Imperial College London

IMPERIAL COLLEGE LONDON

DEPARTMENT OF MATHEMATICS

TOPICS IN STATISTICAL PHYSICS OF ACTIVE AND LIVING SYSTEMS

Author:
Ignacio Antonio Bordeu Weldt

Supervisor:
Dr Gunnar Pruessner

A thesis submitted for the degree of

Doctor of Philosophy of Imperial College London

and the

Diploma of Imperial College

October 18, 2019

Copyright Declaration

Declaration of Originality

I hereby certify that the content of this thesis is the product of my own work. The ideas and work of collaborators, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. I have included in Chapters 2 and 3 part of my own published work, for which I hold permission from the publisher.

Ignacio Bordeu

To my wife, Coni, for embarking on this adventure with me

Acknowledgements

I would first like to thank my supervisor Gunnar Pruessner for his tireless support, through thick and thin, during my years as a PhD student. For the countless conversation about life, family and science, for sharing with his students his love for true understanding, and for his uncanny attention for detail that always lead to further insight about the problems at hand.

I would like to thank my collaborators and friends. In particular, I thank Toru Hiratsuka and Fiona Watt, for giving me the opportunity to work by their side, for sharing their knowledge with me and trusting this physicist-turned-biologist. In particular, I want to thank Fiona for mentioning me in the piece she wrote on *The reward of great collaborations* [1] for what I feel very proud. I want to thank my friend Clare Garcin, for teaching me almost everything I know about embryonic stem cells, cell culture, and the patience and resilience one must have when working with stem cells. It was a pleasure to work by your side.

I must thank my friends from Imperial, Saoirse, Ben, Rosalba, Ziluo, Hardik, Fernando, Luca, Adam, Nanxin, Gcina, Lorenzo, Neil, and (my squash partner) Johannes that together with Hannah were the toughest badminton couple to beat, and all the people from the Complexity Centre. Thanks to Harry, Andrew, Lara and all of Cohort 7 at the TSM CDT. Thanks to Jamie and Tung-Jui from King's for the fun times. I should also thank all the Chemical Engineers that I met over the years, for the good times, and for organising the basketball and football matches.

Finally, and most importantly I want to thank family. First, my wife Coni for her love and friendship, for always being there, for begin relaxed, patient, and for always finding fun things to do together. I thank her for sharing these amazing years in London with me, and for the coming years that we will get to share with Matías, the new member of our growing family, and most amazing and beautiful person in the world, who has taught me so much in such a short time. Let the adventure continue! Secondly, I want to thank my family in Chile and France, my siblings, in-laws, nieces and nephew, my dad, and especially my mom for her encouragement to pursue my dreams, and her superhuman strength to overcome the toughest challenges that life throws at her.

I acknowledge the financial support from CONICYT PhD scholarship (Chile), Beca de Doctorado en el Extranjero No. 72160465, and the support of the Centre for Doctoral Training on Theory and Simulation of Materials at Imperial College London, EPSRC (EP/L015579/1).

Abstract

In this thesis, I describe a range of scenarios, from viral propagation to stem cell dynamics, where techniques and ideas from statistical physics provide a route to describe and quantify the observed dynamics. All the systems considered here lie in the realm of out-of-equilibrium physics, where injection and dissipation of energy, matter and/or momentum drive their spatio-temporal evolution.

This thesis is organised such that the topics are presented from *mainly* theoretical to *mainly* experimental. In Chapter 2, I make use of field-theoretic methods to study a branching random walk. This is a paradigmatic process in the study of viral propagation, however analytic results are very limited. Here I show how a field-theoretic approach provides an a route to obtain exact results for the scaling of the volume explored by such a process. In Chapter 3, I show how branching or self-replication can emerge in large scale ecological systems. I show, numerically, how a spatial instability of the vegetation patches gives rise to their self-replication, and discuss the implications for real ecosystems. In Chapter 4, I dive into the realm of cellular biology, where I performed experimental, analytical and numerical work in order to understand the rich dynamics of the spatio-temporal interactions of mouse embryonic stem cells and localized sources of protein signals, and discuss the implications for multicellular organisation. In Chapter 5, I discuss my work on human Keratinocytes, where I studied the interplay between the pulsatile activity of a specific pathway and differentiation. I introduce a method that allows the construction of a phase diagram from the stem cell state. Combined with numerical simulations, this method allowed the visualisation of the temporal relation between the signals to show how transitions between stem cell states occur.

Finally, in Chapter 6, I discuss the main findings of this thesis. I present the general and specific conclusions and point out some key open problems on each sub-field studied.

Contents

List of Figures

List of Tables

List of Publications

Most of the material in this thesis has been published or submitted for publication as detailed in the following:

Chapter 2 The material presented in this chapter has been published [2] in Scientific Reports as the manuscript entitled *Volume explored by a branching random walk on arbitrary graphs*, **I Bordeu**, S Amarteifio, R Garcia-Millan, B Walter, N Wei and G Pruessner.

I contributed to all theoretical calculations and prepared all figures presented here. The main algorithms were written by S Amarteifio, R Garcia-Millan and B Walter. I performed the data analysis, and wrote the manuscript together with G Pruessner.

Chapter 3 The results presented in this chapter have been published [6] in a Special Issue of *Ecological Indicators* as part of a manuscript entitled *Extended patchy ecosystems may increase their total biomass through self-replication*, M Tlidi, **I Bordeu**, MG Clerc and D Escaff.

I contributed to the theoretical calculations and performed all numerical simulations and data analysis. I wrote the manuscript together with M Tlidi, with input from the other co-authors. I prepared all the figures presented here.

Chapter 4 The material presented in this chapter is under review in Physical Review X as the manuscript entitled *Ageing effective potentials in cell-signal interactions*, **I Bordeu** , C Garcin, SJ Habib and G Pruessner.

The experiments were optimized and performed by C Garcin and myself. I developed the methods for data analysis and theoretical model together with G Pruessner. I carried out all analytic calculations and numerical simulations, and prepared all figures presented here. S Habib designed the experiments in discussion with C Garcin, G Pruessner and myself. I wrote the manuscript. All authors discussed and interpreted the data, and commented on the manuscript.

Chapter 5 The material presented in this chapter is in preparation for submission as part of the work entitled *Pulse kinetics of ERK MAPK controls epidermal stem cell states*, T Hiratsuka, **I Bordeu**, G Pruessner, FM Watt.

I developed the data analysis for the pulses, including the construction of phase diagrams and simulations. Interpreted the results and contributed to witting the manuscript. T Hiratsuka performed all experiments and provided me with the figures indicated.

Other work published during my time at Imperial College not discussed in this Thesis are

- [15] *Linear stability analysis of morphodynamics during tissue regeneration in plants*, AM Reijne, **I Bordeu**, G Pruessner, G Sena. *Journal of Physics D: Applied Physics* **52** (8), 084002 (2018).
- [16] *Self-replication of localized vegetation patches in scarce environments*, **I Bordeu I**, MG Clerc, P Couteron, R Lefever, M Tlidi.*Scientific reports* **6**, 33703 (2016). Submitted during my masters degree at Universidad de Chile.
- [17] *From localized spots to the formation of invaginated labyrinthine structures in a Swift–Hohenberg model*, **I Bordeu**, MG Clerc, R Lefever, M Tlidi. *Communications in Nonlinear Science and Numerical Simulation* **29**(1), 482-487 (2015). Submitted during my masters degree at Universidad de Chile.
- [18] *Rodlike localized structure in isotropic pattern-forming systems*, **I Bordeu** and MG Clerc.*Physical Review E* **92**(4), 042915 (2015). Submitted during my masters degree at Universidad de Chile.

Conference articles

- [19] *Finger Dynamics in Pattern Forming Systems.* **I Bordeu**, MG Clerc, R Lefever, M Tlidi. *Springer Proceedings in Physics* **173**. Springer, Cham, (2016). Submitted during my masters degree at Universidad de Chile.
- [20] *Labyrinthine dissipative patterns*, **I Bordeu.** *Journal of Physics: Conference Series* **720**, 012004 (2016). Submitted during my masters degree at Universidad de Chile.

Chapter 1

Introduction

Statistical mechanics is the branch of Physics that deals with systems composed of large numbers of constitutive particles. In general, a microscopic description at the individual particle level of a many-body system is impractical as number of degrees of freedom—say, of positions and momenta—and corresponding number of equations, explodes with the number of particles. For this reason, methods from statistics and probability theory provide a route to bypass such difficulties, by extracting properties of the systems of interest at the population level.

Depending on the energetic properties of the systems of interest, they can be classified into two main classes: (1) equilibrium and (2) non-equilibrium systems. The first corresponds to the standard class of systems in thermodynamic (mechanical and thermal) equilibrium, where there is no external forcing and the systems maintain, on average, a stationary state. Here, there is no in-flux or out-flux of particles, energy and/or momentum.

A system can be driven away from equilibrium by injection of energy, particles and/or momentum. This corresponds to case (2) above. These systems can split into two groups: close-to-equilibrium and far-from-equilibrium. Systems close to equilibrium are systems whose deviation from equilibrium is small, such that linear response theory is valid and the relaxation back to equilibrium can be analytically studied. These type of systems are also the object of study in non-equilibrium thermodynamics, which attempts to make a connection between equilibrium state variables and properties of matter (slightly) out-of-equilibrium.

Far-from-equilibrium systems are driven away from equilibrium, in many cases beyond the linear response regime, where non linear interactions dominate the dynamics, and a rich diversity of spatio-temporal phenomena take place. Non-linear physics and chaos lie at the heart of far-from equilibrium physics. Despite the fact that most natural phenomena lie far from equilibrium, the theoretical understanding is limited, as very few systems are analytically tractable, and universal laws are rare. Living matter is the perfect example of far-from-equilibrium matter, where nutrients serve as the injection of energy and matter required by cell to reproduce, move, and evolve.

At this point it is important to distinguish between two types of matter, passive and active. A passive system corresponds to collection of atoms, molecules or colloids where any injection of energy occurs through the boundaries of the system by exchanging energy, matter and/or momentum with some heat bath. Active systems are composed of active matter, which can transform chemical or other types of energy into kinetic energy. This phenomena, referred to as *self-propulsion*, can be observed in bacteria, sperm, cells and artificial self-propelling particles, such as Janus particles. An active system is then a system where the injection of energy occurs at the individual particle level. For the sake of this thesis any type of particle that can autonomously self-propel, and/or self-replicate, and/or change shape will be referred to as active.

It turns out that addressing any of the usual questions of statistical mechanics, such as the distribution of states in a non-equilibrium system is extremely challenging from a theoretical perspective. In many situations, the question itself is hard to formulate, as not even a free energy can be defined for the system, these are referred to as non-variational systems.

Biological and in particular cellular systems are becoming the new laboratory for performing non-equilibrium statistical physics experiments. With the advent of new technologies and microscopy techniques, biologists are able to extract evermore detailed data on the cellular behaviour, both in-vitro and in-vivo. Scientists have now access to detailed information on the sub-processes involved in the decision making of cells during development, wound healing, and deregulation during disease. The ability to compare models and simulations with real biological data allows the use of the predictive power of physics to push the field of biology forwards. At the same time, the open questions in biology have inspired physicist to come of with new theoretical frameworks to understand the properties of living systems. This is an idealised view on how collaboration between physics and biology can push both field forwards, simultaneously. However, in most cases the situation is different, what are relevant questions for one field will not necessarily be relevant in aother. In this case, there must be a compromise of one of the parties for the sake of science. When problems are approached with an open mind, it is certain that new questions will arise that are of common interest, and can provide inspiration for the further development of science. Part of the work in this thesis supports this idea, where the focus was not on developing a theoretical framework to understand a biological process, but to develop the tools to understand the biological data in the first place. This sets the stage for further theoretical work, which can then be built from the knowledge obtained from the experimental work.

The work presented in this thesis lies in the interface between biology, physics and applied mathematics, and between theory and experiments. The general objective of this thesis is to develop novel approaches to describe and quantify the spatio-temporal dynamics of active and living matter. These approaches can help us find answers to relevant questions in both mathematics and biology with the use of already existing or new tools.

I start off by describing the mathematical properties of the volume explored by a generic spatial branching process, the *branching random walk*. This is a classical model in theory of stochastic processes, however is has applications in biology and active systems. I later describe how branching might emerge in real ecosystems, where the importance of volume exclusion and supercitical behaviour becomes evident. I then use the ideas from the volume explored by a random walker to quantify the effective spatio temporal interactions of stem cells and biochemical signals in vitro. These biochemical signals control the state of stem cells. Inspired by statistical mechanics, I then construct a phase diagram of skin stem cells, which allow me to study the temporal evolution of the cell state, from stem to differentiated cell.

Altogether the methods developed here provide a toolbox for the study of generic temporal and spatio-temporal processes in active and living systems. Some of these techniques rely heavily on theory, while other focus on data analysis and simulations. A good understanding of the experimental data is a steppingstone for later building a comprehensive theoretical model that is truly connected with the real system under study.

1.1 Objectives

In this thesis I make use of techniques and ideas from statistical mechanics in order to accomplish the following objectives:

- To describe and characterise, with the use of tools from equilibrium and non-equilibrium statistical mechanics, phenomena relevant to biological and ecological systems.
- To extend the current knowledge of equilibrium physics to active systems, such as self-replicating (branching) matter and self-organised spatial and temporal processes.
- To develop new methodologies for extracting relevant physical and biological observables from experimental data.

1.2 Structure of the thesis

The results of this thesis are presented in Chapters 2-5.

- In Chapter 2 field-theoretic methods are used to study statistical properties of a critical branching random walk. In this process, active particles randomly explore the environment, and are subject to self-replication and extinction.
- In Chapter 3 a simple theoretical model is introduced to describe interaction between vegetation. This non-linear model has localised solutions, which can become unstable and self-replicate to repopulate landscapes.
- In Chapter 4 experimental work is presented to understand the interactions between individual embryonic stem cells and localised sources of signal. A suitable spatio-temporal correlation function is introduces to show that cell behave as force-field generators.
- In Chapter 5 the relation between a specific molecular pathway and the cell commitment to differentiation is studied in human skin stem cells. Data analysis tools are developed, which allow the construction of a cell state phase diagram, from which information on the causal relation between these processes can be extracted.

Each chapter start with its own introduction, where the theoretical and experimental background and tools are presented. Then the main results for the chapter are presented. Each chapter ends with a discussion of the main findings.

The general discussion and conclusion are presented in Chapter 6, where open questions and future projects are also discussed.

Chapter 2

Field-theoretic approach to the branching random walk

The results presented in this chapter are under review for publication in *Scientific Reports* as part of a manuscript entitled *Volume explored by a branching random walk on general graphs*, co-authored by Saoirse Amarteifio, Rosalba Garcia-Millan, Benjamin Walter, Nanxin Wei and Gunnar Pruessner. Some parts of this chapter are quoted verbatim from the published manuscript [2].

I contributed to all analytic calculations, ran the numerical simulations and performed the data analysis. I prepared all figures, and wrote the manuscript. The numerical algorithms were written by Saoirse Amarteifio, Rosalba Garcia-Millan, and Benjamin Walter in discussion with all authors.

2.1 Summary

A branching processes is a zero-dimensional stochastic model which describes the evolution of particles that can either branch, a birth event by which a particle self-replicates, or decay, a death event by which a particle becomes *extinct* [21]. A branching particles can be imagined as a bacteria in a medium with a small amount of antibiotics, such that it can divide or die as it randomly explores the environment. This prototypical birth-death process has been extensively used to model diverse social and physical phenomena, from the extinction of family names to fission cascades in nuclear reactions [21]. However, most of the natural phenomena of interest, including nuclear processes, viral epidemics in animal populations or social networks require, for a better description, a spatial embedding of this process. Such embedding referred to as a *branching random walk* (BRW) can be constructed by allowing particles undergoing a branching processes to perform a random walk on a lattice. Despite it being one of the simplest models for spatial propagation of disease, amongst many other applications, the statistical properties of the volume explored by the BRW have until now remained unknown, with exact results limited to a one dimensional regular lattice [22].

