a + b + c + \mu_t) = 1 \\
&a = 1
\end{aligned}$$

We also have the following: $P(x_t = -1, z_0) + P(x_t = 1, z_0) = P(x_t, z_0 = -1) + P(x_t, z_0 = -1) = 1/2$. Hence:

$$\begin{aligned}
&\underbrace{P(x_t = -1, z_0) + P(x_t = 1, z_0)}_{\frac{1}{2}(a+cz_0)} = \underbrace{P(x_t, z_0 = -1) + P(x_t, z_0 = 1)}_{\frac{1}{2}(a+bx_t)} = 1/2 \\
&\Rightarrow 1 + bx_t = 1 \qquad \text{and} \qquad 1 + cz_0 = 1 \qquad\qquad \Rightarrow b = c = 0.
\end{aligned}$$

We have then a parametrization of the probability distribution at any time t: $P(x_t, z_0) = \left(\frac{1+\mu_t}{4}, \frac{1-\mu_t}{4}, \frac{1-\mu_t}{4}, \frac{1+\mu_t}{4}\right)$. This reasoning was conducted under the assumption that the initial probability distribution p_0 yields the symmetry condition as well and thus is also of the form $P(x_0, z_0) = P_0 = \left(\frac{1+\mu_0}{4}, \frac{1-\mu_0}{4}, \frac{1-\mu_0}{4}, \frac{1+\mu_0}{4}\right)$.

Consideration of the initial distribution multiplies the number of models. So far I have introduced two models - simple regulation and one with feedback. However, within both of them, I can either fix the initial distribution or let it be any four-dimension probability vector satisfying the symmetry condition (i.e., I let μ_0 be any number between -1 and 1). From now on, I will use the following notation:

- S - will denote the simple model (no feedback) with the initial probability distribution equal to the stationary state one
- F - will denote the model with feedback with the initial probability distribution equal to the stationary state one
- $\tilde{\text{S}}$ - will denote the simple model (no feedback) with free initial probability distribution
- $\tilde{\text{F}}$ - will denote the model with feedback with free initial probability distribution.

3.3 Basic concepts and the main task

Here I provide definitions of the two main notions, *Mutual Information* and *Entropy production rate*, considered in this chapter, give some intuition behind them, and how they relate to each other. Both of them are derived from the same quantity, *Entropy*, a notion considered both in Mathematics and in Physics. Let me start with its definition.

3.3.1 Entropy

A discrete random variable X is a measurable function defined on a countable probability space. If we conventionally denote the probability of the i-th event by P_i, then the *Entropy* of the random variable X is defined as:

$$S(t) = -\sum_i P_i(t) \log P_i(t). \tag{3.4}$$

Heuristically speaking, it quantifies the amount of *uncertainty* a random variable carries. For example in coin flipping, the entropy is the highest if the coin is fair and decreases to zero with the probability of the heads or tails approaching to 0 or 1.

3.3.2 Entropy production rate

The *entropy production rate* formula is derived directly from *Entropy*. After differentiation of the latter with respect to time we get:

$$\begin{aligned}\dot{S}(t) &= -\sum_i \dot{P}_i(t) \log P_i(t) - \sum_i P_i(t) \frac{1}{P_i(t)} \dot{P}_i(t) \\ &= -\sum_i \dot{P}_i(t) \log P_i(t) - \left(\sum_i P_i(t) \right)'.\end{aligned}$$

Let's denote by w_{ij} the transition rate from state i to state j. We have that $\dot{P}_i(t) = \sum_{j \neq i} w_{ji} P_j(t) - \sum_{j \neq i} w_{ij} P_i(t)$. We define w_{ii} as $-\sum_{j, j \neq i} w_{ij}$, so that we can write compactly $\dot{P}_i(t) = \sum_j P_j(t) w_{ji}$ and the expression for $\dot{S}(t)$ becomes:

$$\begin{aligned}\dot{S}(t) &= -\sum_i \left(\sum_j P_j(t) w_{ji} \right) \log P_i(t) - 0 \\ &= -\sum_{i,j} P_j(t) w_{ji} \log P_i(t).\end{aligned} \tag{3.5}$$

With the definition of w_{ii}, the terms w_{ij} satisfy $\sum_j w_{ij} = 0$. Thus, the following expression:

$-\sum_i P_i(t)\log P_i(t)\sum_j w_{ij} = -\sum_{i,j} P_i(t)w_{ij}\log P_i(t)$ is equal to zero and we subtract it from (3.5) to obtain a compact form:

$$\dot{S}(t) = \left(\sum_{i,j} P_i(t)w_{ij}\log P_i(t) - \sum_{i,j} P_i(t)w_{ij}\log P_j(t)\right) = \sum_{i,j} P_i(t)w_{ij}\log\frac{P_i(t)}{P_j(t)}. \quad (3.6)$$

Further formula manipulation gives:

$$\begin{aligned}\dot{S}(t) &= \frac{1}{2}\sum_{i,j} P_i(t)w_{ij}\log\frac{P_i(t)}{P_j(t)} - \frac{1}{2}\sum_{i,j} P_i(t)w_{ij}\log\frac{P_j(t)}{P_i(t)}\\ &= \frac{1}{2}\sum_{i,j} P_i(t)w_{ij}\log\frac{P_i(t)}{P_j(t)} - \frac{1}{2}\sum_{j,i} P_j(t)w_{ji}\log\frac{P_i(t)}{P_j(t)}\\ &= \frac{1}{2}\sum_{i,j}(P_i(t)w_{ij} - P_j(t)w_{ji})\log\frac{P_i(t)}{P_j(t)}.\end{aligned}$$

This expression can be split in two in the following way:

$$= \underbrace{\frac{1}{2}\sum_{i,j}(P_i(t)w_{ij} - P_j(t)w_{ji})\log\frac{w_{ji}}{w_{ij}}}_{entropy\ flow} + \underbrace{\frac{1}{2}\sum_{i,j}(P_i(t)w_{ij} - P_j(t)w_{ji})\log\frac{P_i(t)w_{ij}}{P_j(t)w_{ji}}}_{entropy\ production\ rate}. \quad (3.7)$$

The difference between the *entropy production rate* and the *entropy flow*, is the rate at which the whole entropy of a system changes [53]. The entropy flow quantifies the flux of entropy from the system to the outside. In the steady state, as the entropy does not change, they are equal, which means that the whole entropy produced by the system is given away.

The second underbracket of (3.7) can be rewritten as:

$$\sigma(t) = \sum_{i,j} P_i(t)w_{ij}\log\frac{P_i(t)w_{ij}}{P_j(t)w_{ji}}. \quad (3.8)$$

This is the formula used from now on for the *entropy production rate* [45].

We will be looking at both *steady state entropy production rate* and *average dissipation*, which is a mean integral, calculated up to some time, of the entropy production rate.

Steady state entropy production rate

The steady state entropy production rate is simply the limit of the time-dependent quantity, entropy production rate, when $t \to \infty$:

$$\begin{aligned}\sigma(t) &= \sum_{i,j} P_i(t) w_{ij} \log \frac{P_i(t) w_{ij}}{P_j(t) w_{ji}} \\ &= \sum_{i,j} P_i(t) w_{ij} \log \frac{w_{ij}}{w_{ji}} + \underbrace{\sum_{i,j} P_i(t) w_{ij} \log \frac{P_i(t)}{P_j(t)}}_{\dot{S}(t)} \quad \xrightarrow[t\to\infty]{} \quad \sum_{i,j} P_i^{\mathrm{ss}} w_{ij} \log \frac{w_{ij}}{w_{ji}} = \sigma^{\mathrm{ss}}.\end{aligned} \tag{3.9}$$

The steady state entropy production rate corresponds, heuristically speaking, to the amount of energy available at every moment. The first question to be asked is what are the circuits that optimally transmit information, given a limited amount of steady state dissipation, σ^{ss}. The energy expense of a circuit that remains in the steady state is well defined by this quantity. However, if the system starts not in the steady state it will dissipate more energy, and a second question arises - what is the "total expense" of such a circuit. In order to answer this question we have to define yet another notion - the **average dissipation**.

Average dissipation

The *average dissipation* will have to be calculated as the integral of the entropy production rate in Eq. (3.8) over the entire time the circuit is active, τ_p, (such as the duration of the cell cycle or the interval between new inputs that kick the system into the initial out-of-steady-state condition). After some time the circuit will relax to the steady state (see the diagram in Fig. 3.3 below), where its energetic expense is given by $\hat{\sigma}^{\mathrm{ss}}$ - a rescaled steady state entropy production rate, see section "Rescaling". But the out-of-steady-state initial condition costs the system some energy. We can compare the performance of circuits that have the same steady state entropy production rate, but different regulatory designs, by considering their average dissipation until a given time τ_p:

$$\Sigma^{\mathrm{avg}}(\tau_p) = \frac{1}{\tau_p} \int_0^{\tau_p} \hat{\sigma}(\tau) d\tau. \tag{3.10}$$

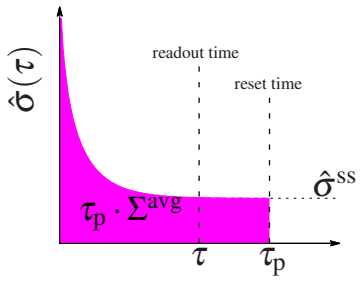

Figure 3.3: Schematic representation of system's relaxation. Entropy production rate, $\hat{\sigma}(\tau)$ relaxes with time to its steady state value, $\hat{\sigma}^{ss}$. At τ_p - the system is "kicked out" or reset, thus the pink area represents the total energy dissipated until that time. The information is collected at an earlier readout time τ.

3.3.3 Mutual Information

The most commonly known definition of mutual information, involving entropy, reads:

$$I(X,Y) = S(X) - S(X|Y), \tag{3.11}$$

where $S(X)$ is the entropy of the random variable X and $S(X|Y)$ is the conditional entropy of X given Y. The conditional entropy is given by:

$$S(X|Y) = \sum_j P(Y=j) S(X|Y=j), \tag{3.12}$$

and it quantifies the uncertainty about X, given that we know Y. So the difference between the entropy of a random variable X and its conditional entropy given another random variable Y is in fact how much we reduced our uncertainty about X assuming we know Y. The definition (3.11) is equivalent to the definition used throughout this work:

$$I[X,Y] = \sum_{i,j} P(x_i, y_j) \log \frac{P(x_i, y_j)}{P(x_i)P(y_j)}. \tag{3.13}$$

This definition of the mutual information measures how "far" is the joint probability distribution of X and Y of the product of their marginal distributions. It is easy to see, that if X and Y are independent, then $P(X,Y) = P(X) \cdot P(Y)$ and the logarithm renders the mutual information to be 0. I show the equivalence of defs. (3.12) and (3.13) in the Appendix to this chapter.

The time dependent Mutual Information for our model of input (z) and output (x) reads:

$$I[X_t, Z_0] = \sum_{x_t, z_0} P(x_t, z_0) \log \frac{P(x_t, z_0)}{P(x_t)P(z_0)}. \tag{3.14}$$

This quantity is a function of the transition rates and time. Its explicit formulae for the model without feedback, as an example, is worked out in the Appendix to this chapter.

3.3.4 Rescaling

Both of these notions, mutual information and entropy production ate are functions of time. In order to analyse the system in its natural timescale, we set $\tau = t \cdot \lambda$, where λ is the inverse of the relaxation time (smallest, non-zero eigenvalue of the matrix $\mathcal{L}$). From now on, we will talk about $I[x_\tau, z_0] = I[x_{\lambda \cdot t}, z_0]$, about $\hat{\sigma}(\tau) = \frac{1}{\lambda}\sigma(\tau/\lambda)$ and about $\hat{\sigma}^{\text{ss}} = \frac{1}{\lambda}\sigma^{\text{ss}}$.

To gain some intuition, I chose to plot in Fig. 3.4 the rescaled entropy production rate, $\hat{\sigma}$, with respect to the rescaled time τ for the model without feedback. $\hat{\sigma}$ is a constant function of τ, if the initial probability distribution is the steady state one. In other cases $\hat{\sigma}$ relaxes to $\hat{\sigma}^{\text{ss}}$ monotonically. Particularly interesting is the fact that the entropy production rate can be negative (for some μ_0 at initial times). I will not discuss the physics behind it, an avid reader can find more details in [6].

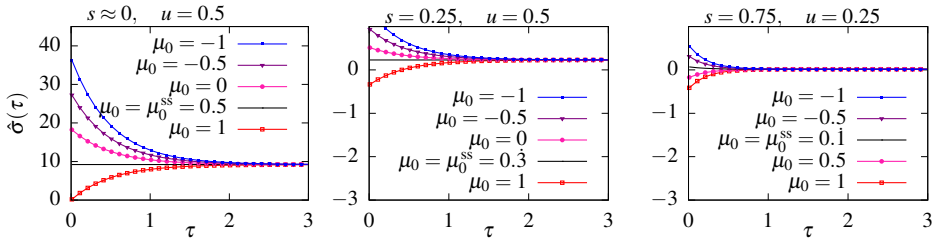

Figure 3.4: Rescaled entropy production rate $\hat{\sigma}$ plotted as function of τ for some chosen values of parameters and different initial distributions (parametrized by μ_0), for the model without feedback.

3.3.5 Task

The task is to find maximal mutual information between the input and the output, with or without constraints, for all model variants, (simple regulation, and the one with feedback; starting at steady state, or starting out of steady state) and compare their performance - the amount of information transmitted and the energy spent. The next section will show the results of the unconstrained optimization, which is numerically least demanding. Then, a constraint will be set on the steady state entropy production rate $\hat{\sigma}^{\text{ss}}$. Mathematically speaking, we are looking for the maximum of the function (mutual information) in the space of parameters (s, u, r, or s, α, y, r). Optimizing with a constraint is looking for the maximum of the function not in the whole parameter space ($\mathbb{R}_+^N$) but on the manifold given by $\sigma^{\text{ss}}(\text{parameters}) = \text{constraint}$. Finally, to compare not only the information transmitted in the models, but also its cost, we will calculate the average dissipation of the models.

3.4 Results

Here I present results of the optimization of mutual information between the input at time 0 and the output at time τ, without and with a constraint set on the steady state entropy production rate, $\hat{\sigma}^{\rm ss}$, and calculate the average dissipation, $\Sigma^{\rm avg}$. The latter quantity is a function of the transition rates, (u, s and r for the model without feedback and α, y, s and r for the model with feedback). We can also treat the initial distribution, (parametrized by a single parameter, μ_0), as an additional constraint - in our particular case we set μ_0 to be equal to $\mu_0^{\rm ss}$, i.e., we fix the initial distribution to be the steady state one. Alternatively, we can relax this constraint and let the initial distribution be different from the steady state one; in other words we can also optimize over μ_0, potentially obtaining higher mutual information.

As mentioned in the Introduction, some of the results presented here were first obtained by Francesca Mancini. In [33] she calculated the optimal information for many variants of the models for the case when there is no constraint on the steady state entropy production rate. In [32] she found the optimal mutual information also for when there is a constraint put on the steady state entropy production rate for models S and F. Here I extend her results to the models $\tilde{S}$ and $\tilde{F}$, i.e., I perform the optimization also with respect to the initial distribution. Then, I calculate the average dissipation of the optimal information (the cost).

3.4.1 Unconstrained optimization

Optimization algorithm

On the example of the simple model, S, I provide the optimization procedure. Recall that the Mutual Information is a function of the transition rates (u, s, r) and of the readout time, τ. First, one of the rates, (r), was for simplicity and without loss of generality set to 1. Then, the readout time τ was fixed and rates u and s returning highest mutual information, and satisfying $\mu_0 = \mu_0^{\rm ss}$, found. This procedure was repeated for all τ. The results are shown in Fig. 3.5: blue stars indicate the highest mutual information computed with the optimal rates u and s found for a fixed τ. Red curves in panel (a) correspond to the mutual information computed for these optimal rates plotted for the whole range of τ. The rate of antialigning, s, was found to maximize mutual information when it is equal to 0 for all times τ. The rate of the independent input flipping, u, increases linearly with τ until when it becomes constant, equal to 0.5.

The optimization algorithm was applied to the three remaining models: $\tilde{S}$, F, and $\tilde{F}$. For the model with feedback and with steady state initial distribution, F, it was exactly the same, there was only one additional rate to be found in the optimization. For the models with initial distribution not fixed, $\tilde{S}$ and $\tilde{F}$, the parameter space had one supplementary dimension, as μ_0 is no longer fixed.

Let me summarize the result of this unconstrained optimization in one graph, Fig. 3.6,

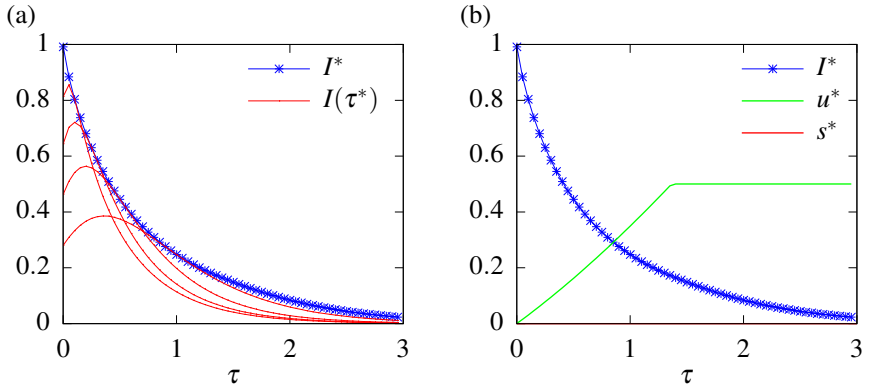

Figure 3.5: Optimal mutual information between the input, z_0, and output x_τ for the simple model, S. In panel (a) optimal mutual information for every τ is marked with blue stars. Red curves are mutual information curves calculated with the optimal parameters, but for the whole range of τ. In panel (b) optimal mutual information and optimal rates u and s are presented

and in one sentence - simple feedback allows for better information transmission only in the case when the initial distribution is fixed to its steady state value. Optimizing over the initial distribution renders the models without and with feedback equivalent. Of course the model with feedback performs better than the model without feedback if the initial distribution is fixed.