In this chapter the Doi-Peliti formalism [23, 24] is used to cast the branching random walk in terms of a field theory. The field theory is then used to calculate the scaling of the volume explored by the BRW. It is shown how these results, initially obtained for regular lattices, can be naturally extended to more general graphs, such as fractals, trees and scale-free networks. With this information, the branching random walk is then used to extract information about real social and metabolic networks.

2.2 Introduction

Over the years a number of models to describe viral propagation have been proposed. Effective (non-spatial) models such as the simplest Susceptible-Infectious (SI) and Susceptible-Infectious-Susceptible (SIS) or Susceptible-Infectious-Recovered (SIR) distinguish between two or three populations, respectively [25]. The SI model considers a species of susceptible individual (S), which becomes infectious (I) at a given rate. In the SIS model, infected individuals can go back to the susceptible (s) state. Like viral infections such as the common cold [25], this model allows for a steady state where a fixed fraction of the population remains infected, while the rest remains susceptible. The SIR model introduces a third population, namely a *recovered* (R) population to which infectious individuals can transition. Recovered individuals are effectively *immune*, thus the state where all individuals are immune acts as an absorbing state. In the steady state of the SIR model, all individuals end in the recovered state, this resembles the effect of childhood illnesses such as measles or rubella. There are many variations of these models, which incorporate more species, multiple sub-groups and ageing [26].

With the advances of imaging and tracking technologies in the case of biology, and the access to detailed data sets of social and communications networks, the importance of the substrate on viral propagation processes became evident [27, 28]. Modern biological and theoretical models of disease propagation incorporate spatial interaction by allowing pathogens to multiply at a host cell and be passed to one of the neighbouring cells, at random. It has been argued that this direct cell-cell transmission is an efficient viral propagation mechanisms in tissues [29]. branching random walkers are also allowed to decay, which amount the the recovery or death (depending on the interpretation) of individuals. The individuals, being cells, animals, or communication devices, are represented by single sites of a lattice. At every lattice site, a virus can either multiply (branch, self-replicate) or be degraded by action of the immune system (become extinct, decay). The total number of individuals affected therefore corresponds to the number of distinct sites visited by a BRW, as also argued by Dumonteil et al. [30] for the case of viral propagation in animal populations on two dimensions. In the continuum space limit, the volume explored by the BRW is sometimes referred to as the *Branching Wiener Sausage* [31, 32].

Accounting for space in viral processes is a key step to describe viral propagation in real systems [33]. The first to introduce a spatial process with birth were Fisher [34] to study the propagation of advantageous genes, and Kolmogorov, Petrovsky and Piskunov [35] in the study of the general problem of diffusion with increase of particles. Both of these seminal papers proposed a continuum limit for the two species process described by the chemical reaction

$$X + Y \to 2X, \tag{2.1}$$

which occurs with rate of unity. Here both particle species can diffuse. The equation describing the evolution of the proportion $h(x,t)$ of particles of type X in the system is the well known F-KPP equation

$$\frac{\partial h}{\partial t}(x,t) = \frac{\partial^2 h}{\partial x^2}(x,t) + h(x,t) - h^2(x,t). \tag{2.2}$$

Eq. (2.1) describes the branching of particles of type X. This branching is limited by the presence (or absence) of partiles of type Y. Hence the quadratic saturation appearing in Eq. (2.2), which also ensures that the maximum value of the fraction $h(x,t)$ is unity. The fact that the F-KPP is a deterministic model, implies that fluctuations are neglected. Thus regimes where fluctuations dominate the dynamics cannot be described by these types of models.

One of the earliest attempts to include space into stochastic epidemic models and compare with epidemiologic data was by the statistician Murice Bartlett [36, 37]. He constructed a stochastic model that accounted for person to person transmission in the propagation of measles. The model considered infected individual that could infect a nearest neighbour at a given rate. Infected individuals were removed from the lattice a given rate. New susceptible individuals were introduced at a certain rate, while there was also spontaneous infection of susceptible sites. Bartlett was able to perform numerical simulations on lattices of up to 6×6 sites (the computational limit at the time), and studied the seasonality of measles outbreak in cities of England. This model served as motivation for the study of further spatial epidemic models. In the mid 70's J. Radcliffe [38] considered the so-called *branching Brownian motion* (BBM), first introduced by Ikeda, Nagasawa and Watanabe in the late 1970's [39], as an approximation to Bartlett's model, and studied its convergence. Similarly to the system described by Eq. (2.1), the simpler BBM describes the evolution of the single species process, without extinction

$$X \to 2X, \tag{2.3}$$

where particles of species X diffuse in space. In one dimension, progress was made in the study of range of the BBM, in particular, a duality was found between the range of the BBM and the F-KPP equation [40]. The branching random walk (BRW) considered in this chapter is an extension of the BBM in the sense that particles are also allowed to decay. In recent years, exact calculations of the range of a 1-dimensional BRW were found by Dumonteil et al. [30]. It has been argued that the BRW provides a good approximation for the propagation of disease in animal epidemics in two dimensions [30], where the area explored maounts to the total population of infected individuals. In this paper, the authors suggest that studying epidemics in humans, and human networks becomes more complicated, because of long range interactions due to air travel, among others.

In this chapter, the BRW is not only studied in lattices of arbitrary dimensions, but it is also shown how the approach followed here can deal with long range communications and human networks such as the ones discussed in [30].

As described above, the BRW consists of a system where random walkers produce descendants or decay at rates given. Descendants are random walkers indistinguishable from the parent one, thus they themselves can undergo branching and extinction, while preforming the random walk. The BRW is a spatial extension of a Galton-Watson-type process [41], a non-spatial stochastic process describing the birth/death dynamics of entities. In the context of epidemic models, each lattice site corresponds to a susceptible individual, however in this case susceptible individuals do not get spontaneously infected, disease propagates along nearest-neighbours represented by the random walkers. Distinct sites visited account for the total number of individuals that were ever affected by the disease.

The average number of descendants produced as a result of branching and extinction, controls a transition between a subcritical and a supercritical phase [21]. In the subcritical regime the average offspring number is lower than one, thus on average the number of particles decays exponentially with time, and the disease ultimately infects only a finite number of sites. In contrast, in the supercritical regime the average offspring number is greater than one, thus the average number of particles in the system grows exponentially, and the virus eventually engulfs almost all available space [21, 30]. When the average offspring number is exactly one, the process is said to be in a critical regime. At criticality, none of the processes, branching or extinction, dominate over the other, thus random fluctuations control the dynamics. The expected fraction of distinct sites visited or the size of the epidemic outbreak can be seen as the order parameter of the process.

The characterisation of the statistical properties of distinct sites visited by a BRW is a long-standing problem of branching processes and random walk theory [42, 30, 22]. Exact results have been obtained for one-dimensional systems by Ramola et al. [22]. They found that the in the critical regime, where the branching and death rates are equal, that the distribution $p(a)$ of number of sites visited a scales like

$$p(a) \sim a^{-3} \tag{2.4}$$

for large span a, and scales linearly, i.e. $p(a) \sim a$, for small span. Studying the volume explored by a BRW in one dimension is simplified by the property that, starting from a single infected site, every site between the leftmost and the rightmost reached sites must have been already visited by an active random walker. However, this method is not directly applicable for higher dimensional systems. In two dimensions Dumonteil et al. [30], by providing some symmetry arguments, derived the properties of the area and perimeter of the convex hull of the cluster of visited sites, and related it to the spreading of animal epidemics. They found that, in two dimensions, the area explored, as measured by the convex hull, grows logarithmically in time [30]

$$\langle a \rangle (t) \sim \log(t), \tag{2.5}$$

and its probability distribution scales like

$$p(a) \sim a^{-2}. \tag{2.6}$$

In three dimensions and below, some results have been obtained for a related system, the so-called tree-indexed random walk by Le Gall et al. [43, 44]. However, this system conditions the process to a given offspring number, thus cannot be directly related to the BRW studied in this chapter.

In the next section an alternative route to calculate the scaling properties of the volume explored by a branching random walk is presented. A reaction-diffusion process is introduced, where active (branching) random **walkers** leave behind immobile **tracer** particles [31], like the breadcrumbs left behind by Hänsel and Gretel [45], c.f. Figure 2.1a. As every site can carry at most one tracer, their total number becomes a counter of the number

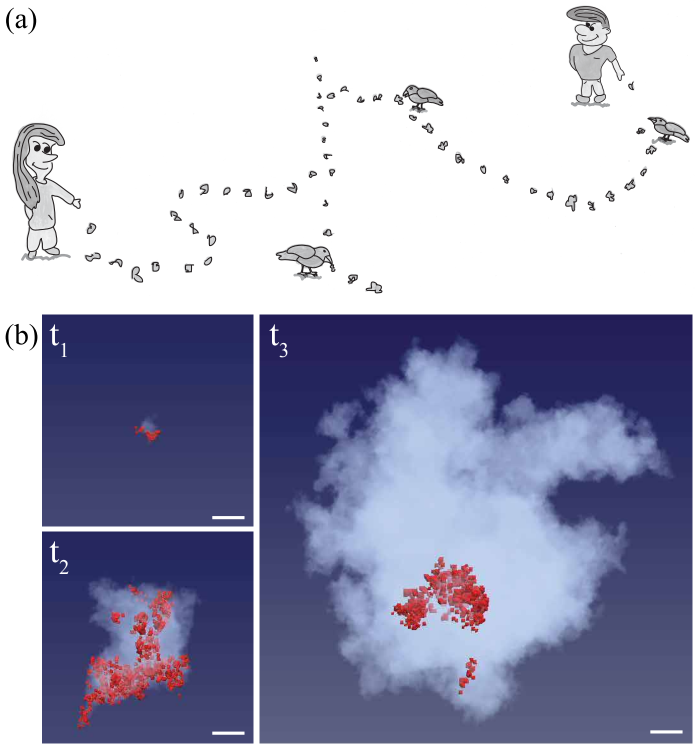


Figure 2.1: **Tracing the path.** (a) The active walkers, Hänsel and Gretel, leave a trace of breadcrumbs along their way to mark the path they have taken. Birds slowly remove the breadcrumbs, as if they were subject to decay (regularisation, see main text). (b) Time evolution of branching random walkers (red) and the cloud of visited sites on a 3d regular lattice at times $t_1 \sim 10^2$, $t_2 \sim 10^3$, and $t_3 \sim 10^4$. Scale bars are equal for all time points. This figure has been published in Sci. Rep. as part of the manuscript [2].

of distinct sites visited, the size of the outbreak. A snapshot of the process in three dimensions can be seen in Fig. 2.1b, which shows the cloud (grey) of visited sites and the branching random walkers (red) on a three dimensional lattice. The master equation of this reaction-diffusion process is then cast into a statistical field theory using the ladder operators introduced by M. Doi [23] and L. Peliti [24] (described in Sec. A). This formalism provides a systematic way for including noise fluctuations and correlations in the calculations of the observables, allowing the study of regimes that would otherwise be ignored. Field theoretic renormalisation in dimensions $d = 4 - \varepsilon$ then allows the derivation of the scaling of the number of distinct sites visited, and the power-law decay of the distribution in arbitrary integer dimensions. Furthermore, it is shown that the results obtained for regular lattices can be extended to general graphs. These findings are supported with numerically simulations in regular lattices, fractals, and artificial and real complex networks.

2.3 Branching random walk

In this section, a mathematical description for the epidemic process is introduced in terms of a reaction-diffusion system. Reactions, such as birth and death take place locally *on-site*, while diffusion emerges from the *off-site* hopping of particles from one site to another. Here, a population of active random **walkers** that hop from their current location $\mathbf{x}$ on a graph to any adjacent site $\mathbf{y}$ with rate H. Additionally, active walkers are subject to two concurrent Poisson processes, namely extinction with rate e and binary branching with rate s. There is no volume exclusion between active walkers, thus during a binary branching event, a single walker at locations $\mathbf{x}$ produces a second identical particle at the same location $\mathbf{x}$, which is indistinguishable from the parent particle. It is branching that renders the particles *active*, as they inject particles to the system spontaneously and at the particle level. This very process distinguishes the BRW from a standard Brownian motion, where there is a fixed number of particles in the system.

The BRW is a *bosonic* process, and can be written in chemical reaction language as

$$\overset{H}{\leftsquigarrow} A \xrightarrow{s} 2A \tag{2.7}$$

$$A \xrightarrow{e} \emptyset \tag{2.8}$$

where the squiggly arrow represents diffusion of particles of species A.

As the observable of interest is the number of distinct sites visited, an immobile **tracer** particle species B is introduced. Tracer particles are deposited by active walkers with rate γ at every lattice site they visit, thereby leaving a trail of tracers behind, similar to the breadcrumbs left by Hänsel and Gretel [45], Fig. 2.1a.

Contrary to the bosonic nature of active walkers, the constraint that at most a single tracer can reside at any given lattice site is imposed. This carrying capacity for tracers implies that the total number of tracers at a given time corresponds exactly to the number of distinct sites visited by the BRW, in the limit of γ going to infinity. Furthermore, a carrying capacity of unity means that the spawning of a tracer is suppressed in the presence of another tracer. As it is shown later, this steric effect generates a series of nonlinear interactions in the field-theory, complicating the calculations significatively. Yet, only with this restriction in place is the number of tracers a measure of the number of distinct sites visited by the walkers. Additionally, for reasons that will become clearer in the field-theoretic formulation, tracer particles B are allowed to become extinct with rate ϵ'. In principle, this would affect the measurement of the number of distinct sites visited, however, the limit $\epsilon' \to 0$ is taken before any observables are calculated. The reactions for tracer particles B can then be expressed as

$$A \xrightarrow{\gamma} A + B \tag{2.9}$$

$$B \xrightarrow{\epsilon'} \emptyset \tag{2.10}$$

As mentioned above, the deposition rate γ has to diverge in order to mark every single site visited by the walkers. However, it turns out that this limit is irrelevant as far as the asymptotic features of this process at large system sizes and long times are concerned [31].

As states earlier, this process can be though as an epidemic spreading on a lattice. In this setup, branching random walkers account for the disease multiplying in a host site and spreading between neighbouring sites. Non-visited lattice sites correspond to susceptible individuals, while visited sites count the number of individuals that have been infected at least once. In the BRW considered here there is no volume exclusion of active walkers, this means that a given site can contain an arbitrary amount of active (infectious) agents, simultaneously. In biological tissues, this can account, for example, for the multiplication of viruses within a cell, which then are transmitted from cell-to-cell. In this context, the total number of sites visited is a natural observable to quantify the range of the viral propagation.

2.3.1 Master equation for the branching random walk

A master equation can be written for the joint probability $\mathcal{P}(\{n\},\{m\};t)$ to find the lattice in a certain configuration $\{n\}$ and $\{m\}$ of occupation numbers $n_{\mathbf{x}}$ and $m_{\mathbf{x}}$ for active walkers and tracers, respectively, at time t

$$\frac{d}{dt}\mathcal{P}(\{n\},\{m\};t) = \dot{\mathcal{P}}_s + \dot{\mathcal{P}}_e + \dot{\mathcal{P}}_{\epsilon'} + \dot{\mathcal{P}}_H + \dot{\mathcal{P}}_\gamma, \tag{2.11}$$

where the terms $\dot{\mathcal{P}}_\bullet = \dot{\mathcal{P}}_\bullet(\{n\},\{m\};t)$ on the right-hand side indicate the contributions from branching s, extinction of active walkers e and tracer particles ϵ', hopping H and deposition γ, respectively. The contribution from branching of active walkers corresponds to

$$\dot{\mathcal{P}}_s(\{n\},\{m\};t) = s\sum_{\mathbf{x}}\Big((n_{\mathbf{x}}-1)\mathcal{P}(\{\dots,n_{\mathbf{x}}-1,\dots\},\{m\};t) - n_{\mathbf{x}}\mathcal{P}(\{n\},\{m\};t)\Big), \tag{2.12}$$

where the sum $\sum_{\mathbf{x}}$ runs through all lattice sites $\mathbf{x}$, and $n_{\mathbf{x}}$ corresponds to the number of active particles at site $\mathbf{x}$. The first term on the right-hand-side (RHS) amounts to the probability of increasing the particle number, though branching with rate s, at the lattice site $\mathbf{x}$ from $n_{\mathbf{x}}-1$ to $n_{\mathbf{x}}$, thus contributing positively to the change in probability $\dot{\mathcal{P}}$. The second term on the RHS corresponds to the probability of increasing the particle number at a given site from $n_{\mathbf{x}}$ to $n_{\mathbf{x}}+1$, thus contributing negatively to the change in probability $\dot{\mathcal{P}}$. In the same way, the contributions to the evolution of the joint probability $\mathcal{P}$ from extinction of active walkers amount to

$$\dot{\mathcal{P}}_e(\{n\},\{m\};t) = e\sum_{\mathbf{x}}\Big((n_{\mathbf{x}}+1)\mathcal{P}(\{\dots,n_{\mathbf{x}}+1,\dots\},\{m\};t) - n_{\mathbf{x}}\mathcal{P}(\{n\},\{m\};t)\Big), \tag{2.13}$$

where e is the extinction rate of walkers. It should be noted that the master equation for the off-lattice binary branching process corresponds to the sum of both Eq. (2.12) and Eq. (2.13), where the summation over lattice sites can be dropped [46].

The contributions to Eq. (2.11) from the spawning of immobile tracer particles by active walkers must take into account the finite carrying capacity $\bar{m}_0$ of each lattice site, which here is set to unity. To account for a finite carrying capacity an effective deposition rate is introduced that decays linearly with the number of tracer particles already present at the site of interest, this may be written as

$$\gamma_{\text{eff}} = \gamma\frac{\bar{m}_0 - m_{\mathbf{x}}}{\bar{m}_0}. \tag{2.14}$$

To study the number of distinct sites visited $\bar{m}_0$ is set to 1. With this constraint in place, each site visited is *marked* with a tracer particle at most once, thus their total number accounts to the number of sites visited by the BRW. With these considerations, the contributions to the master equation from deposition of tracer particles read

$$\dot{\mathcal{P}}_\gamma(\{n\},\{m\};t) = \gamma \sum_{\mathbf{x}} \Big((1-(m_{\mathbf{x}}-1))n_{\mathbf{x}}\mathcal{P}(\{n\},\{\ldots,m_{\mathbf{x}}-1,\ldots\};t) - (1-m_{\mathbf{x}})n_{\mathbf{x}}\mathcal{P}(\{n\},\{m\};t)\Big), \tag{2.15}$$

where $m_{\mathbf{x}}$ corresponds to the number of immobile tracer particles at site $\mathbf{x}$. Analogously to Eq. (2.13), the contribution from extinction of tracer particles reads

$$\dot{\mathcal{P}}_{\epsilon'}(\{n\},\{m\};t) = \epsilon' \sum_{\mathbf{x}} \Big((m_{\mathbf{x}}+1)\mathcal{P}(\{n\},\{\ldots,m_{\mathbf{x}}+1,\ldots\};t) - m_{\mathbf{x}}\mathcal{P}(\{n\},\{m\};t)\Big). \tag{2.16}$$

The terms Eq. (2.12), (2.13), (2.15) and (2.16) correspond to the reaction terms of the BRW reaction-diffusion process, which act locally at every given sites. To account for diffusion we consider that every active walker can hop, with equal rate H, from its current lattice site $\mathbf{x}$, to any of the $q = 2d$ nearest-neighbour sites $\mathbf{y}$ of $\mathbf{x}$ (denoted by $\mathbf{y}.nn.\mathbf{x}$ in the sum), where d corresponds to the dimension of the lattice. With this in place, we can write the contribution to the master equation from hopping as

$$\dot{\mathcal{P}}_H(\{n\},\{m\};t) = \frac{H}{q} \sum_{\mathbf{x}} \sum_{\mathbf{y}.nn.\mathbf{x}} \Big((n_{\mathbf{y}}+1)\mathcal{P}(\{\ldots,n_{\mathbf{x}}-1,\ldots,n_{\mathbf{y}}+1,\ldots\},\{m\};t) - n_{\mathbf{x}}\mathcal{P}(\{n\},\{m\};t)\Big). \tag{2.17}$$

The sum of all contributions described above provides the master equation of the BRW with deposition of immobile tracers. It is noteworthy that the model described by Eq. (2.11) is valid in general graphs, where the topological properties of the graph can be incorporated by introducing a suitable adjacency matrix on the hopping process. The Doi-Peliti formalism can be now used to cast the master equation Eq. (2.11) into a statistical field theory, as described in the next section.

2.3.2 Field theory of the BRW

The steps one must follow to cast the master equation of a reaction-diffusion process in terms of a Doi-Peliti field theory are the following:

1. Construct the master equation for the reaction-diffusion process of interest. In principle, reactions must be local in time and space. However, there are extensions of the formalism for more general scenarios.
2. The master equation is turn into Schrödinger type equation using the ladder operators introduced by M. Doi [23] and L. Peliti [24].
3. The action $\mathcal{A}[a^\dagger, a]$ in terms of ladder operators, is written in terms of time-continuous fields, by associating to each creation operator $a^\dagger$ a corresponding creation field $\phi^\dagger$, and to each annihilation operator a an annihilation field ϕ.
4. It is useful, in some systems, to introduce a *Doi-shifted* creation fields $\tilde{\phi} = \phi^\dagger - 1$. This can simplify the

action, and the subsequent calculation of observables.

5. Once the Doi-shifted field-theoretic action has been obtained, observables can be calculated though perturbative approaches.

Master equation to ladder operators

To cast the master equation for the BRW, Eq. (2.11), in terms of operators, the ladder operators $a^\dagger$, $b^\dagger$, which define the creators, and a and b the annihilators for particle species A and B, respectively, are introduced. Following the step (1) to (3) described above (details in appendix A) the action $\mathcal{A}$ for the BRW, in terms of the ladder operators, may be written as

$$\begin{aligned}
\mathcal{A} =& s\sum_{\mathbf{x}}\Big(a^{\dagger^2}(\mathbf{x})a(\mathbf{x}) - a^\dagger(\mathbf{x})a(\mathbf{x})\Big)\\
&+e\sum_{\mathbf{x}}\Big(a(\mathbf{x}) - a^\dagger(\mathbf{x})a(\mathbf{x})\Big)\\
&+\gamma\sum_{\mathbf{x}}\Big((b^\dagger - 1)a^\dagger a - (b^\dagger - 1)b^\dagger b a^\dagger a\Big)\\
&+\epsilon'\sum_{\mathbf{x}}\Big(b(\mathbf{x}) - b^\dagger(\mathbf{x})b(\mathbf{x})\Big)\\
&-\frac{H}{2q}\sum_{\mathbf{x}}\sum_{\mathbf{y}.nn.\mathbf{x}}\Big(a^\dagger(\mathbf{y}) - a^\dagger(\mathbf{x})\Big)\left(a(\mathbf{y}) - a(\mathbf{x})\right).
\end{aligned} \tag{2.18}$$

It now becomes clear that the Doi-Shift can simplify some of the terms in the action.

2.3.3 Field theory and distinct sites visited

The Doi-shift [47] is performed by introducing the *shifted* creation operators $\tilde{a} = a^\dagger - 1$ and $\tilde{b} = b^\dagger - 1$. In the case of the BRW is useful to go to the continuum limit $\Delta x \to 0$, where Δx is the lattice spacing, and $\sum_{\mathbf{x}} \to \int dx$. In the continuum limit the Doi-shifted operators are mapped into corresponding creation fields $\tilde{a} \to \tilde{\phi}(\mathbf{x}, t)$ and $\tilde{b} \to \tilde{\psi}(\mathbf{x}, t)$, and annihilation fields $a \to \phi(\mathbf{x}, t)$ and $b \to \psi(\mathbf{x}, t)$. The Doi-shifted field-theoretic action may be written in two parts

$$\mathcal{A}_{DS} = \mathcal{A}_0 + \mathcal{A}_{NL}, \tag{2.19}$$

where, the bilinear $\mathcal{A}_0$ and nonlinear $\mathcal{A}_{NL}$ parts are given by

$$\mathcal{A}_0 = \int \mathrm{d}^d x\mathrm{d}t \underbrace{\left[-\tilde{\phi}\frac{d\phi}{dt} + D\tilde{\phi}\nabla^2\phi - r\tilde{\phi}\phi - \tilde{\psi}\frac{d\psi}{dt} - \epsilon'\tilde{\psi}\psi + \tau\tilde{\psi}\phi\right]}_{\mathcal{L}_0(\phi,\psi,\tilde{\phi},\tilde{\psi})} \tag{2.20}$$

$$\mathcal{A}_{NL} = \int \mathrm{d}^d x\mathrm{d}t \underbrace{\left[s\tilde{\phi}^2\phi + \sigma\tilde{\psi}\tilde{\phi}\phi - \lambda\tilde{\psi}\psi\phi - \xi\tilde{\psi}^2\psi\tilde{\phi}\phi - \kappa\tilde{\psi}\psi\tilde{\phi}\phi - \chi\tilde{\psi}^2\psi\phi\right]}_{\mathcal{L}_{NL}(\phi,\psi,\tilde{\phi},\tilde{\psi})}, \tag{2.21}$$

where d represents the spatial dimension of the process, and the Laplacian terms is generated when taking the continuum limit in space of the hopping term Eq. (2.18). The integrands $\mathcal{L}_0$ and $\mathcal{L}_{NL}$ correspond to the bilinear and non-linear part of the Liouvillian $\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_{NL}$. The space and time integrated Liouvillian produces the field-theoretic action $\mathcal{A} = \int \mathrm{d}^d x \mathrm{d}t\, \mathcal{L}$, whose exponential $\mathrm{e}^{\mathcal{A}}$ enters into the path integral formulation. The diffusion constant is $D = (H/q)\Delta x^2$, and $H \propto \Delta x^{-2}$ when the limit $\Delta x \to 0$ is taken, in order to maintain finite diffusivity.

The control parameter of the process corresponds to the *net* extinction rate $r = e - s$. As described in Sec. 2.2, when $r = 0$ the expected offspring number is one i.e, the system is in the critical regime. For $r < 0$, the system is in the supercritical regime, where the number of particles in the system on average increases, while $r > 0$ corresponds to the subcritical regime, were the number of walker on average decays with time. Hereafter, the focus is put on the critical case $r = 0$, where fluctuations dominate the dynamics, and the behaviour becomes unpredictable and highly volatile.

Through field-theoretic renormalisation in dimensions $d = 4 - \varepsilon$ the exact scaling behaviour of the number of distinct sites visited by the walkers can be determined. In the field theory, all large scale phenomena will be controlled by $r \to 0^+$, which corresponds to the onset of epidemics, the limit studied here. The large scale limit is usually referred to as the infrared, as in Fourier space, this limit is obtained by looking a short (infrared) wave-vectors, i.e. the limit $\mathbf{k} \to 0$.

The Liouvillian constructed above is the object that allows the exact calculation of the scaling exponents of the p-th moment of the volume explored by a branching random walk $\langle a^p \rangle (t, L)$, in time t, and the linear system size L. In terms of operators, these observables corresponds to powers of the total number of tracer particles at time t, summed over all lattice sites, given that the system is initialised with a single active walker at position $\mathbf{x}_0$ at time $t = 0$. The first moment, is then obtained by integrating $\left\langle a(\mathbf{x}, t) a^\dagger(\mathbf{x}_0, 0)\right\rangle$ over all space, which in field-theoretic formulation reads

$$\langle a \rangle (t, L) = \int \mathrm{d}^d x \left\langle \psi(\mathbf{x}, t) \phi^\dagger(\mathbf{x}_0, 0)\right\rangle \tag{2.22}$$

$$= \int \mathrm{d}^d x \left\langle \psi(\mathbf{x}, t) \tilde{\phi}(\mathbf{x}_0, 0)\right\rangle, \tag{2.23}$$

where we have used the fact that $\langle \phi \rangle = 0$. The moments of the number of sites are dominated by integrals of the form

$$\langle a^p \rangle (t, L) \sim \int \mathrm{d}^d x_p \ldots \mathrm{d}^d x_1 \left\langle \psi(\mathbf{x}_p, t) \ldots \psi(\mathbf{x}_1, t) \tilde{\phi}(\mathbf{x}_0, 0)\right\rangle, \tag{2.24}$$

or equivalently, by evaluating the Fourier transform at spatial momentum $\mathbf{k} = 0$. These are functions of the couplings introduced above, but to leading order not of the walker's initial position $\mathbf{x}_0$. Here, the brakets $\langle \bullet \rangle$ correspond to the path integral

$$\left\langle \psi(\mathbf{x}_p, t) \ldots \psi(\mathbf{x}_1, t) \tilde{\phi}(\mathbf{x}_0, 0)\right\rangle = \int \mathcal{D}\Pi \left(\psi(\mathbf{x}_p, t) \ldots \psi(\mathbf{x}_1, t) \tilde{\phi}(\mathbf{x}_0, 0)\right) \mathrm{e}^{\mathcal{A}}, \tag{2.25}$$

which measures the p-point correlation function of tracers at $(\mathbf{x}_i, t)$, $i = 1, 2, \ldots, p$ in response to the creation of a walker at $(\mathbf{x}_0, t = 0)$. Here, the integration measure is $\mathcal{D}\Pi = \mathcal{D}\phi\, \mathcal{D}\tilde{\phi}\, \mathcal{D}\psi\, \mathcal{D}\tilde{\psi}$.

2.3.4 Propagators, couplings and observables

As stated earlier, the aim is to calculate the scaling of the moments of the number of distinct sites visited by the critical BRW at long time and large system size limits.

In order to render the Laplacian term in Eq. (2.19) local, the action is rewritten in Fourier space, where the momentum **k** is the conjugate of position **x** and the frequency ω is the conjugate of time t (as defined in appendix B.4). Then the bilinear part of the Liouvillian reads

$$\mathcal{L}_0 = -\mathring{\imath}\omega\hat{\tilde{\phi}}\hat{\phi} + D\mathbf{k}^2\hat{\tilde{\phi}}\hat{\phi} + r\hat{\tilde{\phi}}\hat{\phi} - \mathring{\imath}\omega\hat{\tilde{\psi}}\hat{\psi} + \epsilon'\hat{\tilde{\psi}}\hat{\psi} + \tau\hat{\tilde{\psi}}\hat{\phi}, \tag{2.26}$$

where $\hat{\bullet}$ indicate the Fourier-transformed fields. From Eq. (2.26) we can read directly the propagators for active walkers

$$\left\langle\hat{\phi}(\mathbf{k},\omega)\hat{\tilde{\phi}}(\mathbf{k}',\omega')\right\rangle_0 = \frac{\bar{\delta}(\mathbf{k}+\mathbf{k}')\bar{\delta}(\omega+\omega')}{-\mathring{\imath}\omega + Dk^2 + r} = \tag{2.27}$$


,

where $\bar{\delta}(\mathbf{k}+\mathbf{k}') = (2\pi)^d\delta(\mathbf{k}+\mathbf{k}')$ denotes a scaled d-dimensional Dirac-δ function, and correspondingly for $\bar{\delta}(\omega+\omega')$. Diagrammatically, the propagator is shown as a straight line, where time is to be read from right to left. For the tracers the propagator corresponds to

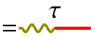
$$\left\langle\hat{\psi}(\mathbf{k},\omega)\hat{\tilde{\psi}}(\mathbf{k}',\omega')\right\rangle_0 = \frac{\bar{\delta}(\mathbf{k}+\mathbf{k}')\bar{\delta}(\omega+\omega')}{-\mathring{\imath}\omega + \epsilon'} = \tag{2.28}$$


,

diagrammatically shown as a wavy line, also to be read from right to left. Both propagators carry a positive mass, $r = e - s$ in Eq. (2.27) and ϵ' in Eq. (2.28). The mass of tracer particles ϵ' in Eq. (2.28) regularises the propagators of the immobile particles in the field theory, guaranteeing causality as the inverse Fourier transform will generate a Heaviside-θ function in time. This is pictorially represented by the birds that eat the breadcrumbs left behind by Hänsel and Gretel's to trace their path (Fig. 2.1a). The limit $\epsilon' \to 0$ is taken before any observable is evaluated.