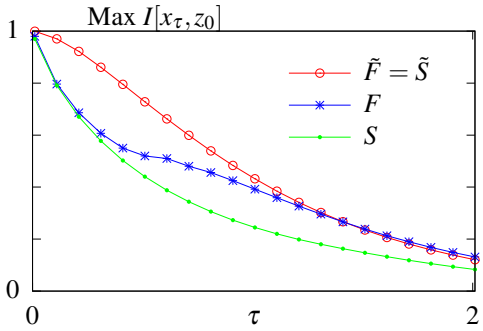

Figure 3.6: Results of the unconstrained optimization - mutual information for the simple models (S and $\tilde{S}$) and with feedback (F and $\tilde{F}$) with respect to the readout time τ. Optimization done both when the initial distribution is fixed to its steady state value (no tilde) and when the parameter is subjected to optimization as well (with tilde).

The conclusion about the equivalence of the simple model and the model with feedback no longer holds when we constrain $\sigma^{\hat{\text{ss}}}$.

3.4.2 Constraining $\hat{\sigma}^{ss}$

Optimization algorithm

The optimization in the case when we set a constraint on the steady state entropy production rate, $\hat{\sigma}^{ss}$, was numerically much more demanding, although the concept remains the same - we fix the time τ at which we read out the information, and look for rates that maximize $I[x_\tau, z_0]$ and that yield $\hat{\sigma}^{ss}$ = given constraint.

Let me begin by showing two instructive plots done for the simple models, S and $\tilde{S}$, before I summarize the results for all four models in one graph and a table.

In Fig. 3.7, optimal mutual information is plotted in the conventional way, as function of the readout time, τ. Not surprisingly, it is a decreasing function of τ for both models. Also, higher $\hat{\sigma}^{ss}$ allows for more information transmitted, and of course model $\tilde{S}$ performs better (transmits more information) than model S (solid lines are above the corresponding dashed lines). The second plot presents the same result in a different way – in Fig. 3.8, optimal

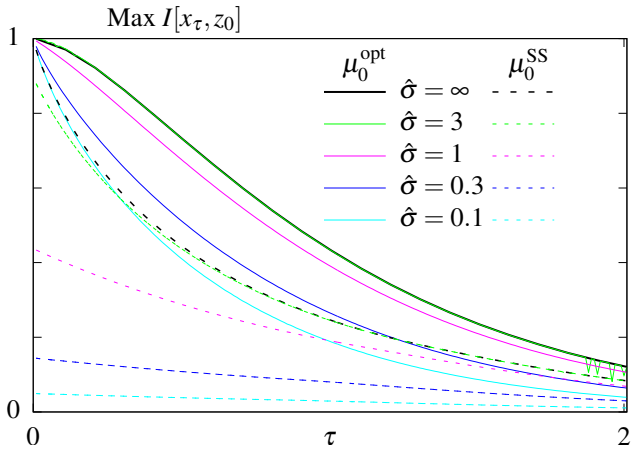

Figure 3.7: Optimal mutual information as function of the readout time, τ, for different steady state entropy production rates, $\hat{\sigma}^{ss}$, for the model S (dashed lines) an $\tilde{S}$ (solid lines).

mutual information for models S and $\tilde{S}$ is shown as function of the steady state entropy production rate, $\hat{\sigma}^{ss}$, for fixed values of the readout time, τ. I plotted in separate panels the results for models S and $\tilde{S}$, but the scales on the axis are the same, so that it is easy to compare between them. One obvious observation is that higher the steady state entropy production rate is, higher information is transmitted. Also, choosing the initial distribution, instead of starting from the steady state, significantly increases information transmitted.
Finally I present the results obtained for all four models in one plot, Fig. 3.9. The difference between optimal mutual information transmitted in models $\tilde{S}$ and $\tilde{F}$ is higher for smaller $\hat{\sigma}^{ss}$, and, as shown previously, this difference vanishes as $\hat{\sigma}^{ss} \to \infty$. The conclusion about the model with feedback performing better (transmitting more information) than the models

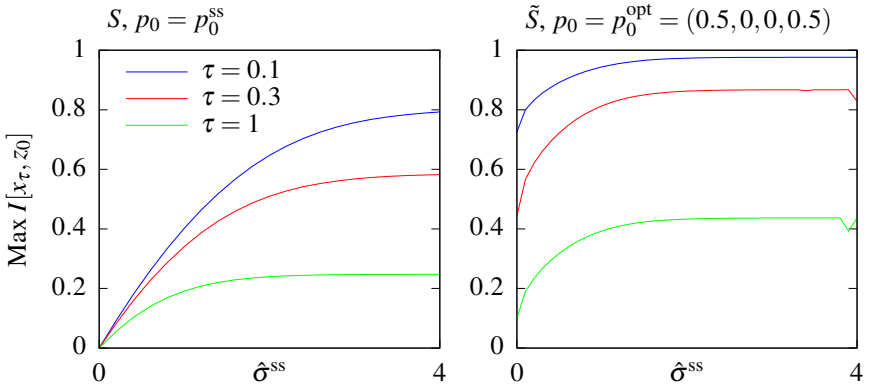

Figure 3.8: Optimal mutual information as function of the steady state entropy production rate $\hat{\sigma}^{ss}$ for different readout times, τ for the simple models. In panel (a) the initial distribution, p_0 is fixed to the steady state value p_0^{st}, in panel (b) it was subjected to optimization.

without feedback holds, as well as the observation (which was also true for the unconstrained optimization case) that the models with free initial distribution outrank the variants with the initial distribution fixed to its steady state value.

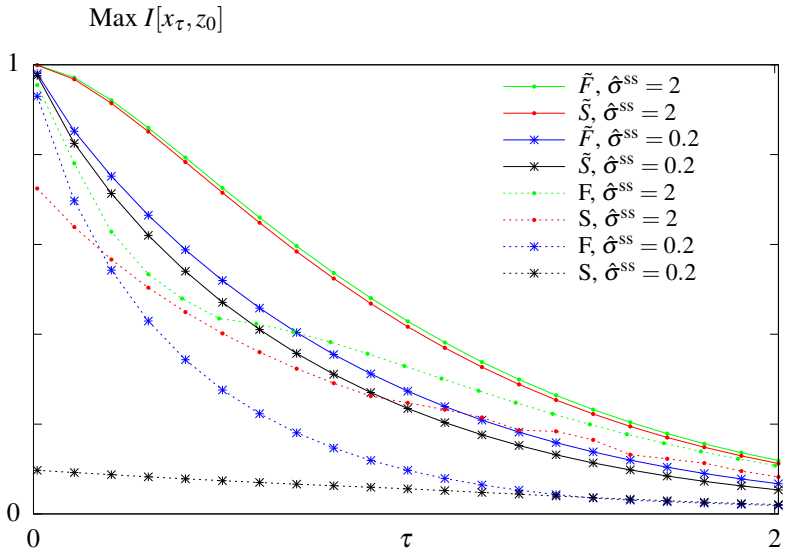

Figure 3.9: Optimal mutual information as function of the readout time, τ, for two different steady state entropy production rates, $\hat{\sigma}^{ss}$, for the models S and F (dashed lines), and the models $\tilde{S}$ and $\tilde{F}$ (solid lines).

Let me summarize all the results obtained so far in a symbolic form. I denote the optimal mutual information for each model by $I(W)$, where $W \in \{S, \tilde{S}, F, \tilde{F}\}$.

- $I(S) < I(\tilde{S}), \quad I(F) < I(\tilde{F}) \quad \forall \hat{\sigma}^{\mathrm{ss}} < \infty$, as well as without constraint on $\hat{\sigma}^{\mathrm{ss}}$,
- $I(S) < I(F) \quad \forall \hat{\sigma}^{\mathrm{ss}} < \infty$, as well as without constraint on $\hat{\sigma}^{\mathrm{ss}}$,
- $I(\tilde{S}) < I(\tilde{F}) \quad \forall \hat{\sigma}^{\mathrm{ss}} < \infty, \qquad I(\tilde{S}) \xrightarrow[\hat{\sigma}^{\mathrm{ss}} \to \infty]{} I(\tilde{F})$.

3.4.3 Cost of optimal information

We have seen that the for both models, if we can choose the initial distribution, instead of starting from the steady state, we can significantly increase information transmitted. Now a question arises if this choice of initial distribution "costs" us something? We will be calculating the average dissipation, $\Sigma^{\mathrm{avg}}(\tau_p)$, and look for the highest mutual information attainable for a given steady state entropy production rate if we allow the initial condition to be out of the steady state. As argued already, the systems that start at steady state, i.e., for which $\mu_0 = \mu_0^{\mathrm{ss}}$, will not pay additional cost (see Fig. 3.4, for $\mu_0 = \mu_0^{\mathrm{ss}}$ the function of $\hat{\sigma}(\tau)$ is constant, equal to $\hat{\sigma}^{\mathrm{ss}}$). This means that also the mean integral, $\Sigma^{\mathrm{avg}}(\tau_p)$, will be equal to $\hat{\sigma}^{\mathrm{ss}}$.

Optimization algorithm

First of all, the steady state entropy production rate, $\hat{\sigma}^{\mathrm{ss}}$, as it is an inevitable expense we pay constantly, is fixed. Then we fix the reset time, τ_p, until which Σ^{avg} will be calculated. Next, we fix Σ^{avg} itself and then transition rates returning optimal mutual information for a chosen readout time $\tau \leq \tau$ are found. With this procedure, we find Σ^{avg} for which the mutual information is highest for a given $\hat{\sigma}^{\mathrm{ss}}$. We shall call this quantity the *cost*. The result of this analysis are shown in Fig. 3.10, where for a chosen $\hat{\sigma}^{\mathrm{ss}} = 0.1$, optimal mutual information is plotted against $\Sigma^{\mathrm{avg}}(\tau_p)$. Depending on the value of τ, mutual information has either two peaks (smaller τ) or one peak (bigger τ). The higher peak, or only one, requires $\mu_0 = 1$, i.e., the initial distribution to be equal to $(0.5, 0, 0, 0.5)$. The second peak (if arises) is when $\mu_0 = -1$, which in turn means that the initial distribution is $(0, 0.5, 0.5, 0)$. This result about the optimal μ_0 holds for all $\hat{\sigma}^{\mathrm{ss}}$.