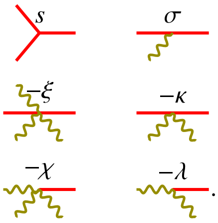
Finally, from Eq. (2.19) the *transmutation* vertex reads

$$\left\langle\hat{\psi}(\mathbf{k},\omega)\hat{\tilde{\phi}}(\mathbf{k}',\omega')\right\rangle_0 = \tau\frac{\bar{\delta}(\mathbf{k}+\mathbf{k}')\bar{\delta}(\omega+\omega')}{(-\mathring{\imath}\omega + \epsilon')(-\mathring{\imath}\omega + Dk^2 + r)} = \tag{2.29}$$



and signals the appearance of a tracer particle (wavy) in response to the presence of a walker (straight), as time is to be read from right to left.

The non-linear part of the Liouvillian, $\mathcal{L}_{NL}$, contributes with six interaction vertices, which diagrammatically read

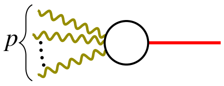


(2.30)

(2.31)

(2.32)

Finally, the observables of the form of Eq. (2.25) have the diagrammatic structure



. (2.33)

Their scaling in time and finite size can be extracted from the scaling of the vertex generating function, which is the standard object of field-theoretic renormalisation. In the next section all possible infrared-relevant interactions are described.

The transmutation rate τ, corresponding to γ on the lattice, and the branching rate s of the active particles (s on the lattice), and diffusion with constant D are the three processes that are expected to govern all infrared behaviour in all dimensions and are therefore assumed to be dimensionless. As it is shown in the following section, these two choices determine the engineering dimension [48] of all other couplings, resulting in ξ, κ and χ being infrared irrelevant. Together with λ, these four couplings are due to the suppression of the spawning of tracers when a site is occupied already. At the upper critical dimension, $d_c = 4$, the coupling λ is marginally relevant, being infrared irrelevant above and relevant below.

2.4 Field-theoretic renormalisation

In the following, the field-theoretical procedure to obtain the scaling of the number of distinct sites visited by the BRW is presented.

2.4.1 Relevant interactions

As discussed in the previous section, the deposition of tracer particles introduces a series of non-linear interaction, which are represented diagrammatically in Eq. (2.30). However, by dimensional analysis arguments it can be shown that beyond dimensions $d = 4$ all such non-linearities become irrelevant, i.e. correlations vanish at long time and large system size. For this reason $d_c = d = 4$ is usually referred to as the *upper critical dimension* of the process. To put this in context, for a non branching random walk, correlations vanish above $d = 2$, thus it is the branching process that *pushes* the critical dimension up.

2.4.2 Dimensional analysis of the couplings

To compute the upper critical dimension of the process described by the Liouvillian $\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_1$, Eqs. (2.20) and (2.21), and to extract the relevant interactions, i.e the couplings that remain relevant in every spatial dimension, the engineering dimensions (here, represented by $[\cdot]$) of every coupling in the action are computed. It is expected that the long range physics in time and space is governed by three processes: diffusion with constant D, branching with rate s, and transmutation with rate τ. Introducing three independent dimensions, namely $\mathtt{A}$, $\mathtt{B}$ and $\mathtt{C}$, and imposing

$$[\tau] = \mathtt{A}, \quad [s] = \mathtt{B}, \quad \text{and} \quad [D] = \mathtt{C}, \tag{2.34}$$

with $[\mathbf{x}] = \mathtt{L}$, $[t] = \mathtt{T}$, and $[\partial_t] = [D\nabla^2]$ it follows that $\mathtt{T} = \mathtt{CL}^2$ is not an independent dimension. As the action, $\mathcal{A} = \int \mathrm{d}^d x \mathrm{d}t \, \mathcal{L}$, itself must be dimensionless, i.e. $[\mathcal{A}] = 1$, we obtain $[r] = \mathtt{T}^{-1} = \mathtt{CL}^{-2}$ and

$$\left[\tilde{\phi}\right] = \mathtt{B}^{-1}\mathtt{CL}^{-2}, \qquad [\phi] = \mathtt{BC}^{-1}\mathtt{L}^{2-d}, \qquad \left[\tilde{\psi}\right] = \mathtt{A}^{-1}\mathtt{B}^{-1}\mathtt{C}^{2}\mathtt{T}^{-2}, \qquad [\psi] = \mathtt{ABC}^{-2}\mathtt{L}^{4-d} \tag{2.35}$$

for the fields in real time and space, such that $[\tilde{\phi}\phi] = [\tilde{\psi}\psi] = \mathtt{L}^{-d}$. The engineering dimensions of the couplings follow:

$$[\lambda] = \mathtt{B}^{-1}\mathtt{C}^2\mathtt{L}^{d-4} \qquad [\sigma] = \mathtt{ABC}^{-1}\mathtt{L}^2 \qquad [\chi] = \mathtt{AL}^d \tag{2.36a}$$

$$[\kappa] = \mathtt{CL}^{d-2} \qquad [\xi] = \mathtt{ABC}^{-1}\mathtt{L}^{d+2}. \tag{2.36b}$$

Setting $\mathtt{A} = \mathtt{B} = \mathtt{C} = 1$, the critical dimension $d_c = 4$ is found, above which all of these interactions become irrelevant. At the critical dimension $d = d_c = 4$ the couplings σ, χ, κ, and ξ remain irrelevant, while λ becomes marginal. To regularise the ultraviolet, short scales, the renormalisation is performed in dimensions $d = 4 - \epsilon < 4$. This regularisation allows the evaluation of loop integrals that would, in $d = 4$, be logarithmically divergent.

To summarize, from the dimensional it is obtained that together with the propagators, the transmutation τ and the interaction vertices s of branching and $-\lambda$ of suppression of spawning represent all relevant couplings at the critical dimension $d_c = 4$.

Performing field theoretic renormalisation in dimensions $d = 4 - \epsilon$ (as detailed in Sec. B.1.1) the scaling of the number of distinct sites visited by the BRW can be obtained. These results are presented in the following section.

2.5 Scaling of the number of sites visited

Following the field-theoretic approach it is obtained that in the thermodynamic limit at long times t, the expected number of distinct sites visited or the volume explored, $\langle a \rangle (t, L)$, scales like $t^{(d-2)/2}$ in dimensions $d < 4$. In dimensions $d < 2$ this volume remains finite in large t. The scaling of the pth moment of the number of distinct sites visited follows,

$$\langle a^p \rangle (t, L) \propto t^{(pd-2)/2} \quad \text{for} \quad Dt \ll L^2 \tag{2.37a}$$

$$\langle a^p \rangle (t, L) \propto L^{(pd-2)} \quad \text{for} \quad Dt \gg L^2 \tag{2.37b}$$

in $d < 4$ provided that $pd - 2 > 0$. The gap-exponent [49] of $\left\langle a^{p+1} \right\rangle / \langle a^p \rangle$ for the scaling in L, which can be thought of as the effective dimension of the cluster of visited sites, is therefore d in dimensions less than $d_c = 4$.

To support these results, numerical simulations of the BRW were performed on regular lattices (details in Sec. B.6). For dimensions $d = 1, 2,$ and 3, a perfect agreement is found as shown in Figs. 2.2a, 2.2b, and 2.2c, where, after an initial transient, the moments scale according to Eq. (2.37) in time and system size. The comparison between the analytic scalings, Eq. (2.37), and the power-law fittings obtained from the numerics are shown in

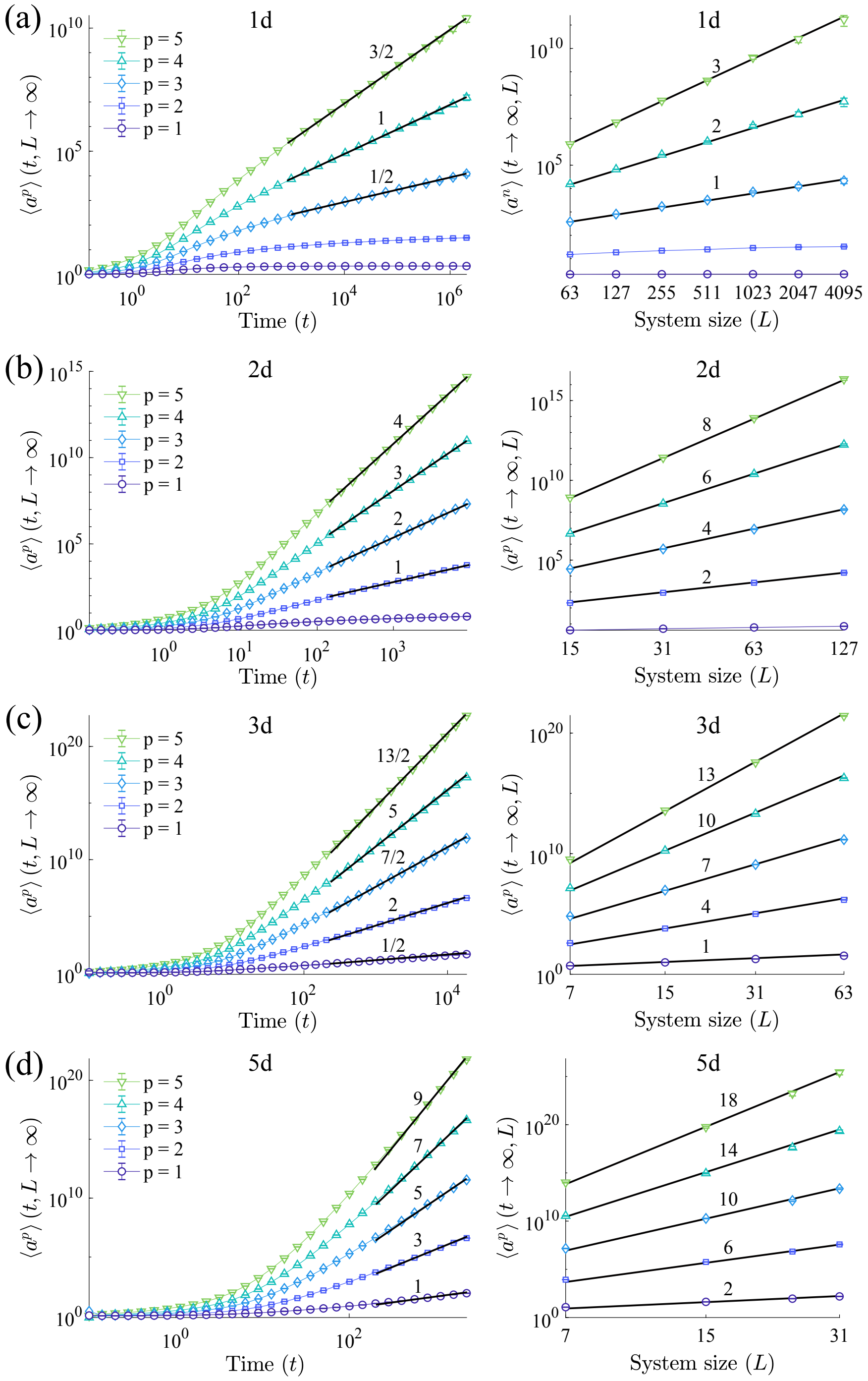


Figure 2.2: **Distinct sites visited by the BRW on regular lattices.** Scaling of the moments of the numbers of distinct sites visited in time (left) and system size (right) for (a) 1d, (a) 2d, (c) 3d, and (d) 5d regular lattices. Solid black lines represent the theoretical exponents given by Eq. (2.37) for $d < 4$, and Eq. (2.38) for $d > 4$. Simulations parameters: $H = 0.1$, $s = e = 0.45$, $\epsilon' = 0$, and $\gamma \to \infty$. This figure has been published in Sci. Rep. as part of the manuscript [2].

exponent	d=1		d=2		d=3		d=5	
	num	theo	num	theo	num	theo	num	theo
α_1					0.47(2)	1/2	1.0(2)	1
α_2			0.98(3)	1	2.0(1)	2	2.9(3)	3
α_3	0.48(4)	1/2	2.0(1)	2	3.5(1)	7/2	4.8(4)	5
α_4	1.0(1)	1	2.9(1)	3	5.0(2)	5	6.7(7)	7
α_5	1.5(1)	3/2	3.9(1)	4	6.4(2)	13/2	9(1)	9
gap	0.5(1)	1/2	1.0(1)	1	1.5(2)	3/2	2.0(5)	2

Table 2.1: **Scaling of visited sites in time.** Scaling in time, $\langle a^p \rangle (t) \sim t^{\alpha_p}$, of the p-th moment of the number of distinct sites visited for regular lattices of integer dimension, d, as indicated. The columns marked *num* shows the numerical results, the columns marked *theo* show theoretical results according to Eqs. (2.37a). The row marked *mean gap* show the average gap-exponent, $\left(\alpha_5 - \alpha_{p_{\text{low}}}\right)/(5 - p_{\text{low}})$, for the corresponding lattice, where p_{low} is the lowest moment displaying algebraic divergence.

Tabs. 2.3 and 2.2.

The process is *free* beyond $d_c = 4$ dimensions, where the probability of any walker or any of its ancestors or descendants ever to return to a previously visited site drops below unity, and the scaling becomes independent of the dimension,

$$\langle a^p \rangle (t, L) \propto t^{2p-1} \quad \text{for} \quad Dt \ll L^2 \tag{2.38a}$$

$$\langle a^p \rangle (t, L) \propto L^{4p-2} \quad \text{for} \quad Dt \gg L^2 \tag{2.38b}$$

with logarithmic corrections in $d = d_c = 4$. The system size gap-exponent,in dimensions greater than $d_c = 4$ is thus 4, as confirmed by numerical observations in dimension $d = 5$ (see Fig. 2.2d and Tab. 2.4).

As correlations become irrelevant (in $d > 4$), this is usually referred to as mean-field behaviour. The set of sites visited may thus be regarded as a four-dimensional object, projected into the d-dimensional lattice considered. Intuitively, these results show that in dimensions $d \leq 4$ a BRW will explore the environment in a compact way, without leaving holes in the cloud of visited sites. Effectively infecting all sites in a neighbourhood. This is a feature of the branching dynamics that allows the process to efficiently cover lattices of dimensions $d = 4$. However, in higher dimensions, branching random walkers can get effectively *lost* in the lattice, due to the large number of dimensions. Here different offspring of the BRW might explore space without ever retracing their path, or even without ever meeting another sibling. This means that in $d > 4$ the cloud of visited sites is not necessarily compact, and might have a more "dendritic" appearance. Furthermore, the results presented in Fig. 2.2 show that higher dimensional lattices ($d > 2$) promote a scaling of the mean ($p = 1$) that grows indefinitely in time, up to finite size effects. This is also true for higher moments, where, for example, the variance, related to the second moment ($p = 2$), scales faster in higher dimension of the lattice, up to dimensions $d = 4$, indicating that a viral process becomes more volatile in higher dimensional systems.

Focusing on dimensions below $d_c = 4$, the distribution of the number of distinct sites visited, a, follows a power law,

$$\mathcal{P}(a) = Aa^{-(1+2/d)}\mathcal{G}(a/a_c) \tag{2.39}$$

exponent	d=1		d=2		d=3		d=5	
	num	theo	num	theo	num	theo	num	theo
β_1					0.97(4)	1	1.9(2)*	2
β_2			2.1(2)	2	3.9(1)*	4	5.7(5)*	6
β_3	0.96(4)	1	4.2(3)	4	6.8(2)*	7	10(1)	10
β_4	1.93(7)	2	6.1(3)	6	9.8(3)*	10	14(2)	14
β_5	2.93(8)	3	8.1(4)	8	12.7(4)*	13	17(3)	18
gap	0.9(1)	1	2.0(3)	2	2.9(2)	3	4(1)	4
fit range	[255, 4095]		[15, 127]		[7, 127]		[7, 31]	

Table 2.2: **Scaling of visited sites by a BRW as function of the system size.** Scaling, $\langle a^p \rangle (t) \sim L^{\beta_p}$, of the p-th moment of the number of distinct sites visited as function of the system size L, for regular lattices of integer dimension d as indicated. The columns marked *num* show the numerical results, the columns marked *theo* show theoretical results according to Eq. (2.37b). The row marked *mean gap* shows the average gap-exponent, $\left(\beta_5 - \beta_{p_{\text{low}}}\right)/(5 - p_{\text{low}})$, for the corresponding lattice, where p_{low} is the lowest moment displaying algebraic divergence. *Goodness of fit < 0.05.

with metric factor A and cutoff $a_c \sim (Dt)^{d/2}$ for $Dt \ll L^2$ and $a_c \sim L^d$ otherwise. This results (see Fig. 2.3a) shows how increasing the dimensionality of the lattice promotes the appearance of larger events, evidencing the relevance of dimension when, for example, attempting to extend *in vitro* (plated, 2D) to *in vivo* (bulk tissue, 3D) conditions in biological systems and material science. The fact that in a 3D tissue the coordination number is higher compared to the 2D case, implies that a virus has the means to spread faster, and there is a higher risk for the virus to extend to a larger fraction of the tissue in the 3D case.

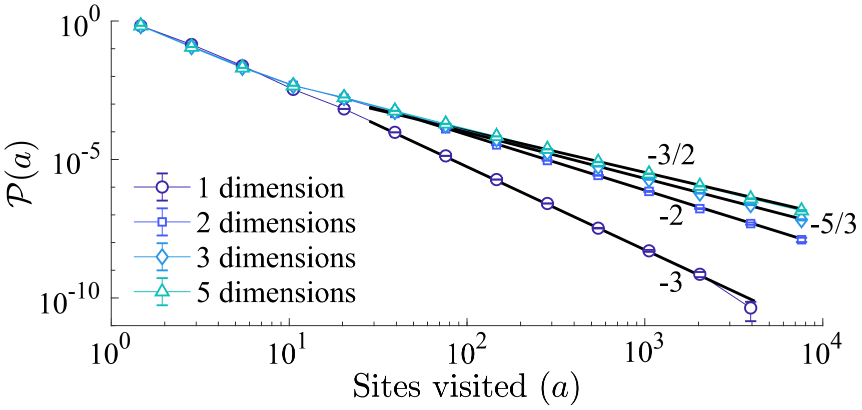


Figure 2.3: **Probability distribution of the number of visited sites** for regular lattices of dimensions $d = 1, 2, 3$ and 5. The solid black lines represent the predicted scaling given by Eq. (2.39). Simulations parameters: $H = 0.1$, $s = e = 0.45$, $\epsilon' = 0$, and $\gamma \to \infty$. This figure has been published in Sci. Rep. as part of the manuscript [2].

In dimensions $d \geq d_c = 4$ the resulting scaling of the distribution is that of Eq. (2.39) at $d = 4$, where the probability distribution decays the slowest, with exponent $a^{-3/2}$. To numerically construct the distribution $\mathcal{P}(a)$ of sites visited, the total number of distinct sites visited by the process was recorded for 10^6 realisations of the process. The numerical results coincide with our theoretical predictions, as shown in Fig. 2.3a.

The exponents found above for $d = 1$ are in agreement with the exact solution by Ramola et al. [22], Eq. (2.4),

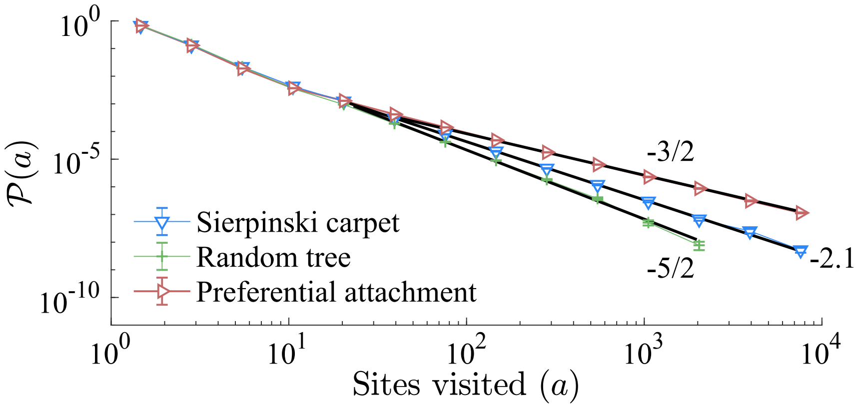


Figure 2.4: **Probability distribution of the number of visited sites** for Sierpinski carpet, random tree, and a preferential attachment (scale-free) networks. The solid black lines represent the predicted scaling given by Eq. (2.39). Simulations parameters: $H = 0.1$, $s = e = 0.45$, $\epsilon' = 0$, and $\gamma \to \infty$. This figure has been published in Sci. Rep. as part of the manuscript [2].

where $\mathcal{P}(a)$ decays as a^{-3}. In two dimensions the power-law tail of the number of distinct sites visited decays as a^{-2}, which coincides with the decay of the 2d convex hull area distribution [30], Eq. (2.6).

2.6 Extension to general graphs

In the field theoretic approach followed to find the scaling in Sec. 2.5 the spatial dimension of the lattice enters only in as far as its spectral dimension is concerned, which characterises the density of eigenvalues of the Laplace operator on the graph considered. The results presented earlier extend naturally to all translational invariant lattices and graphs, by replacing the dimension d of the lattice in Eqs. (2.37), (2.38) and (2.39) by the spectral dimension d_s of the graph, as detailed in Sec. B.5.

This holds true, more generally, as long as the lattice Laplacian itself does not undergo renormalisation, *i.e.* in the absence of an anomalous dimension [50]. In the study of networks the number of nodes N, is a more natural measure of the size of the network than the linear size L. Using $L \sim N^{1/d_s}$ we can write the scaling of the BRW in time and number of nodes as

$$\langle a^p \rangle (t, N) \propto t^{(pd_s-2)/2} \quad \text{for} \quad Dt \ll N^{2/d_s} \tag{2.40a}$$

$$\langle a^p \rangle (t, N) \propto N^{(p-2/d_s)} \quad \text{for} \quad Dt \gg N^{2/d_s}. \tag{2.40b}$$

In this expression the gap-exponent for the scaling in number of nodes always unity, independent of the network. This extension to graphs allowed the prediction of the behavior of the BRW spreading in both artificial graphs relevant for social and biological sciences, and complex systems in general [51, 52, 28, 53], as well as real networks. To illustrate this and validate these results, the BRW was studied on both a Sierpinski carpet (SC) (Fig. 2.5a, methods Sec. B.6), and a random tree network (RTs) (Fig. 2.6a and methods Sec. B.6). Both of these graphs are widely applied in the context of porous media [54] and percolation [55], and have known spectral dimension: $d_s \approx 1.86$

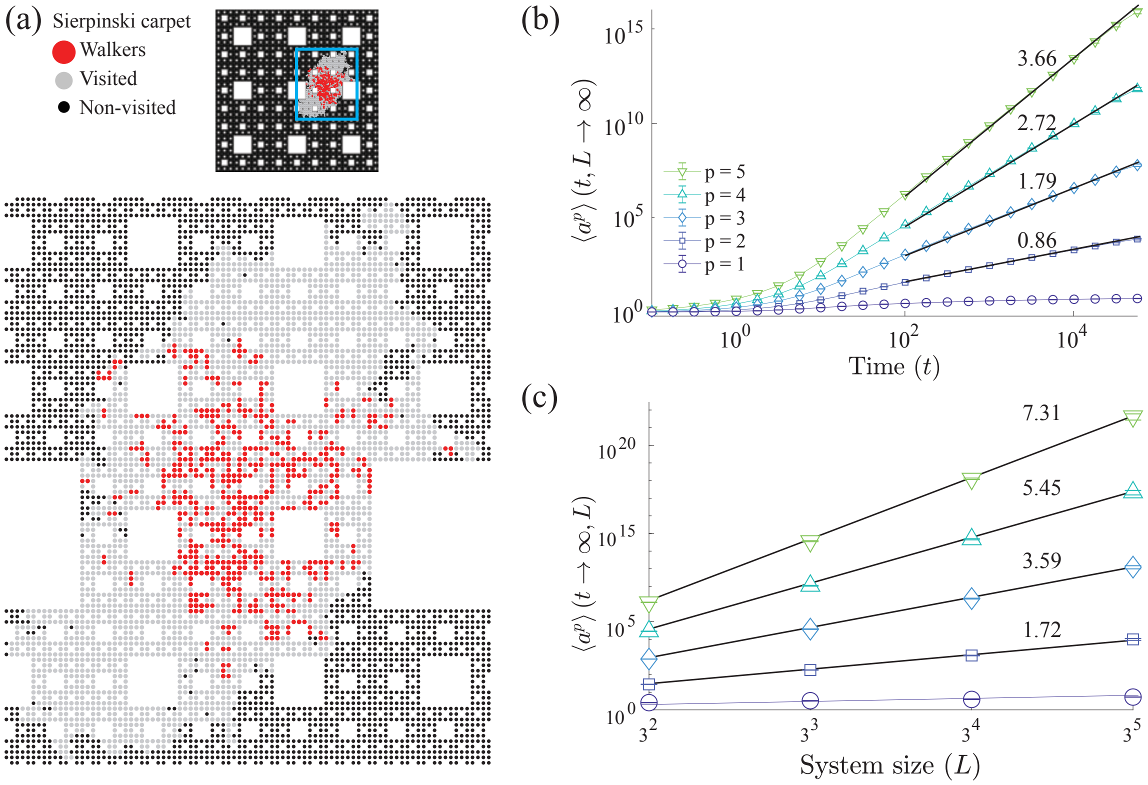


Figure 2.5: **Scaling on general graphs** on the Sierpinski carpet. (a) Shows representative states, with the full Sierpinski carpet shown on inset, indicating walkers (red), visited sites (grey) and non-visited sites (black). (b) Shows the scaling of moments of the number of distinct sites visited as a function of time, and (c) linear system size. The solid black lines represent the predicted scaling from Eq. (2.37). Simulation parameters: $H = 0.1$, $s = e = 0.45$, $\epsilon' = 0$, and $\gamma \to \infty$. This figure has been published in Sci. Rep. as part of the manuscript [2].

for the SC [56], and $d_s = 4/3$ for RTs [57]. Considering Eq. (2.37) with $d = d_s$, for the SC, and (2.40) for the RT accurate predictions for the spreading dynamics were obtained, as confirmed by numerical simulations, Figs. 2.5 and 2.6a. These theoretical predictions extend also to the distribution of visited sites (see Fig. 2.4a), by setting $d = d_s$ in (2.39).

Furthermore, because of their relevance, the spreading of the BRW was studied on a class of scale free networks [58]. Since their introduction, scale free graphs have been observed to describe a plethora of natural phenomena, including the World-Wide-Web [59], transportation [60], and metabolic networks [61], to name but a few. Here, a preferential attachment scheme was considered [58] (Fig. 2.6b, see methods Sec. B.6) to construct networks with power-law degree distribution (Fig. B.1). The existence of a finite spectral gap in these networks, which indicates slow decay of return times of a random walker [62, 63], suggests that the BRW process is bound to exhibit mean-field behaviour, i.e. $d_s \geq 4$. This is confirmed by numerical simulations, where the probability distribution of visited sites (Fig. 2.4a) was found to have a power-law decay with exponent $-1.52(2) \approx -3/2$, and the scaling in time and system size (Fig. 2.6b and Tab. 2.1) follow mean-field behaviour as predicted by (2.40) for $d_s = 4$. In practical terms, this means that a viral process in well connected networks, such as the World-Wide-Web, are

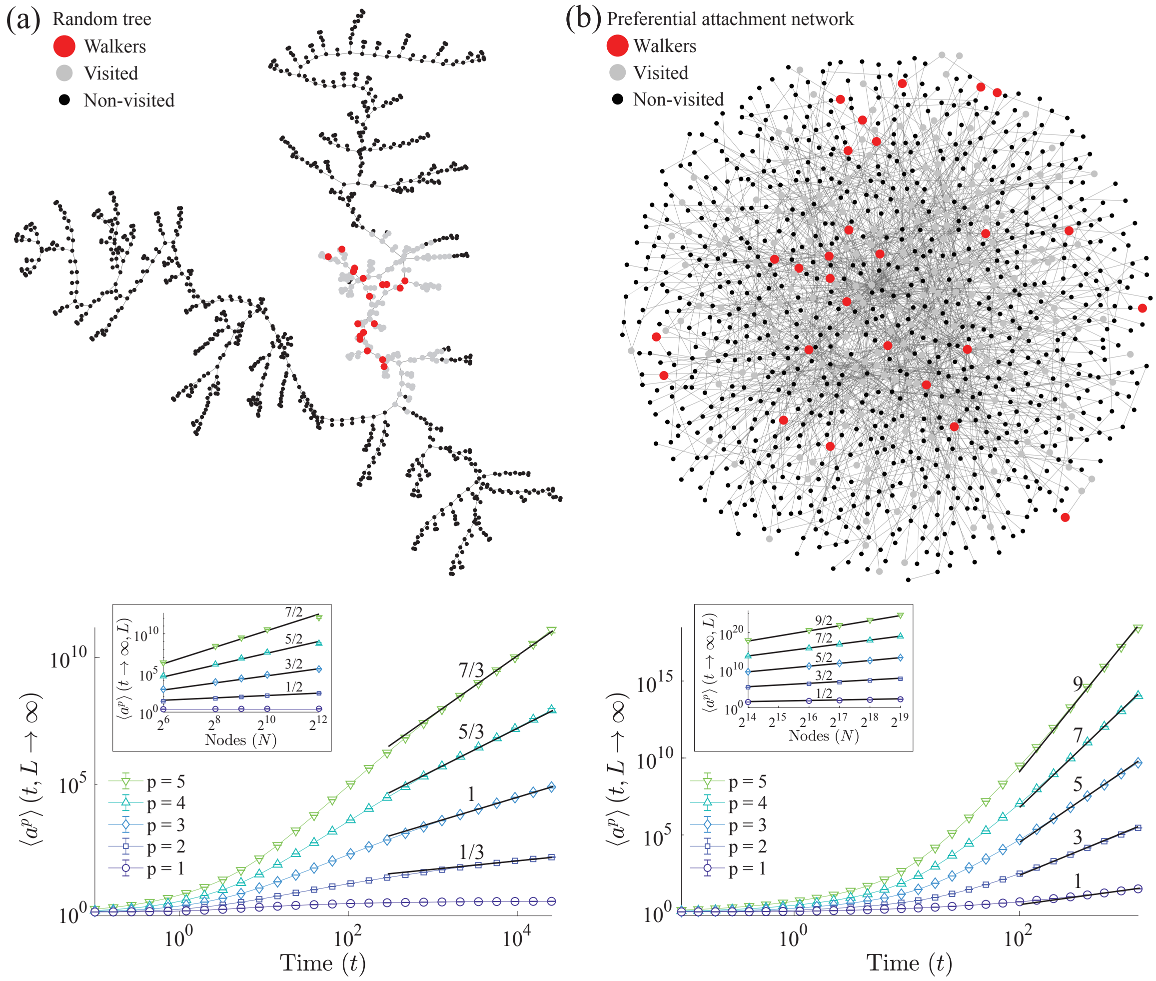


Figure 2.6: **Scaling on general graphs** on (a) random tree and, (b) preferential attachment networks. The top row shows representative states, indicating walkers (red), visited sites (grey) and non-visited sites (black). The bottom row shows the scaling of moments of the number of distinct sites visited as a function of time, and number of nodes. The solid black lines represent the predicted scaling from Eqs. (2.40). Simulation parameters: $H = 0.1$, $s = e = 0.45$, $\epsilon' = 0$, and $\gamma \to \infty$. This figure has been published in Sci. Rep. as part of the manuscript [2].

highly volatile. There, the risk is high that, for example, from a ensemble of viruses and malware, at some of them will infect a large proportion of the network.