I present the interplay of the three quantities - optimal mutual information, steady state entropy production rate, and cost, in one graph, Fig. 3.12, for different readout and reset times, τ and τ_{p}. Let me introduce one more handy term that will characterise the models - the *relaxation cost*. It is schematically depicted in Fig. 3.11, which is analogous to the cartoon in Fig. 3.3. This time the pink area shades the difference between the total cost paid until time τ_{p}, $\tau_{\mathrm{p}}\Sigma^{\mathrm{avg}}$ and the energy that would be anyhow dissipated if the system started at steady state, $\tau_{\mathrm{p}}\hat{\sigma}^{\mathrm{ss}}$.

I can now compare models S with $\tilde{S}$ in terms of the cost of information transmission and the readout time τ. We know from the previous section that model $\tilde{S}$ transmits more information than model S, however it is only now that we can compare the costs. As shown in Fig. 3.12, the total cost (z-axis, in colour) generated was only slightly bigger for $\tilde{S}$ than for

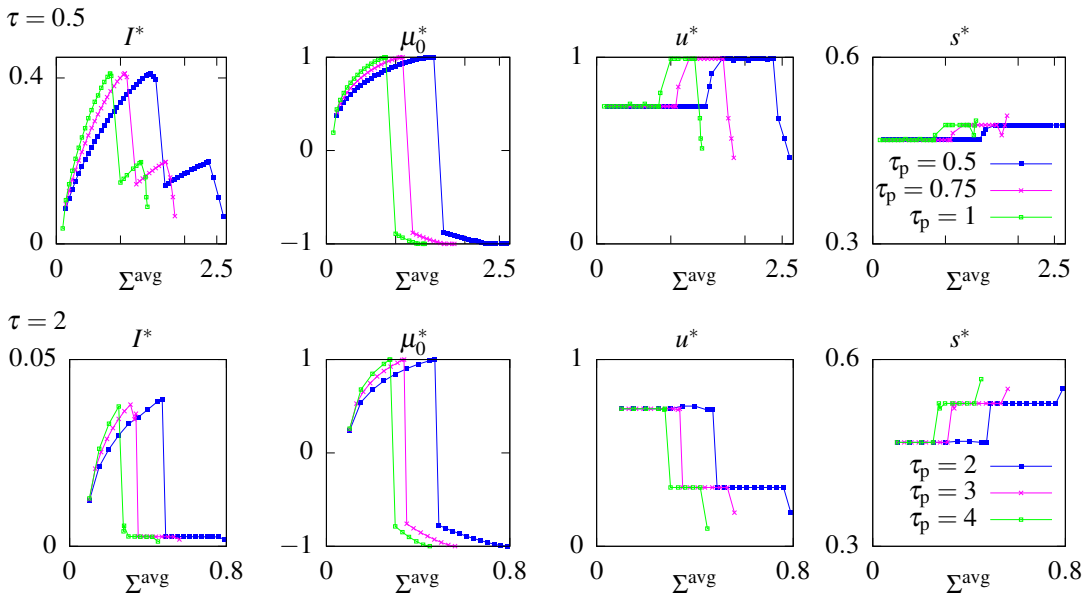

Figure 3.10: Optimal mutual information and optimal parameters μ_0, u and s for the simple model as function of the average dissipation, Σ^{avg}, for two values of the readout time, $\tau = 0.5$ (upper panels), and $\tau = 2$ bottom panels and three values of the reset time, τ_{p} (different colours of curves). Steady state entropy production rate, $\hat{\sigma}^{\mathrm{ss}}$, was fixed to 0.1.

S and the difference is more pronounced only for relatively small $\hat{\sigma}^{\mathrm{ss}}$. This holds for both sets of parameters ($\tau = \tau_{\mathrm{p}} = 0.5$ and $\tau = 1, \tau_{\mathrm{p}} = 2$), but is more visible for $\tau = \tau_{\mathrm{p}} = 0.5$. In order to quantify the intuition that $\tilde{S}$ transmits more information than S at a small price, I plotted in panel (b) of Fig. 3.12 the *information gain*, $I^* - I^{\mathrm{ss}}$, and the *relaxation cost* with respect to $\hat{\sigma}^{\mathrm{ss}}$. $I^* - I^{\mathrm{ss}}$ is the difference between the optimal information when the initial distribution is free to be optimized over ($\tilde{S}$) and the optimal information for the system with steady state initial distribution (S). In other words, I looked into how much the system pays additionally for better information if the initial condition is optimal. We see that the relaxation cost is almost the same regardless of the reset time, τ_{p}. Of course the relaxation cost decreases along with increasing steady state entropy production rate, $\hat{\sigma}^{\mathrm{ss}}$.

This analysis leads to a result stating that higher optimal mutual information obtained when the optimization over the initial distribution is allowed does not generate significantly higher costs. The same result holds for models F and $\tilde{F}$, but instead of plotting analogous figures for the model with feedback I performed another cross-comparison, this time between the models with steady state initial distribution, S with F, and the models where the initial distribution is the optimization variable, $\tilde{S}$ with $\tilde{F}$. Figure 3.13 completes previous results about mutual information. Let's recall that the model with feedback performed

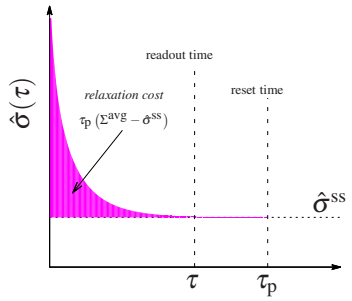

Figure 3.11: Cartoon of the relaxation of the system when the initial distribution is not the steady state one, and thus $\hat{\sigma}(\tau) \neq \hat{\sigma^{ss}}$. The pink area is called the *relaxation cost*.

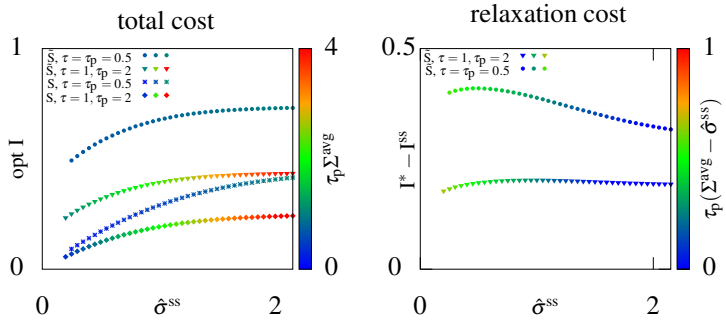

Figure 3.12: Left figure: Total cost, $\tau_p \Sigma_{avg}$, of the optimal information transmitted with respect to the steady state entropy production rate, $\tau_p \hat{\sigma}^{ss}$, for the model without feedback, that starts at the steady state distribution, S, and that optimizes the initial distribution, $\tilde{S}$. Right figure: relaxation cost, $\tau_p\,(\Sigma^{avg} - \hat{\sigma}^{ss})$, of the gain in information, $I^* - I^{ss}$, with respect to the steady state entropy production rate.

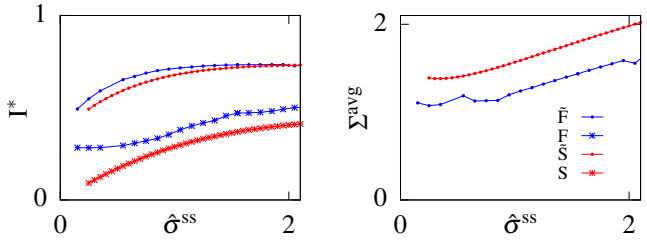

Figure 3.13: Comparison of all four models: simple (S, $\tilde{S}$) and with feedback (F, $\tilde{F}$), with the initial distribution equal to the steady state one (S, F) or optimized over ($\tilde{S}$, $\tilde{F}$). Left panel - optimal mutual information, right panel - corresponding cost, Σ^{avg}. The values of τ and τ_p were fixed to 0.5.

better (transmitted more information) than the model without feedback. However, there was a subtlety to be taken into account when distinguishing the models with optimal initial

distribution ($\tilde{S}$ and $\tilde{F}$) and steady state initial distribution (S and F). The comparison of the models with steady state initial distribution showed significant difference in the amount of information transmitted, in contrast to the variants with optimal initial distribution. In fact, the optimal mutual information for sufficiently large entropy production rates for models $\tilde{S}$ and $\tilde{F}$ was nearly the same for all times τ. This was not the case for models S and F, where for small τ, regardless of the steady state entropy production rate $\hat{\sigma}^{\mathrm{ss}}$, the optimal mutual information for model S was significantly smaller than for F.