The spectral dimension gives information on the behaviour of dynamical processes on graphs. Hence, the BRW can be used to characterise real-world networks through the power-law decay of the distribution of visited sites. For example, the BRW exhibits near mean field-behaviour on a subset of the Facebook network, which has been characterised as scale-free [3]. As a result, the BRW evidences a large effective (spectral) dimension, $d_s = 3.9(1)$, indicating a fast spreading of the viral process on this network.

It should be emphasized that the spectral dimension is sensitive to changes in network topology and connectivity. To exemplify this two publicly available datasets for the yeast protein interaction network were considered. Even though both network describe subsets of the same biochemical network, namely the complete yeast protein

exponent	S.C.		R.T.		P.A.	
	num	theo	num	theo	num	theo
α_1					1.0(1)	1
α_2	0.81(5)	0.86	0.35(7)	1/3	2.8(2)	3
α_3	1.71(7)	1.79	0.9(1)	1	4.8(4)	5
α_4	2.62(10)	2.72	1.6(2)	5/3	6.6(5)	7
α_5	3.54(14)	3.66	2.2(4)	7/3	8.5(9)	9
gap	0.91(9)	0.93	0.6(2)	2/3	1.9(4)	2

Table 2.3: **Scaling of visited sites in time.** Scaling in time, $\langle a^p \rangle (t) \sim t^{\alpha_p}$, of the p-th moment of the number of distinct sites visited for the Sierpinski carpet (S.C., $d_s \approx 1.86$), the random tree (R.T., $d_s = 4/3$), and preferential attachment (P.A., $d_s > 4$) networks. The columns marked *num* shows the numerical results, the columns marked *theo* show theoretical results according to Eqs. (2.37a) with d replaced by the spectral dimension d_s. The row marked *mean gap* show the average gap-exponent, $\left(\alpha_5 - \alpha_{p_{\text{low}}}\right)/(5 - p_{\text{low}})$, for the corresponding graph, where p_{low} is the lowest moment displaying algebraic divergence.

exponent	S.C.		R.T.		P.A.	
	num	theo	num	theo	num	theo
β_1					0.49(1)	1/2
β_2	1.9(1)	1.72	0.58(6)	1/2	1.49(1)	3/2
β_3	3.8(2)	3.59	1.6(1)	3/2	2.49(2)	5/2
β_4	5.7(3)	5.45	2.6(2)	5/2	3.49(2)	7/2
β_5	7.5(4)	7.31	3.7(2)	7/2	4.49(2)	9/2
gap	1.9(3)	1.86	1.0(1)	1	1.00(2)	1
fit range	[9, 243]		$[2^6 - 1, 2^{12} - 1]$		$[2^{14} - 1, 2^{19} - 1]$	

Table 2.4: **Scaling of visited sites by a BRW as function of the system size.** Scaling, $\langle a^p \rangle (t) \sim L^{\beta_p}$, of the p-th moment of the number of distinct sites visited as function of the system size L, for the Sierpinski carpet (S.C., $d_s \approx 1.86$), the random tree (R.T., $d_s = 4/3$), and preferential attachment (P.A., $d_s > 4$) networks. The columns marked *num* show the numerical results, the columns marked *theo* show theoretical results according to Eq. (2.37b) for the S.C. (with d replaced by the spectral dimension d_s), and according to Eq. (2.40) for random tree (R.T.) and preferential attachment (P.A.). The row marked *mean gap* shows the average gap-exponent, $\left(\beta_5 - \beta_{p_{\text{low}}}\right)/(5 - p_{\text{low}})$, for the corresponding graph, where p_{low} is the lowest moment displaying algebraic divergence. *Goodness of fit < 0.05.

interactome, the spectral dimensions in both cases were found to be significantly different, $d_s = 3.0(1)$ for the network with $N = 1870$ nodes [4], and $d_s = 3.8(1)$ the for the larger network of $N = 2559$ nodes [5], leading to differences in properties of the spreading process among the two. The discrepancy points to differences in the connectivities of both networks and shows the importance of having access to the complete network in order to provide a reliable analysis of their properties, which may have biological implications [64, 65].

2.7 Generalisation to *k* offspring

The results presented above for the binary branching process, where walkers branch into exactly two new walkers, apply equally to more general branching processes, where the number of offspring in each birth event is given by a distribution. This can be seen, for example, in real-world scenarios where a single infected individual or device infects a whole neighbourhood around them, or in the case of signal propagation in protein networks, where the activation of one node (or chemical reaction) can activate a whole fraction of its neighboring nodes.

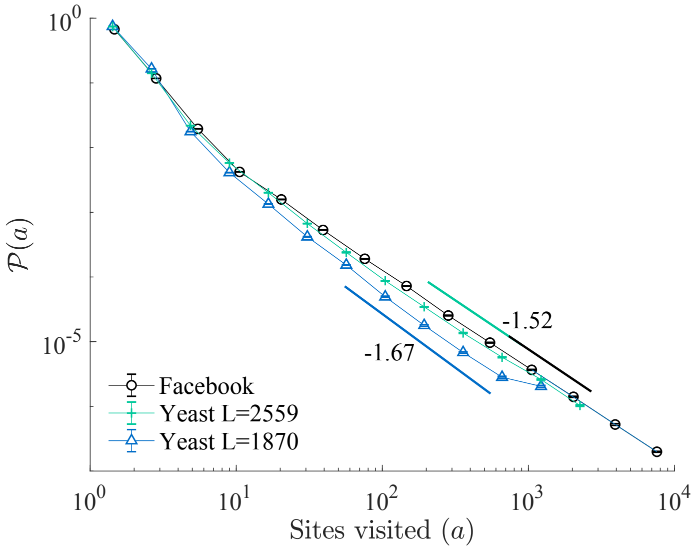


Figure 2.7: **Probability distribution of number of distinct sites visited** $\mathcal{P}(a)$, for the Facebook network (L = 63730 nodes) [3], and yeast protein interaction networks with $L = 1870$ [4], and $L = 2559$ nodes [5]. According to Eq. (2.39) the power-law decay of $\mathcal{P}(a)$ is $a^{-(1+2/d_s)}$ provided $d_s \leq 4$ suggesting a spectral dimension $d_s \sim 3.83(6)$ for the large subset of the yeast protein network, $d_s = 2.97(5)$ for the small subset, and $d_s = 3.85(6)$ for the Facebook network. This figure has been published in Sci. Rep. as part of the manuscript [2].

Considering the case where the offspring number is a random number, it is shown here that this process lies in the same universality class as binary branching [66, 67]. Instead of two distinct processes for branching into two active walkers (with rate s above) and getting extinguished (with rate e above), the latter can be considered as branching into $k = 0$ walkers and the former generalised to branching into any number k of walkers. Each of these processes may occur with rate σ_k, which can always be written as $\sigma_k = \sigma p_k$ with p_k the normalised probability for branching into k walkers and σ the rate with which any such processes take place.

The two contributions $\mathcal{P}_s$, Eq. (2.12), and $\mathcal{P}_e$, Eq. (2.13), are thus subsumed and generalised by

$$\dot{\mathcal{P}}_c(\{n\}, \{m\}; t) = \sigma \sum_{k=0}^{\infty} \sum_{\mathbf{x}} p_k \Big((n_{\mathbf{x}} - k + 1) \mathcal{P}(\{\ldots, n_{\mathbf{x}} - k + 1, \ldots\}, \{m\}; t) - n_{\mathbf{x}} \mathcal{P}(\{n\}, \{m\}; t) \Big), \qquad (2.41)$$

which allows for p_1, but the process of branching into a single particle has no bearing on the master equation.

In the field theory, the mass of the propagator for active walkers becomes [46]

$$r = -\sigma \sum_{k=0}^{\infty} p_k (k-1) = \sigma(1 - \bar{k}), \qquad (2.42)$$

where $\bar{k} = \sum_{k=0}^{\infty} p_k k$ is the average offspring number, which again, defines a subcritical ($r > 0$), a critical ($r = 0$), and a supercritical ($r < 0$) regime.

In the case of generalised branching, the non-linear part of the action contains contributions of the form $\tilde{\phi}^k \phi$ for all $k \geq 2$ as soon as there is any $k \geq 2$ with $p_k > 0$ [46]. However, terms with $k > 2$ turn out to be infrared irrelevant as their couplings have dimension $\mathsf{B}^{k-1}\mathsf{C}^{2-k}\mathsf{L}^{2(k-2)}$. The field theoretic results above for binary branching therefore govern also branching processes with generalised offspring distribution.

2.8 Discussion and outlook

In this chapter, it has been shown how field theoretic methods provide a versatile and powerful method to study reaction diffusion processes, and to extract otherwise inaccessible observables of interest.

With the use of field-theoretic renormalisation, the scaling properties of the number of distinct sites visited by a branching random walk where obtained, which are relevant for a wide range of problems. This process can be interpreted as the simplest model for spatial viral propagation, where lattice sites correspond to susceptible individuals and active branching walker model model the viral spreading. For this reason, the results presented here are of particular importance for the study of viral propagation in 3d tissues and high dimensional networks, such as the internet. In this context, Eq. (2.37) describes the rapid propagation of the average number of infected individuals $\langle a \rangle \sim t^{d/2-1}$ in dimensions $d > 2$. At the same time, the fat-tailed distribution of visited sites, Eq. (2.39), (exponent $-1 - 2/d_s$) results in the high variability of the viral spread at the onset of an epidemic. This onset, corresponds to the critical value of parameters ($s = e$), where a viral process transition from a controlled (subcritical) disease, to a large scale (supercritical) epidemic.

The fact that the BRW does not consider volume exclusion for the infection (active particles), means that an infected individual (site) can carry an arbitrary number of active particles at a given time. Moreover, infected (visited) sites are not removed from the system, as they would on a standard SIR model, and are not different, as far as the active particles are concerned, from non-visited sites. This mean that an infected individual can be re-infected (revisited) not only an arbitrary number of times but also arbitrarily fast. This is different, for example from an SIS model, where infected individuals would become susceptible again at a given rate, preventing immediate re-infection. With these differences in mind, the BRW can be though as a simplified description of viral transmission, multiplication and deletion in a population of hosts. As discussed in previous sections, in a cellular tissue, a virus might multiply in a host cell, and be transmitted to a neighbouring one by direct cell-to-cell contact. This process, in principle, does not prevent the virus from returning to the original host, and does not limit the number of viral units within each host. In this context the number of sites visited, corresponds exactly to the number of cells ever affected by the virus, and the results presented here provide theoretical bounds for the spreading speed, fluctuations, and higher moments.

The field theoretic approach followed in this chapter provides a quantitative measure to explore and determine the spectral dimension of artificial and real networks. This is useful when the spectral dimension is greater or equal to 2, where the traditional approach of exploring graphs, based on simple random walks [68, 63], fails to extract their spectral properties. The predicted relation between dimensionality of the graph and the propagation dynamics can contribute to the design of more robust and reliable information and chemical networks, as variations in a networks topology directly affects the spreading properties. In summary these results shed new light on the properties of spatial branching processes on general graphs, and their applicability in the study of real complex networks, and provide observables of broad interest for the characterisation of real world lattices, tissues, and networks.

In the next chapter, branching is studied in a two dimensional system. In that case a mean field approach is followed, as the dynamics of interest is not dominated by the fluctuations. A F-KPP-type equation is introduced to model the super-critical branching of particle-like solutions.

Chapter 3

Branching in ecology: the case of patch self-replication

The results presented in this chapter have been published in a Special Issue of *Ecological Indicators* as part of a manuscript entitled *Extended patchy ecosystems may increase their total biomass through self-replication* in collaboration with Mustapha Tlidi, Marcel G.Clerc and Daniel Escaff.

I contributed to all analytic calculations, performed the numerical simulations and data analysis. I prepared all figures presented here, and wrote the manuscript together with M. Tlidi, with input from all coauthors.

3.1 Summary

The branching random walk studied in Chapter 2 corresponds to the simplest spatial extension of a branching process, where volume exclusion is absent. Thus active particles do not interact. Branching phenomena like the one discribed earlier can originate, in real systems, from a variety of self-organising processes. At the nano-scale molecular cooperation can give rise to replication of polynucleotides [69]. At the microscopic, cellular, level molecular pathways activate an intricate machinery that can initiate mitosis, the process by which living cells self-replicate. At a much larger scale, such as vegetated landscapes, it has been proposed that non-local interactions between individual plants can lead to spatial self-organisation and patterning of real landscapes [70]. For many of these processes, volume exclusion plays a role in shaping the spatial distribution of particles or patterning. In this chapter it is shown that non-local interactions allow for the formation of localised vegetation patched, which correspond to dense clusters of trees, shrubs or grass. Building up on previous work (Bordeu et al. [16]), it is shown that the circular shape of vegetation patches may become unstable, deforming, subsequently splitting into two new, indistinguishable patches. These patches naturally repel each other, as if they where subject to volume exclusion. This growth mechanism, is often referred to as *self-replication*. The self-replication process can continue until the whole system is covered by patches of vegetation. It is also shown how the *reverse* process, by which a homogeneous vegetation cover destabilises, can lead to the formation of a patched landscape. By relating the parameters of the theoretical model with the resource availability or *aridity*, a connection can be made with desertification and the dynamics of real ecosystems. With the support of satellite images, it is argued that the organisation of real ecosystems is indeed a consequence of self-organisation, and that self-replication might be a real phenomena at large scales.

3.2 Introduction

In 1952 Alan Turing first proposed the ideas of reaction-diffusion systems and pattern formation [71], and what was later termed *Turing instability*, by which a homogeneous solution of a dynamical system can lose its stability via a finite wavelength perturbation, giving rise to a spatially extended self-organised pattern. Since then, our knowledge on the theoretical origins of pattern formation and selection have increase considerably [72]. Pattern formation mechanisms have been proposed to control the formation of animal skin prints [73], the spatial organisation of sand dunes [74], and the organisation of large scale ecosystems [70], to name but a few.

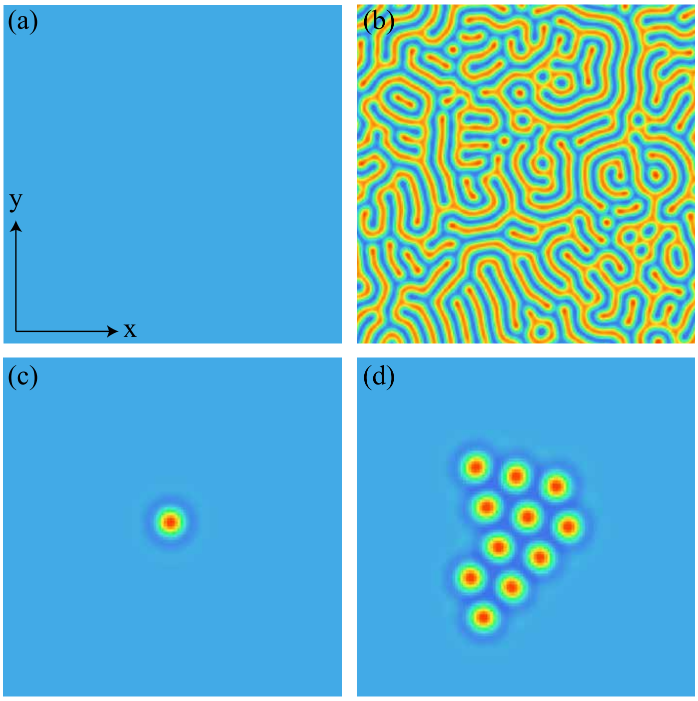


Figure 3.1: **Two dimensional extended patterns and localised structures.** (a) Shows a stable homogeneous solution, and (b) the pattern state resulting from a Turing instability of the homogeneous state in (a). (c) and (d) show a localised solution and a bound state, respectively. Colours indicate the amplitude of the field from low (blue) to high (red).