It is thus interesting to see how the cost of optimal mutual information will classify these four models. In the right panel of Fig. 3.13 I plotted the cost of the optimal mutual information for model $\tilde{S}$ and $\tilde{F}$. I omitted plotting the results for models S and F, as we learnt already that the cost of optimal mutual information for these models is the same, equal to $\hat{\sigma}^{\mathrm{ss}}$. I gathered all the above results, along with the previous ones about optimal mutual information in Table 3.1. I use again the convention: $I(W)$ is the optimal mutual information for model W, $W \in \{S, \tilde{S}, F, \tilde{F}\}$ and $C(W)$ is the cost of the optimal information transmitted in model W.

	I^{opt}	Cost
S, F	$I(S) < I(F)$	$C(S) = C(F)$
$\tilde{S}$, $\tilde{F}$	$I(\tilde{S}) \leq I(\tilde{F})$	$C(\tilde{S}) > C(\tilde{F})$

Table 3.1: Comparison between the four models, S, F, $\tilde{S}$, and $\tilde{F}$ in terms of optimal mutual information, I, and the cost (value of Σ^{avg} calculated with optimal rates), C.

3.5 Short conclusions

In this chapter I looked into the third aspect of mathematical modelling of signalling pathways, information transmission, and presented a model of two binary random variables that regulate each other either without or with feedback. For these two models I looked for the optimal mutual information between the random variables called the input and the output. More precisely, the mutual information is a function of the transition rates between the states (four states, as the variables are binary), the time at which this information is "measured" and the initial probability distribution (assumed to be parametrized by one parameter). On the whole, there were four models considered: one simple (without feedback) with the initial distribution equal the steady state one (called S), one simple but with the initial distribution free to be optimized over ($\tilde{S}$), one with feedback and the initial distribution equal the steady state one (F) and finally one with feedback with the initial distribution free to be optimized

over $(\tilde{F})$.

The optimization was performed numerically by first fixing the time at which the output is read and then looking for the parameters that yield the highest mutual information. I also performed optimizations with a constraint set on the steady state *entropy production rate*. This thorough analysis enabled to compare the four models. We verified that the models with the initial distribution subjected to optimization transmit more information than their equivalents with the initial distribution fixed to the steady state value. Secondly, we found that the model without feedback and steady state initial distribution is worse in terms of information transmitted than the model with feedback and steady state initial distribution. Finally, the same result was shown for the models with the optimized initial distribution with such a difference that the optimal mutual information for the variant without feedback was close to the optimal mutual information for the variant with feedback for even small steady state entropy production rate.

Next I considered *average dissipation* - the mean integral of the time dependent entropy production rate, calculated up to some reset time τ_{p}. The optimal mutual information is attained for some finite average dissipation, called the *cost*. We then asked questions about the interplay of the optimal mutual information and its cost. The conclusion completes neatly the results about optimal information: **for the case of steady state initial distribution, the feedback model outranks significantly the no-feedback model in terms of optimal information, but the respective costs of optimal information are the same. In the case of the initial distribution subjected to optimization there is less difference in the optimal information, but the cost remains highly larger for the no-feedback model than for the one with feedback**.

One could look at these results from two perspectives - first would be to focus on the preeminance of the feedback and argue for its better performance either in terms of information transmitted (in the case of steady state initial distribution) or its frugality in expenses (in the case of optimized initial distribution). Conversely, one could defend the no-feedback model stating that it is only slightly worse in terms of information transmission (optimized initial distribution case) and spends exactly the same amount of energy (steady state initial distribution). But of course the global conclusion about whether feedback is beneficial or not is that it certainly performs better (or equally well) than a simple regulatory system with no feedback. And this holds both for information transmission, and the cost of transmitting this optimal information.

3.6 Appendix

3.6.1 Equivalence of two definitions of Mutual Information

Here I provide formula manipulations that lead from the most known definition of Mutual Information:

$$I[X,Y] = S(X) - S(X|Y),$$

to the definition I used in my analysis:

$$I[X,Y] = \sum_{i,j} p(x_i, y_j) \log \frac{p(x_i, y_j)}{p(x_i)p(y_j)}. \tag{3.15}$$

$$\begin{aligned}
I[X,Y] &= S(X) - S(X|Y) = \\
&= -\sum_i P(X=i) \log P(X=i) - \sum_j P(Y=j) S(X|Y=j) \\
&= -\sum_i p(x_i) \log p(x_i) + \sum_j p(y_j) \sum_i p(x_i|y_j) \log p(x_i|y_j) \\
&= -\sum_i p(x_i) \log p(x_i) + \sum_{i,j} p(y_j) \frac{p(x_i, y_j)}{p(y_j)} \log p(x_i|y_j) \\
&= -\sum_i \left(\sum_j p(x_i, y_j) \right) \log p(x_i) + \sum_{i,j} p(y_j) \frac{p(x_i, y_j)}{p(y_j)} \log p(x_i|y_j) \\
&= -\sum_{i,j} p(x_i, y_j) \log p(x_i) + \sum_{i,j} p(x_i, y_j) \log \frac{p(x_i, y_j)}{p(y_j)} \\
&= \sum_{i,j} \left(p(x_i, y_j) \log \frac{1}{p(x_i)} + p(x_i, y_j) \log \frac{p(x_i, y_j)}{p(y_j)} \right) \\
&= \sum_{i,j} p(x_i, y_j) \log \frac{p(x_i, y_j)}{p(x_i) \cdot p(y_j)}.
\end{aligned}$$

3.6.2 Calculations for the model without feedback

Let me present here some calculations and explicit formulae obtained for the simpler model, without feedback. I would like to show that some work might be done still with a pencil and a piece of paper, with no numeric power at hand. These calculations were originally conducted by Francesca Mancini in [33], for a simple case when s was set to 0 and they were shown again in [32] but for $\mu_0 = \mu_0^{\text{ss}}$. I reproduce the formulae for model S with no assumptions on rates (except that $r = 1$ as everywhere before, without loss of generality) and the initial distribution.

The normalized eigenvector corresponding to the zero eigenvalue gives the stationary probability distribution, i.e., we solve the equation for the stationary probability of a Markov Chain: $P \cdot Q_{\rm S} = 0$, where $Q_{\rm S}$ reads:

$$Q_{\rm S} = \begin{pmatrix} -(u+s) & u & s & 0 \\ u & -(u+r) & 0 & r \\ r & 0 & -(u+r) & u \\ 0 & s & u & -(u+s) \end{pmatrix} \tag{3.16}$$

We obtain:

$$\begin{aligned} &P^{\rm ss}(--,-+,+-,++) \\ &= \left(\frac{1+u}{2(1+s+2u)}, \frac{s+u}{2(1+s+2u)}, \frac{s+u}{2(1+s+2u)}, \frac{1+u}{2(1+s+2u)} \right). \end{aligned} \tag{3.17}$$

With the above we have straightaway the steady state entropy production rate, $\sigma^{\rm ss}$:

$$\begin{aligned} \sigma^{\rm ss} &= \sum_{i,j} P_i^{\rm ss} w_{ij} \log \frac{w_{ij}}{w_{ji}} \\ &= 2\frac{1+u}{2(1+s+2u)}(-s)\log\frac{s}{1} + 2\frac{s+u}{2(1+s+2u)}(-1)\log\frac{1}{s} = \frac{u(s-1)\log s}{1+s+2u}. \end{aligned} \tag{3.18}$$

Calculation of Mutual Information:
This calculation has been thoroughly done in the Appendix in [33]; I provide here my results for the simple model S with $r = 1$. The joint distribution $p(x_t, z_0)$ is computed, after necessary manipulations, from the transition rate matrix $Q_{\rm S}$, as the exponential of its eigenvalues multiplied by the left and right eigenvectors. In our case $P(x_t, z_0)$ is:

$$\frac{e^{-t(s+2u+1)}\left(-e^{(s+1)t}(s-1) + e^{t(s+2u+1)}(s-2u+1) + e^{2tu}(s+\mu_0(s-2u+1)-1)\right)}{4(s-2u+1)}, \tag{3.19}$$

in the first and fourth component, and

$$\frac{e^{-t(s+2u+1)}\left(e^{(s+1)t}(s-1) + e^{t(s+2u+1)}(s-2u+1) - e^{2tu}(s+\mu_0(s-2u+1)-1)\right)}{4(s-2u+1)} \tag{3.20}$$

in the second and third component.

Summing the above by z_0 and by x_t, we obtain $P(x_t)$ and $P(z_0)$, respectively. These three quantities are all we need to plug into the definition of mutual information. We obtain:

$$I[x_t, z_0] = \frac{1}{2}\left((1-A-B)\log(1-A-B) + (1+A+B)\log(1+A+B)\right), \tag{3.21}$$

where:

$$A = \frac{-e^{-2ut}(1-s)}{1+s-2u}, \tag{3.22}$$

$$B = \frac{e^{-(1+s)t}\left(\mu_0(1+s-2u)-(1-s)\right)}{1+s-2u}. \tag{3.23}$$

The above can be simplified a lot. If we substitute by K the following expression:

$$\begin{aligned} K = A + B &= \frac{-e^{-2ut}(1-s)}{1+s-2u} + \frac{e^{-(1+s)t}\left(\mu_0(1+s-2u)-(1-s)\right)}{1+s-2u} \\ &= \mu_0 e^{-(1+s)t} + \frac{1-s}{1+s-2u}\left(e^{-(1+s)t} - e^{-(1+s)t}\right), \end{aligned} \tag{3.24}$$

we obtain a very neat expression for $I[x_t, z_0]$:

$$I[x_t, z_0] = \frac{1}{2}\left((1+K)\log(1+K) + (1-K)\log(1-K)\right). \tag{3.25}$$

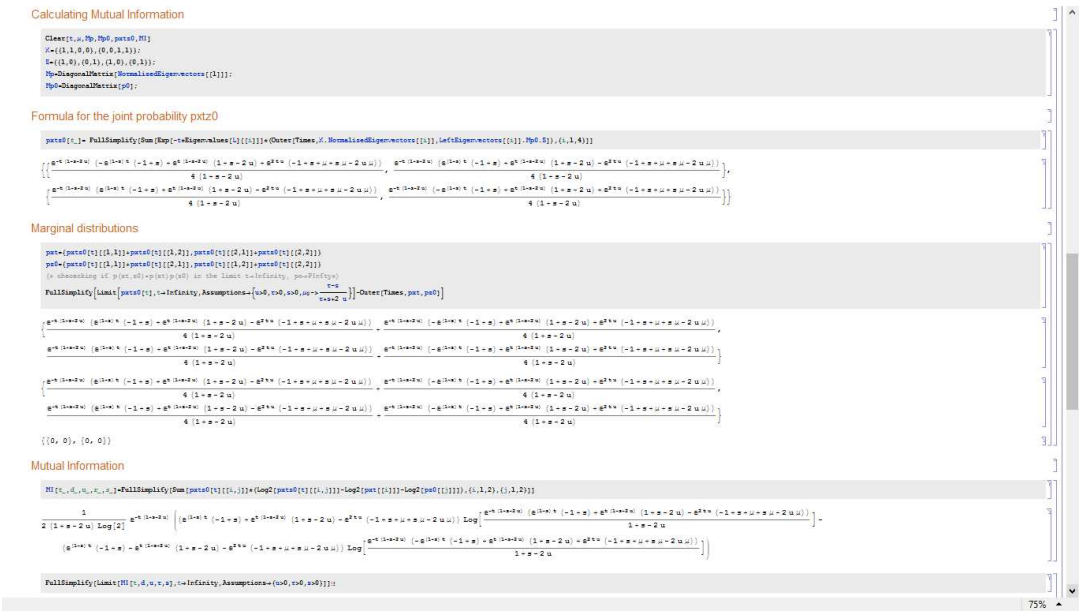

Figure 3.14: Screenshot of Mathematica file I used for computations. Probability $p(x_t, z_0)$ and Mutual Information.

I prepared a plot for the reader to gain some intuition about the mutual information. It doesn't show the optimization results, just the behaviour of the model. In Fig. 3.15 mutual information is plotted with respect to the rescaled time, τ. Intuitively, after some time it decreases for all values of the parameters, but what might be surprising is the peak of

information that appears for $\tau > 0$. The appearance of the maximum is explained for example in [41]. An intuitive reasoning gives a hint - the system needs some time to learn about itself. Not everything is known at $t = 0$ and we learn about the transition rates if we let the system evolve, before it decorrelates and the mutual information decreases.

The procedure of finding the optimal mutual information is a standard method of differentiating the function, equating the gradient to 0 and checking that the given extremum is a maximum. First the time, τ, was fixed, and then for that τ the parameters (u, s or α, y, s) that return the maximum mutual information were found. If there was a constraint set on the steady state entropy production rate, $\hat{\sigma}^{\mathrm{ss}}$ then the parameters had to satisfy a condition. All these steps are done in Mathematica software, a platform for symbolic computations.

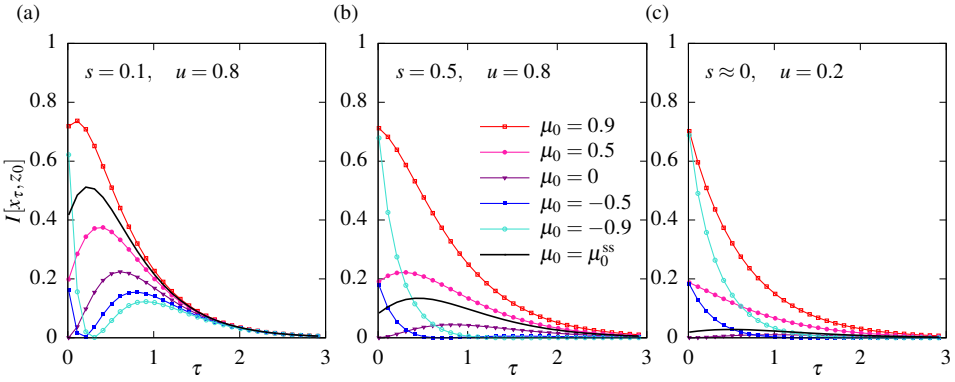

Figure 3.15: Mutual Information plotted with respect to time τ for chosen values of the parameters and different initial distributions (parametrized by μ_0), model without feedback.

3.6.3 A remark about the range of mutual information

Another detail to which I would like to draw the reader's attention is the range of the function I plotted. You observe that in all graphs the mutual information never exceeds the value of 1. It is not due to a normalization! In our particular model it comes from the fact that the variables are binary.

Lets look at the conditional entropy of X given Y. Intuitively, this is a weighted sum of $S(X|Y = j)$, where $P(Y = j)$ are the weights. Further writing gives:

$$\begin{aligned} S(X|Y) &= -\sum_j P(Y=j) \sum_i P(X=i|Y=j) \log P(X=i|Y=j) \\ &= -\sum_{i,j} P(Y=j) P(X=i|Y=j) \log P(X=i|Y=j) \end{aligned}$$

$$= -\sum_{i,j} P(X=i, Y=j) \log \frac{P(X=i, Y=j)}{P(Y=j)}.$$

This quantity is of course non-negative, since the logarithm of a quantity smaller or equal to 1 is non-positive (and the expression under the logarithm is indeed smaller or equal to 1, as it is $\frac{P(X=i,Y=j)}{P(Y=j)} = P(X=i|Y=j) \leq 1$), so the negative sign in front of the whole sum guarantees that $S(X|Y) \geq 0$.

We have then that:

$$I(X,Y) = S(X) - S(X|Y) \leq S(X) = -\sum_i p_i \log p_i, \tag{3.26}$$

and since the function $f(x) = x \log x$ is a convex function for $x \in [0,1]$, we can use Jensen's inequality (that holds for any convex function f):

$$\frac{f(x_1) + f(x_2) + ... + f(x_n)}{n} \geq f\left(\frac{x_1 + x_2 + ... + x_n}{n}\right). \tag{3.27}$$

Taking $x_i = p_i$ we obtain:

$$\frac{\sum_i^n p_i \log p_i}{n} \leq \frac{p_1 + ... + p_n}{n} \log \frac{p_1 + ... + p_n}{n}, \tag{3.28}$$

and this gives a very useful boundary:

$$S(X) = -\sum_x p_i \log p_i \leq -\log(\frac{1}{n}) = \log(n). \tag{3.29}$$

This means that the mutual information is bounded by the logarithm of the number of states of the variables

$$I(X,Y) \leq \min(S(X), S(Y)) \leq \log(\min(|X|, |Y|)). \tag{3.30}$$

In our case both variables are binary, the logarithm used here is a base 2 logarithm, thus we have:

$$I(x_\tau, z_0) \leq \log_2 2 = 1. \tag{3.31}$$

This part of the appendix, hopefully instructive, was to make sure we know why the range of mutual information on all plots is contained in $[0,1]$. One can easily get used to the range $[0,1]$, as normalizations are a natural rescaling done, but here it is not a normalization, but a boundary resulting from the nature of the variables - having two binary variables allows to transmit at most one bit of information.

Summary

In the presented three-part dissertation I investigated three "features" of signalling pathways and tried to mathematically embrace them on a canvas of different models. Their common trait is that they are all modelled by a Markov Chain with continuous time.

In the chapter entitled "Space" I investigated the correspondence between microscopic and macroscopic reaction rate coefficients in a model of a phosphorylation-dephosphorylation cycle with respect to diffusion (motility). Spatiality refers to the environment in which the reactions occur - a biological membrane (a sphere), modelled by a two-dimensional triangular lattice where molecules are allowed to move with given motilities and react when in adjacent lattice sites with given propensities: microscopic reaction rate constants. The state space of the underlying Markov Chain is huge and it was by numerical simulations that we found steady states of the system (fraction of phosphorylated substrates), as well as effective macroscopic reaction rate constants, EMRRCs, as functions of reaction propensities, fractional densities of substrates, and motility. Analytical approach was done thoroughly in limiting cases of infinite and zero motility, which agree with numerical simulations. As an extension, I proposed another analytical approach to the case of finite, non-zero motility, for a slightly modified model (the change consisted in letting two molecules enter the same lattice site and allowing reaction to take place only if the substrate and enzyme entered the same lattice site). In this part not only did I take into account spatiality by modelling the biological membrane to be a (triangular) lattice, but I also investigated phenomena directly related to spatiality - the importance of the lattice size and the molecular crowding.