The Turing mechanism explains the formation of system-wide (extended) patterns, where the whole system transitions from an homogeneous state (Fig. 3.1a) to a pattern solution (Fig. 3.1b). Localised structures (LSs), on the other hand, correspond to solutions that usually appear as circular *bumps* or *dips* with respect to the homoge-

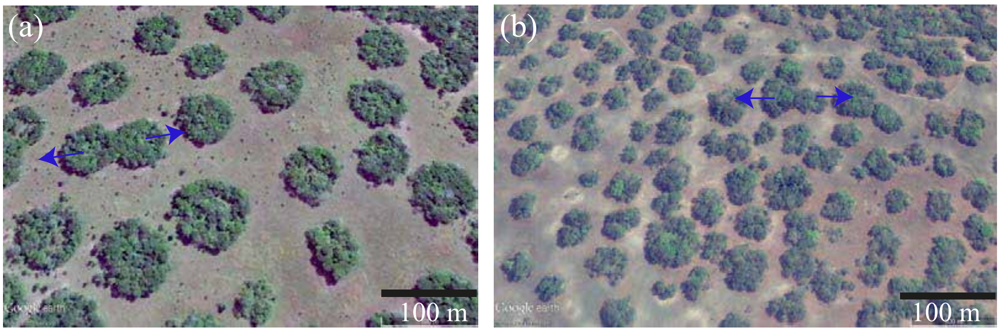


Figure 3.2: **Satellite images of vegetation patches** in (a) the Mufumbwe District in the North-Western Province of Zambia [13°46′39.83″S, 25°16′39.59″E], and (b) the Fombeni, Mozambique [18°41′02.17″S, 35°31′55.95″E]. The images were obtained using *Google Earth Pro*. The blue arrows indicate overlapping patches, possibly undergoing self-replications. This figure, for which I hold the copyright, has been adapted from [6].

neous state that supports them (Fig. 3.1c). There is a large zoo of localised structures, including structures with broken azimuthal symmetry (Bordeu and Clerc [18]) and structures that support permanent dynamics [75]. LSs are particle-like solutions, they interact, and can form bound states (see Fig. 3.1d).

In this chapter, a pattern forming model is introduced to describe the interaction and propagation of vegetation in scarce environments. In the context of vegetation, bare terrain (with zero total biomass) and the uniformly vegetated states (maximal biomass) correspond to two homogeneous solutions of the system. Localised structures represent patches of vegetation if they are supported by the bare state, or gaps if they appear as holes on the vegetated state. Spontaneous transitions between uniform vegetation and fragmented ecosystem, where bare terrain coexists with vegetation, has been widely observed in real ecosystems around the globe, examples are tiger bush [76], fairy circles [77] and labyrinthine patterns ([78] and Bordeu [20]).This transition may occur due to changes in the ecosystem, such as nutrient limitation, water deprivation or changes in the interaction between vegetation. It is widely accepted that facilitative and competitive interactions between individual plants, together with a seed dispersal mechanism provide sufficient conditions for the formation of large scale vegetation patches and extended patterns [76, 79, 80, 81, 77, 82]. In real ecosystems, the spatial properties of the observed structures depend on climate, seasonality, nutrient availability, terrain anisotropy, or anthropogenic and animal influence [83, 84, 85, 86]. Quantitative studies based on field observations have been made on the Sahelian gapped patterns formed by Combretum micranthum (trees) [87, 88, 89], patches of vegetation in arid high altitude environments in the tropical alpine ecosystems of the Andes formed by Festuca orthophylla (Bordeu et al. [16]), and for grasses and Pycnophyllum cushions in Bolivia [90].

It has been recently shown, in a simplified model for vegetation dynamics, that localised patches can be destabilized by changes in aridity, where circular LSs elongate, leading either to the formation of labyrinthine patterns (Bordeu [20]) or to self-replication (Bordeu et al. [16]). The self-replication mechanism allows a single localised structure to transition into an extended, hexagonal, periodic pattern of vegetation. During this transition the total biomass in the system increases monotonically as newly formed patches contribute to the re-population of the territory accessible to vegetation. From a theoretical point of view, self-replication is a well known patterning phenomenon in chemical reactions [91, 92, 93, 94], it has also been reported for models in nonlinear optics [95] and vegetation dynamics [96, 16].

Here a general integro-differential model is considered, where non-local interactions are taken explicitly into account. The non-localities arise from the facilitating mechanisms, i.e. promoters of vegetation growth, competition mechanisms, which limit vegetation growth, and dispersion effects. This model corresponds to a variant of the theory of vegetation patterning established by R. Lefever [76], which focuses on the relationship between the structure of individual plants and the facilitation-competition interactions existing within plant communities.

Through numerical simulations it is shown how self-replication provides a mechanism for scarce resource ecosystems, populated with localised patches of vegetation, to transition to higher biomass states, ending in hexagonal patterns of vegetation patches. Moreover, it is shown that this kind of patterns may be obtained through the decay of a homogeneous vegetated landscape towards a less populated fragmented state, where hydric stress induces the contraction of vegetated areas. Depending on the levels of aridity, the ecosystems may decay to different types of patterned states, or even collapse to the bare, desert state.

The characteristics of both self-replication (forward) and fragmentation (backwards) processes, are studied analytically and numerically. It is shown that, depending on the levels of aridity, localised patches can be more or less stable than the periodic pattern, a phenomenon previously studied in simpler local models [97]. In an attempt to relate the theoretical observations presented here with real observations, satellite images from Zambia and Mozambique are considered. These landscapes exhibit large vegetation patches, each in the range of $10-50$ meters in diameter (see Fig. 3.2). Statistical analysis of the satellite images is then preformed to shows that the patch sizes correlates with the inter-patch distances, which together with the existence of a characteristic wavelength in the system supports the hypothesis that the vegetation self-organises into the observed configurations. This opens the possibility that patch self-replication is effectively taking place in these ecosystems.

This chapter is organized as follows: in Sec. 3.3 the theoretical model is introduced. In Sec. 3.4 the results are presented, which include the linear stability analysis of the model, numerical analysis of the self replication process, the analysis of real spatial patterns. The discussion and outlook are left for Sec. 3.5.

3.3 The model

Modeling vegetation dynamic is a challenging and complex problem. The approach considered depends on the level of description one wants to achieve, the different approaches can be classified into three categories [98]. The first class considers the modeling of vegetation as front dynamics with non-local interactions. The second class considers the local coupling of above-ground vegetation with underground and/or above-ground water transport in the form of a coupled reaction-diffusion system. The third class corresponds to a stochastic approach, where the main ingredient are the environmental fluctuations that trigger the patterning of the vegetation landscape.

In this chapter, the focus is on the relationship between the above-ground facilitation mechanisms and the below ground (root-driven) competitive interaction of vegetation clusters. For this reason, the first approach is adopted, where the order parameter is the vegetation density and the single-plant level, and a there is no requirement of additional state variables [98]. The model considered is a variant of the theory of vegetation patterns established by R. Lefever et al [76] in the late nineties. The spatio-temporal dynamics of vegetation is described in continuous space and time, facilitative and non-local competitive plant-to-plant interactions are introduced through exponential interaction kernels [76, 77, 89, 98]. The model described a single vegetation species, and assumes a flat landscape with isotropic and homogeneous environmental conditions. Further, it assumes that all plants are mature, i.e. age classes are neglected. This approximation can be justified by the fact that individual plants grow on much faster time scale compared to the time scale of the formation of regular vegetation pattern, thus plant maturation is considered instantaneous.

The order parameter of the system corresponds to the vegetation biomass density $b = b(\mathbf{r}, t)$, which is a function of the position $\mathbf{r}$ and time t. The evolution of the biomass density b is given by [77, 89]

$$\frac{\partial b}{\partial t} = \underbrace{b[1-b]M_f(\mathbf{r},t) - \mu b M_c(B(\mathbf{r},t))}_{G(b,B)} + D\nabla^2 b, \tag{3.1}$$

where the phenomenological parameter μ measures the biomass loss/gain ratio as measured for a single, isolated plant, which can be viewed as an indirect measure of resource scarcity or stress that limits net biomass production. In what follows μ is referred to as the *aridity* parameter. The two terms of the function $G(b, B)$ in Eq. (3.1) account for the plant-to-plant facilitation and competition feedback, respectively. They describe the spatial extension of feedback effects in terms of the characteristic ranges L_f and L_c over which facilitative and competitive interactions operate. The facilitative interactions take place on the level of the aerial plant structure (crown), they involve sheltering, litter, water funneling or any other effect that contributes to the biomass natural growth [99]. On resource-limited environments, plants compete for their survival, in face of climate change and increasing drought periods, plants must adapt their root structures to overcome resource scarcity [100, 87]. Through their root structures (rhizosphere), each individual plant tends to deprive its neighbors of vital resources, such as water or nutrients, opposing the facilitation mechanisms and impeding vegetation growth. Field measurements of roots lateral spread indicate that they may extend up to an order of magnitude further than the aerial structure (crown) [87, 88, 89, 90]. Some observations suggest and exponential decay with distance of the competitive interactions [88]. The diffusion term in Eq. (3.1) describes the spatial propagation of vegetation via seed dispersion with diffusion constant D, which measures the strength of dispersive processes. The competitive plant-to-plant interaction are assumed to be of the form $M_c = \exp(\chi_c B(\mathbf{r}, t))$, where the non-locality is accounted by the term

$$B(\mathbf{r},t) = \left(\frac{1}{N_d}\int e^{-|\mathbf{r}'|/L_c}\, b(|\mathbf{r}+\mathbf{r}'|,\, t)\, d\mathbf{r}'\right), \tag{3.2}$$

where, χ_c is the strength of the competitive interaction, and L_c is the characteristic range of the competition, usually equated to the radius of the rhizosphere. The normalisation constant N_d depends on the dimension d, in two dimensions it corresponds to $N_2 = 2\pi L_c^2$. To simplify the problem, the facilitation is considered to be the local process

$$M_f(\mathbf{r},t) = \exp(\chi_f b), \tag{3.3}$$

where χ_f is the strength of the facilitative interaction. This assumption is justified by the fact that in real ecosystems, the rhizosphere, which accounts for competitive interactions, can extends more than an order of magnitude beyond the facilitative interaction range.

In what follows, the linear stability analysis of Eq. (3.1) is presented, and the existence of localised solutions is discussed.

3.4 Results

3.4.1 Stability analysis of the integro-differential model

In the absence of non-local interactions, i.e. $\chi_c = \chi_f = 0$, which implies $M_c = M_f = 1$ Eq. (3.1) becomes

$$\frac{\partial b}{\partial t}(\mathbf{r}, t) = b(\mathbf{r}, t)\,[\bar{\mu} - b(\mathbf{r}, t)] + D\nabla^2 b(\mathbf{r}, t), \tag{3.4}$$

where $\bar{\mu} = 1 - \mu$. The local Eq. (3.4) correspond to a Fisher-KPP equation [34, 35], which describes a transcritical bifurcation. The homogeneous solution $b = 0$ is stable and $b = \bar{\mu}$ unstable for $\bar{\mu} < 0$, for values of $\bar{\mu}$ above the critical point $\bar{\mu} = 0$ the stability shifts, $b = 0$ becomes unstable and $b = \bar{\mu}$ stable. Due to the physical nature of vegetation, negative values of b, which describe the biomass density are unphysical. This equation describes the front propagation of a stable state over an unstable state, yet does not exhibit stable spatial patterns.

For positive values of χ_c and χ_f, the stationary homogeneous states of Eq. (3.1) correspond to the solutions of

$$0 = b\,[1 - b]\exp(\chi_f b) - \mu b \exp(\chi_c b), \tag{3.5}$$

where the competition term acquires the simple form $M_c = \exp(\chi_c b)$. The solutions of Eq. (3.5) correspond to the bare state $b_0 = b = 0$, and the uniform vegetation covers $b = b_s$, given by the solutions of

$$\mu = (1 - b_s)\exp(\Lambda b_s), \tag{3.6}$$

where the feedback difference $\Lambda = \chi_f - \chi_c$ indicates if the community is *cooperative* ($\Lambda > 1$) or *anti-cooperative* ($\Lambda < 1$). The bare state b_0 represents a territory devoid of vegetation, which is stable for $\mu > 1$, while b_s, solutions of Eq. (3.6), correspond to uniform vegetation covers that can be either monostable for $\Lambda < 1$ or bistable for $\Lambda > 1$ (see Figure 3.3). In the monostable regime ($\Lambda < 1$) the biomass density decreases monotonously with increasing aridity μ and vanishes at $\mu = 1$. In the bistable regime ($\Lambda > 1$) the uniform plant distributions exists up to the tipping point

$$b_l = (\Lambda - 1)/\Lambda \quad \text{and} \quad \mu_l = \exp(\Lambda - 1)/\Lambda, \tag{3.7}$$

obtained by setting $\frac{\mathrm{d}\mu}{db_s}|_{b_l} = 0$ in Eq. (3.6).

The spatial stability of the stationary homogeneous states b_s[1] is studied by introducing a small amplitude

---

[1] Here b_s corresponds to the only solution of Eq. (3.6) in the monostable case ($\Lambda < 1$), and the upper branch for the bistable regime ($\Lambda > 1$).

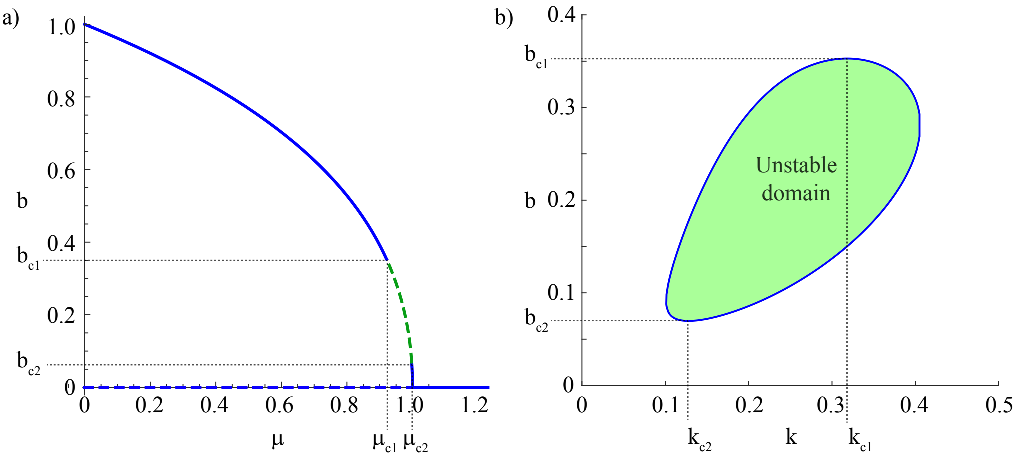


Figure 3.3: **Stability diagrams for Eq. (3.1).** (a) Shows the stability of the homogeneous states b_0 and b_s, as a function of the aridity parameter μ, with respect to homogeneous and inhomogeneous perturbations. Stable states are indicated by solid line, and unstable ones are represented by dotted lines. The Turing-unstable regime is shown in green, where the limits are given by Eq. (3.18). (b) Shows the marginal stability curve of the biomass b with respect to the perturbation wavenumber k. The green shaded area corresponds to the domain of instability delimited by the zeros of Eq. (3.12). Parameter are $\chi_f = 2$, $\chi_c = 1$, $L_c = 4$, and $D = 1$. This figure, for which I hold the copyright, has been adapted from [6].

($\varepsilon \ll 1$) finite-wavelength perturbation away from b_s,

$$b(\mathbf{r},t) = b_s + \underbrace{\varepsilon e^{\sigma(k)t+\hat{\imath}\mathbf{k}\cdot\mathbf{r}} + \text{c.c.}}_{\Delta} \tag{3.8}$$

The deviation from b_s is expressed in terms of normal modes $\exp(\hat{\imath}\mathbf{k}\cdot\mathbf{r})$ as function of the wavevector $\mathbf{k}$. The perturbation is considered along the x-axis, $\mathbf{k}\cdot\mathbf{r} = kx$, without loss of generality. The stability of the perturbation is then obtained by finding the regions where the perturbation decays exponentially in time, i.e $\sigma(k)$ is negative and the state is linearly stable, or the perturbation grows exponentially, where $\sigma(k)$ is positive and the state is linearly unstable.