The chapter "Noise" dives into the branch of modelling that assumes events occur randomly with a probability, not deterministically with an intensity. It is already accepted that this stochastic setting, in contrast to the deterministic one, models more accurately many of the genetic phenomena, but in this chapter something more has been shown - the title character of this chapter, i.e., stochastic fluctuations were exploited to drive individual, yet identical cells into a pre-specified, desired state. This capability was illustrated on models of cell populations, where dynamics are governed by a self-activating gene scheme and the classical genetic toggle switch. A single universal input, in this case UV radiation, was applied

simultaneously to the whole population of cells, whose protein level fluctuates according to discrete stochastic reactions. Here again a lot was done by simulations, since the Markov Chain is defined on a big state space, but we were able to write explicitly its transition rate matrix. We quantified the effectiveness of the control law and concluded that a meticulous, but yet very simple design of a control law, one chosen cell exhibited a qualitatively different phenotype than the others, i.e., synthesised more proteins than the competing cells. This control law was applied to different populations sizes and with different time delays.

In the chapter "Information" I considered how information is transmitted in signalling pathways. In particular, I analyzed a model consisting of an input and an output, both binary random variables, so that the state space of the Markov Chain here was a four element set, with explicit transition rates, and I was able to calculate by hand the steady state probability distribution. The title character of this chapter was the mutual information calculated between these two variables. The model on which I worked was either a simple regulation of the output by the input, or a regulation with feedback from the output to the input, both with free or fixed initial distribution; on the whole four models variants. The transition rates of the Markov Chain defined not only the steady state probability distribution of the system, but also the mutual information, which is a function of time, and these transition rates. We looked for its optimal value under a constraint set on entropy production rate, which is also a function of the transition rates. I classified the four models in terms of optimal information transmitted (highest mutual information) and the cost of this transmission (defined as the average entropy production rate calculated up to a fixed time). I found that the models with the initial distribution subjected to optimization are more informative (have higher optimal mutual information) than their equivalents with the initial distribution fixed to the steady state value. Also, the model without feedback and the steady state initial distribution is less informative than the model with feedback and steady state initial distribution. Finally, the same result was shown for the models with the optimized initial distribution with such a difference that the optimal mutual information for the variant without feedback was close to the optimal mutual information for the variant with feedback for even small steady state entropy production rate. When it comes to the cost, in the case of steady state initial distribution, although the feedback model outranks significantly the no-feedback model in terms of optimal information, the respective costs of optimal information are the same. In the case of the initial distribution subjected to optimization there is less difference in the optimal information, but the cost remains highly larger for the no-feedback model than for the one with feedback.

Appendix

The common feature of all models presented in this thesis is the mathematical tool used to build and analyze them – Markov Chains with continuous time, defined on a finite or countable state space – the *Markov Jump Processes*.

A Markov Jump Process, $(X_t)_{t\geq 0}$, is defined by the state space, Ω, and the transition rates between states, and an initial probability distribution [55]. The transition rates form the matrix Q:

$$[Q]_{ij} = \begin{cases} q_{ij} & \textit{transition rate} \text{ from state } i \text{ to } j,\ i \neq j \\ -\sum_{j,j\neq i} q_{ij} & \text{on the diagonal,} \end{cases} \tag{3.32}$$

where $i, j \in \Omega$.

The probability that the system is at time t in a given state is given by the vector $P(t)$, which satisfies, along with the transition rate matrix Q, the *master equation*, [55]:

$$P'(t) = P(t)Q. \tag{3.33}$$

With an initial condition $P(0)$ provided, one could solve the above master equation and obtain the time-dependent probabilities $P(t)$. However, we were not looking for the whole evolution in time, but the behaviour in the stationary state, i.e., for such a vector P^{st} that satisfies:

$$P^{\text{st}}Q = 0. \tag{3.34}$$

The i-th component of P^{st} gives the stationary probability of residing in the i-th state [55].

In all our models, the Markov Chains are aperiodic and irreducible, i.e., we can get form any state i to every state j. Such chains are called ergodic and they satisfy the *ergodic theorem*: *If $X(t)$ is ergodic, then the random variable*

$$\eta_T := \frac{1}{T} \int_{t_0}^{t_0+T} X(t)dt, \tag{3.35}$$

tends to $\mathbb{E}X(t)$ *as* $T \to +\infty$.
This theorem allows to compute the ensemble average, $\mathbb{E}X(t)$, by taking a sufficiently long time-average η_T from a single realization (trajectory) of $X(t)$ [9]. Ergodic chains have a unique stationary probability distribution.

In chapter 1 the average number of molecules (of phosphorylated or dephosphorylated substrate) were found by taking one, sufficiently long trajectory. The same was done in chapter 2, where we computed the average numbers of proteins in cells. In chapter 3 we were interested in the stationary probability distribution, which was computed analytically.

I provide the state spaces, Ω_i, $i = 1, 2, 3$, for every model analyzed in chapters 1, 2, and 3.

In the first chapter the state space was the set of functions attributing to every lattice site a given molecule occupying it or an empty space if there was no molecule. Recall that in the basic model of this chapter, there were four types of molecules: kinase K, phosphatase P, unphosphorylated substrate $\mathrm{S_u}$ and phosphorylated substrate $\mathrm{S_p}$. Therefore the state space, Ω_1, is the Cartesian product of a five element set:

$$\Omega_1 = \{\varnothing, \mathrm{K}, \mathrm{P}, \mathrm{S_u}, \mathrm{S_p}\}^{\mathrm{L}} \tag{3.36}$$

where L is a 2-dimensional triangular lattice, forming a square domain with periodic boundary conditions. The set of possible transition rates contains only three elements: c - phosphorylation, d - dephosphorylation, and $m/6$ - hopping to adjacent empty site.

Further on in this chapter we investigated the molecular crowding effect, i.e., we added to the set of molecules an additional, non-reagent molecule. Thus the state space Ω_1^{C} is:

$$\Omega_1^{\mathrm{C}} = \{\varnothing, \mathrm{K}, \mathrm{P}, \mathrm{S_u}, \mathrm{S_p}, \mathrm{C}\}^{\mathrm{L}}, \tag{3.37}$$

where C is the crowder molecule. The set of possible transition rates contains an additional element, m_{C} - the motility of crowder molecules.

In the model of transient enzyme–substrate complexes the state space is enlarged by additional states accounting for complex formation and disruption - a pair of proper enzyme molecule and substrate molecule can form a complex or drift apart, and a complex can either fire a reaction or disrupt. The transition rates set is also larger: $\{c_1, c_2, c_3, c_4, d_1, d_2, d_3, d_4, m/6\}$, see (1.25).

Finally, in the "multiple lattice occupancy" model, we allow a substrate molecule and an enzyme molecule to enter the same lattice site. However, neither two substrate molecules, nor two enzyme molecules can occupy the same lattice site. Therefore for this model variant

the state space Ω_1^{M} is:

$$\Omega_1^{\mathrm{M}} = \{\varnothing, \mathrm{K}, \mathrm{P}, \mathrm{S_u}, \mathrm{S_p}, \mathrm{KS_u}, \mathrm{KS_p}, \mathrm{PS_u}, \mathrm{PS_p}\}^{\mathrm{L}}. \tag{3.38}$$

The set of possible transitions remains the same, as there are no additional types of reactions.

In the second chapter, we investigated a population of N identical cells containing proteins. The state space Ω_2 is:

$$\Omega_2 = \{[x_1, \ldots, x_N] | x_h \in \mathrm{N} \quad \text{is the number of proteins in the } h\text{-th cell}\}. \tag{3.39}$$

This time we are able to write explicitly the transition rate matrix Q, see (2.4).

In the third chapter we analyzed a system of two discrete random variables, which can take only two values: "+" (active state) and $''-''$ (inactive state). Thus the state space of this Markov Jump Process, Ω_3, is simply:

$$\Omega_3 = \{(-,-), (-,+), (+,-), (+,+)\}. \tag{3.40}$$

The transition rate matrix for the simple model, Q_{S}, is given in (3.1), and the transition rate matrix for the model with feedback, Q_{F}, is given in (3.2).

Bibliography

[1] J. G. Albeck, J. M. Burke, S. L. Spencer, D. A. Lauffenburger, and P. K. Sorger. Modeling a snap-action, variable-delay switch controlling extrinsic cell death. *PLOS Biol*, 6(12):e299, 2008.

[2] P. F. F. Almeida and W. L. C. Vaz. Lateral diffusion in membranes. *Handbook of Biological Physics*, 1:305–357, 1995.

[3] U. Alon. *An introduction to systems biology: design principles of biological circuits.* CRC press, 2006.

[4] A. Arkin, J. Ross, and H. H McAdams. Stochastic kinetic analysis of developmental pathway bifurcation in phage λ-infected escherichia coli cells. *Genetics*, 149(4):1633–1648, 1998.

[5] N. Q. Balaban, J. Merrin, R. Chait, L. Kowalik, and S. Leibler. Bacterial persistence as a phenotypic switch. *Science*, 305(5690):1622–1625, 2004.

[6] J. I. Belandria. Positive and negative entropy production in an ideal-gas expansion. *EPL*, 70(4):446, 2005.

[7] W. Bialek. Biophysics: Searching for principles. *Q Rev Biol*, 88(2):129–129, 2013.

[8] W. J. Blake, M. Kærn, C. R. Cantor, and J. J. Collins. Noise in eukaryotic gene expression. *Nature*, 422(6932):633–637, 2003.

[9] P. Brémaud. *Markov chains: Gibbs fields, Monte Carlo simulation, and queues*, volume 31. Springer Science & Business Media, 2013.

[10] R. J. Brezski and J. G. Monroe. B-cell receptor. In *Multichain Immune Recognition Receptor Signaling*, pages 12–21. Springer, 2008.

[11] R. E. Campbell, O. Tour, A. E. Palmer, P. A. Steinbach, G. S. Baird, D. A. Zacharias, and R. Y. Tsien. A monomeric red fluorescent protein. *Proc Natl Acad Sci U S A*, 99(12):7877–7882, 2002.

[12] F. C. Collins and G. E. l. Kimball. Diffusion-controlled reaction rates. *J Colloid Sci*, 4:425–437, 1949.

[13] K. Compaan and Y. Haven. Correlation factors for diffusion in solids. *Trans Faraday Soc*, 52:786–801, 1956.

[14] A. Eldar and M. B. Elowitz. Functional roles for noise in genetic circuits. *Nature*, 467(7312):167–173, 2010.

[15] R. J. Ellis. Macromolecular crowding: an important but neglected aspect of the intracellular environment. *Curr Opin Struc Biol*, 11(1):114–119, 2001.

[16] T. S. Gardner, C. R. Cantor, and J. J. Collins. Construction of a genetic toggle switch in escherichia coli. *Nature*, 403(6767):339–342, 2000.

[17] D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *J Phys Chem*, 81(25):2340–2361, 1977.

[18] D. Hall and A. P. Minton. Macromolecular crowding: qualitative and semiquantitative successes, quantitative challenges. *Biochim Biophys Acta*, 1649:127–139, 2003.

[19] N. E. Harwood and F. D. Batista. Early events in b cell activation. *Annu Rev Immunol*, 28:185–210, 2009.

[20] J. Hasty, J. Pradines, M. Dolnik, and J. J. Collins. Noise-based switches and amplifiers for gene expression. *Proc Natl Acad Sci U S A*, 97(5):2075–2080, 2000.

[21] B. Hat, B. Kazmierczak, and T. Lipniacki. B cell activation triggered by the formation of the small receptor cluster: A computational study. *PLOS Comput Biol*, 7(10):e1002197, 2011.

[22] H. Husebye, Ø. Halaas, H. Stenmark, G. Tunheim, Ø. Sandanger, B. Bogen, A. Brech, E. Latz, and T. Espevik. Endocytic pathways regulate toll-like receptor 4 signaling and link innate and adaptive immunity. *EMBO J*, 25(4):683–692, 2006.

[23] M. Kærn, T. C. Elston, W. J. Blake, and J. J. Collins. Stochasticity in gene expression: from theories to phenotypes. *Nat Rev Gen*, 6(6):451–464, 2005.

[24] J. S. Kim and A. Yethiraj. Effect of macromolecular crowding on reaction rates: a computational and theoretical study. *Biophys J*, 96(4):1333–1340, 2009.

[25] H. Kobayashi, M. Kærn, M. Araki, K. Chung, T. S. Gardner, C. R. Cantor, and J. J. Collins. Programmable cells: interfacing natural and engineered gene networks. *Proc Natl Acad Sci U S A*, 101(22):8414–8419, 2004.

[26] M. Kochańczyk, W. Hlavacek, and T. Lipniacki. Spatkin: a simulator for rule-based modeling of biomolecular site dynamics on surfaces. *Bioinformatics*, 33(22):3667–3669, 2017.

[27] M. Kochańczyk, J. Jaruszewicz, and T. Lipniacki. Stochastic transitions in a bistable reaction system on the membrane. *J R Soc Interface*, 10(84):20130151, 2013.

[28] E. Kussell and S. Leibler. Phenotypic diversity, population growth, and information in fluctuating environments. *Science*, 309(5743):2075–2078, 2005.

[29] A. Lipshtat, A. Loinger, N. Q. Balaban, and O. Biham. Genetic toggle switch without cooperative binding. *Phys Rev Lett*, 96(18):188101, 2006.

[30] L. López-Maury, S. Marguerat, and J. Bähler. Tuning gene expression to changing environments: from rapid responses to evolutionary adaptation. *Nat Rev Genet*, 9(8):583–593, 2008.

[31] R. Losick and C. Desplan. Stochasticity and cell fate. *Science*, 320(5872):65–68, 2008.

[32] F. Mancini, M. Marsili, and A. M. Walczak. Trade-offs in delayed information transmission in biochemical networks. *J Stat Phys*, 162(5):1088–1129, 2016.

[33] F. Mancini, C. H. Wiggins, M. Marsili, and A. M. Walczak. Time-dependent information transmission in a model regulatory circuit. *Phys Rev E*, 88(2):022708, 2013.

[34] J. Miękisz and P. Szymańska. Gene expression in self-repressing system with multiple gene copies. *Bull Math Biol*, 75(2):317–330, 2013.

[35] E. W. Montroll. Random walks on lattices. III. Calculation of first-passage times with application to exciton trapping on photosynthetic units. *J Math Phys*, 10(4):753–765, 1969.

[36] B. Munsky and M. Khammash. The finite state projection algorithm for the solution of the chemical master equation. *J Chem Phys*, 124(4):044104, 2006.

[37] B. Munsky and M. Khammash. Guidelines for the identification of a stochastic model for the genetic toggle switch. *IET Sys Biol*, 4:356–366, 2010.

[38] B. Munsky, G. Neuert, and A. van Oudenaarden. Using gene expression noise to understand gene regulation. *Science*, 336(6078):183–187, 2012.

[39] T. Nagai, K. Ibata, E. S. Park, M. Kubota, K. Mikoshiba, and A. Miyawaki. A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nat Biotechnol*, 20(1):87–90, 2002.

[40] P. Nałęcz-Jawecki, P. Szymańska, M. Kochańczyk, J. Miękisz, and T. Lipniacki. Effective reaction rates for diffusion-limited reaction cycles. *J Chem Phys*, 143(21):215102, 2015.

[41] I. Nemenman. Gain control in molecular information processing: lessons from neuroscience. *Phys Biol*, 9:026003, 2012.

[42] J. Pękalski, A. Ciach, and N. G. Almarza. Periodic ordering of clusters and stripes in a two-dimensional lattice model. I. Ground state, mean-field phase diagram and structure of the disordered phases. *J Chem Phys*, 10(114701), 2014.

[43] A. Raj and A. van Oudenaarden. Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell*, 135(2):216–226, 2008.

[44] J. M. Raser and E. K. O'Shea. Noise in gene expression: origins, consequences, and control. *Science*, 309(5743):2010–2013, 2005.

[45] J. Schnakenberg. Network theory of microscopic and macroscopic behavior of master equation systems. *Rev Mod Phys*, 48(4):571, 1976.

[46] N. C. Shaner, P. A. Steinbach, and R. Y. Tsien. A guide to choosing fluorescent proteins. *Nat Methods*, 2(12):905–909, 2005.

[47] A. Singh and M. Soltani. Quantifying intrinsic and extrinsic variability in stochastic gene expression models. *PLOS One*, 8(12):e84301, 2013.

[48] W. K. Smits, O. P. Kuipers, and J.-W. Veening. Phenotypic variation in bacteria: the role of feedback regulation. *Nature Rev Microbiol*, 4(4):259–271, 2006.

[49] S. L. Spencer and P. K. Sorger. Measuring and modeling apoptosis in single cells. *Cell*, 144(6):926–939, 2011.

[50] P. Szymańska, M. Kochańczyk, J. Miękisz, and T. Lipniacki. Effective reaction rates in diffusion-limited phosphorylation-dephosphorylation cycles. *Phys Rev E*, 91:022702, 2015.

[51] P. Szymańska, N. Gritti, J. M. Keegstra, M. Soltani, and B. Munsky. Using noise to control heterogeneity of isogenic populations in homogenous environments. *Phys Biol*, 12(4):045003, 2015.

[52] P. Tolar, H. W. Sohn, W. Liu, and S. K. Pierce. The molecular assembly and organization of signaling active b-cell receptor oligomers. *Immunol Rev*, 232(1):34–41, 2009.

[53] T. Tomé and M. J. de Oliveira. Entropy production in nonequilibrium systems at stationary states. *Phys Rev Lett*, 108(2):020601, 2012.

[54] H. van Beijeren and R. Kutner. Mean square displacement of a tracer particle in a hard-core lattice gas. *Phys Rev Lett*, 55(2):238, 1985.

[55] N. G. Van Kampen. *Stochastic processes in physics and chemistry*, volume 1. Elsevier, 1992.

[56] J.-W. Veening, W. K. Smits, and O. P. Kuipers. Bistability, epigenetics, and bet-hedging in bacteria. *Annu Rev Microbiol*, 62:193–210, 2008.

[57] P. B. Warren and P. R. ten Wolde. Chemical models of genetic toggle switches. *J Phys Chem B*, 109(14):6812–6823, 2005.

[58] L. S. Weinberger, J. C. Burnett, J. E. Toettcher, A. P. Arkin, and D. V. Schaffer. Stochastic gene expression in a lentiviral positive-feedback loop: Hiv-1 tat fluctuations drive phenotypic diversity. *Cell*, 122(2):169–182, 2005.

[59] L. S. Weinberger, R. D. Dar, and M. L. Simpson. Transient-mediated fate determination in a transcriptional circuit of hiv. *Nat Genet*, 40(4):466–470, 2008.

[60] P. J. Żuk, M. Kochańczyk, J. Jaruszewicz, W. Bednorz, and T. Lipniacki. Dynamics of a stochastic spatially extended system predicted by comparing deterministic and stochastic attractors of the corresponding birth-death process. *Phys Biol*, 9(5):055002, 2012.