Introducing Eq. (3.8) in Eq. (3.1) and linearising around the homogeneous stationary state b_s,

$$\begin{aligned} \sigma(k)\Delta &= G(b_s,b_s) + \left.\frac{\partial G}{\partial b}\right|_{b=B=b_s}(b-b_s) + \left.\frac{\partial G}{\partial B}\right|_{b=B=b_s}(B-b_s) - Dk^2\Delta \qquad (3.9)\\ \Longrightarrow\ \sigma(k) &= \left[\chi_f(1-b_s) - 1 - \chi_c(1-b_s)\mathcal{I}_d\right] b_s \mathrm{e}^{\chi_f b_s} - Dk^2 \qquad (3.10)\end{aligned}$$

where $G(b_s,b_s) = 0$ and Eq. (3.6) was used to remove the μ dependence from the second equation. The integral $\mathcal{I}_d$ is dimension-dependent

$$\mathcal{I}_d = \frac{1}{N_d}\int \mathrm{e}^{-|\mathbf{r}'|/L_c + \hat{\imath}\mathbf{k}\cdot(\mathbf{r}'+\mathbf{r})}\mathrm{d}\mathbf{r}' = \frac{\mathrm{e}^{\hat{\imath}\mathbf{k}\cdot\mathbf{x}}}{(1+(kL_c)^2)^{(1+d)/2}} \tag{3.11}$$

and represents the non-local contributions of the perturbation Δ. Finally, the dispersion relation for $\sigma(k)$ for the two-dimensional system reads

$$\sigma(k) = \left[\chi_f(1-b_s) - 1 - \frac{\chi_c(1-b_s)}{(1+(kL_c)^2)^{3/2}}\right] b_s e^{\chi_f b_s} - Dk^2. \tag{3.12}$$

The critical points associated with the Turing instability are obtained from imposing the conditions

$$\left.\frac{\partial \sigma}{\partial k}\right|_{b_s=b_c,\ k=k_c} = 0 \tag{3.13}$$

$$\sigma(k_c) = 0, \tag{3.14}$$

simultaneously. These two conditions provide the thresholds b_c and the most unstable wavenumber k_c associated with the Turing instability. The critical wavenumber at the onset of the instability, obtained from Eq. (3.13), is[2]

$$k_c = \frac{1}{L_c}\left[\left(\frac{3\chi_c b_c(1-b_c)e^{\chi_f b_c} L_c^2}{2D}\right)^{\frac{2}{5}} - 1\right]^{\frac{1}{2}}. \tag{3.15}$$

As a function of the aridity parameters μ, Eq. (3.15) reads

$$k_c = \frac{1}{L_c}\left[\left(\frac{3\mu\chi_c b_c e^{\chi_c b_c} L_c^2}{2D}\right)^{\frac{2}{5}} - 1\right]^{\frac{1}{2}}, \tag{3.16}$$

The characteristic wavelength of the Turing pattern is given by $\lambda = 2\pi/k_c$. This, together with Eq. (3.16), implies that the wavelength of the system and size of the patches scale inversely with the aridity μ. The value of the critical state b_c is obtained from Eq. (3.14)

$$[\chi_f(1-b_s)-1]b_s \mathrm{e}^{\chi_f b_s} - \left(\chi_c(1-b_s)b_s \mathrm{e}^{\chi_f b_s}\right)^{\frac{2}{5}} \left(\frac{D}{L_c^2}\right)^{\frac{3}{5}} \left(\left(\frac{2}{3}\right)^{\frac{3}{5}} + \left(\frac{3}{2}\right)^{\frac{2}{5}}\right) + \frac{D}{L_c^2} = 0. \tag{3.17}$$

This equation can be solved numerically, with the values obtained for the density b_c the values for the critical aridity for which the Turing instability exists can be obtained from Eq. (3.6). In the critical regime ($\Lambda = 1$), with $\chi_f = 2, \chi_c = 1$, $L_c = 4$ and difussivity $D = 1$, Eq. (3.17) has two finite-wavelength solutions

$$(\mu_{c1}, b_{c1}) = (0.921, 0.353) \tag{3.18}$$

$$(\mu_{c2}, b_{c2}) = (0.997, 0.070), \tag{3.19}$$

with critical wavenumbers $k_{c1} = 0.318$ and $k_{c2} = 0.127$, respectively.

The results of the linear stability analysis are summarized in Fig. 3.3. As the aridity level μ increases, the vegetation evolves toward extinction. For $0 < \Lambda \leq 1$, the biomass decreases monotonously. The bare state density b_0 is reached at the switching point $\mu = 1$, where the uniformly vegetated state is destabilized through the Turing instability [71, 101]. In the range $\mu_{c1} < \mu < \mu_{c2}$, the uniform biomass state is linearly unstable and the uniform cover may destabilise into a pattern state. In real ecosystems, extremely arid ecosystems do not support vegetation (desert state), while nutrient-rich environments can support a uniform vegetation cover (forest, jungle or grassland state). The Turing instability occurs when the aridity is decreased a desert state decreases, such that a perturbation,

---

[2] As a function of the dimension d Eq. (3.15) reads $k_c = \frac{1}{L_c}\left[\left(\frac{(d+1)\chi_c b_c(1-b_c)e^{\chi_f b_c} L_c^2}{2D}\right)^{\frac{2}{d+3}} - 1\right]^{\frac{1}{2}}$.

i.e. the deposition of a seed in the terrain, grows. In this way that the desert state is lost. Moreover, the Turing instability occurs in a region where the homogeneous cover is not stable either, as there is not enough nutrients or water to support it. This competition between growth of vegetation and competition for nutrients that leads to the spatial patterning observed in nature.

When increasing the stress or aridity, the wavelength of the corresponding pattern increases (see Eq. (3.16)), simultaneously the morphology of the vegetation pattern changes. In general, as the aridity parameter is increased, the pattern transitions from spots of low biomass density, which form hexagonal patterns, to periodic stripes alternating regions of high and low biomass density, to spots of high biomass density forming hexagonal patterns [102, 103]. This behavior has also been observed in mathematical models that include water transport by underground diffusion and/or above ground run-off [81, 104, 105, 78].

L_c	μ_c	b_c	λ_c
3.5	0.942	0.308	20.0
	0.994	0.106	36.7
4.5	0.907	0.379	20/1
	0.999	0.0505	61.7
5.5	0.888	0.411	21.3
	0.9995	0.0305	88.4
6.5	0.877	0.430	22.7
	0.9998	0.021	117.5

Table 3.1: **Effect of the competition range in the critical values of the parameters**. Each rows show, for the indicated competition range (first column) the critical values of aridity μ_{c1} (top) and μ_{c2} (bottom), biomass density b_{c1} (top) and b_{c2} (bottom) found by solving Eqs. (3.17) and (3.5), and pattern wavelengths $\lambda_{c1} = 2\pi/k_c$ (top) and $\lambda_{c2} = 2\pi/k_c$ (bottom) (Eq. (3.15)). L_c and λ_c are measured in units of distance, b_c in units of density, and μ_c in units of inverse time, accounting for the drain of resources from the system.

Table 3.1 shows how different interaction ranges affect the range of parameters where patterns emerge, delimited by μ_{c1} and μ_{c2}, as shown in Figure 3.3. Given an interaction range, which in real ecosystems depends on the vegetation species and in the properties of the soil. In hard terrain with low infiltration roots tend to be superficial and extend horizontally [88]. Vegetation with longer competitive range have a larger unstable region (see Tab. 3.1, second column). As it is argued later, the relation between the competitive interaction and the pattern wavelength provides means to compare the model with real observations.

3.4.2 Localised vegetation patches

As mentioned earlier, periodic states and localised patches emerge as a self-organised response by the system to changes in the parameters. Aridity is of particular interest as it can be directly related to field measurements and changes in the soil or environmental conditions.An increase in aridity, lowering of hydric resources or other stress factors may cause a contraction of savannas and woodlands in a process called fragmentation, which may lead to desertification. This triggers the formation of vegetation patterns and patches. Theoretically, the necessary and sufficient condition for the formation of localised patches is the coexistence between a homogeneous state and a periodic pattern. This implies the existence of a hysteresis loop, inside of which there is a *pinning* range of the aridity parameter where localised gaps or patches are stable [106, 107, 77]. This comes from the fact that a portion

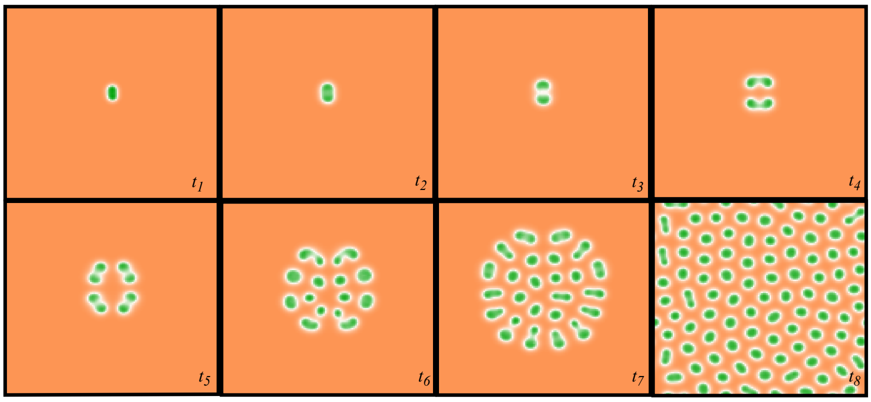


Figure 3.4: **Evolution of a self-replicating vegetation patch**, for a sequence of time points $t_1 < t_2 < ... < t_8$. Obtained through direct simulation of Eq. (3.1), for a system of size 128 × 128-points, and parameters: $\mu = 1.02$, $\chi_f = 2, \chi_c = 1$, $L_c = 4.5$, $D = 1$.This figure, for which I hold the copyright, has been published in [6].

of the pattern state can be connected to the homogeneous state by stationary fronts, getting *pinned*, forming a localised pattern. Similar pinning behavior occurs in many spatially extended systems where a homogeneous steady state coexists with a spatially periodic state [108, 109, 110, 111]. It was first reported by Y. Pomeau [108] for front solutions in hydrodynamics, and has since then also been applied to vegetation patterns [97]. localised structures and localised patterns are a well documented phenomenon, concerning almost all fields of natural science including chemistry, biology, ecology, physics, fluid mechanics, and optics [112, 113, 70].

The stability analysis of Eq. (3.1) shows that for some range of the aridity parameter, there is coexistence between homogeneous and pattern states. This implies the existence of stable localised vegetation patches in that parameter range. These stable circular localised structures are supported by one of the homogeneous steady state, $b_0 = 0$ (bare state) or b_s (homogeneously vegetated). In the context of vegetation dynamics, localised structures that emerge from the unpopulated state $b_0 = 0$, correspond to circular patches of vegetation that are surrounded by bare terrain.

3.4.3 Self-replication: from a localised vegetation patch to an extended pattern

When the aridity is tuned beyond the regime of stability of the localised vegetation patches they may be affected by a modulational instability, where the patch elongates, increases in size to finally split through the decay of the central "bridge", into two new, identical patches. This dynamic process, which resembles cellular mitosis, occurs repeatedly to each of the new patches allowing a single initial patch to end up covering the whole system, as can be observed in numerical simulations (Fig. 3.4). In the long time limit, the system relaxes to a periodic hexagonal pattern of vegetation patches. This is the same phenomenon recently described in a related model for vegetation dynamics (Bordeu et al. [16]). However, in that case the authors considered a simplified model with only local interactions.

Here it is hypothesised that if present in real vegetation, self-replication may allow the total biomass of the ecosystem to increase even when the aridity is too high for the existence of a stable homogeneous cover of vegetation. Fig. 3.5 shows the increase of the total biomass on the transition from a single patch to four patches

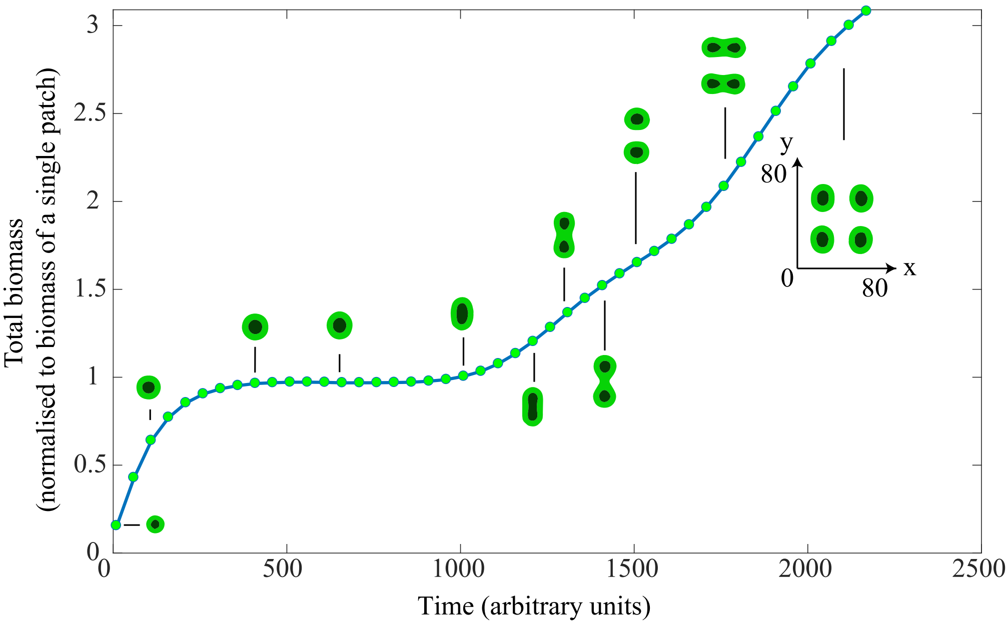


Figure 3.5: **Evolution of the total biomass** during the transition from a single vegetation patch to four patches through self replication. Points obtained through direct simulation of Eq. (3.1), for a system of size 128 × 128-points, and parameters: $\mu = 1.02$, $L_c = 4.50$, $\chi_f = 2$, $\chi_c = 1$, $D = 1$.Representative images of the localised structures are shown at the different stages. The total biomass is normalized with respect to the biomass of a mature patch. Thus, during maturation a single spot reaches a normalized biomass of 1, to replicate into two patches to reach a total of almost 2. After a second replication (four patches) the system reaches a normalized biomass of 3. This figure, for which I hold the copyright, has been adapted from [6].

through self replication, measured in units of a single mature localised patch. It should also be pointed out that in this model, Eq. (3.1), self-replication occurs only in regions with available terrain. This means patches on the periphery of the patch cluster are much more likely to self-replicate than patches lying in the bulk (see Fig. 3.4). For this reason, the growth rate of the total number of patches is proportional to the perimeter of the patch cluster, thus grow linearly with time in the limit of large cluster size[3].

As mentioned earlier, self-replication occurs through the destabilisation of a localised structure, where there is an exponential growth of a modulational perturbation of the localised structure. The growth rate σ' of the perturbation gives information on the time scales of self-replication and can be extracted from direct numerical simulations of Eq. (3.1). For a fixed competition range L_c in Eq. (3.1), there is a range of aridity μ for which an initial localised structure destabilizes and undergoes self-replication (see Fig. 3.6). As expected, the growth rate σ' is faster for lower levels of aridity μ. For a fixed value of the aridity parameter μ, a longer competition range L_c leads to a faster destabilisation of the localised patch, as can be seen in Fig. 3.6. This shows that the non-local competition between vegetation not only has the intuitive effect of decreasing the growth of neighbouring vegetation, yet also contributes to a faster destabilisation of localised patches, accelerating the self-replication process.

Moreover, for values of the aridity parameter μ below the self-replicating region (above the top dashed line in

---

[3] This is related to growth of the perimeter of the area explored by a branching process, as described in Chapter 2, in the super-critical regime [22].

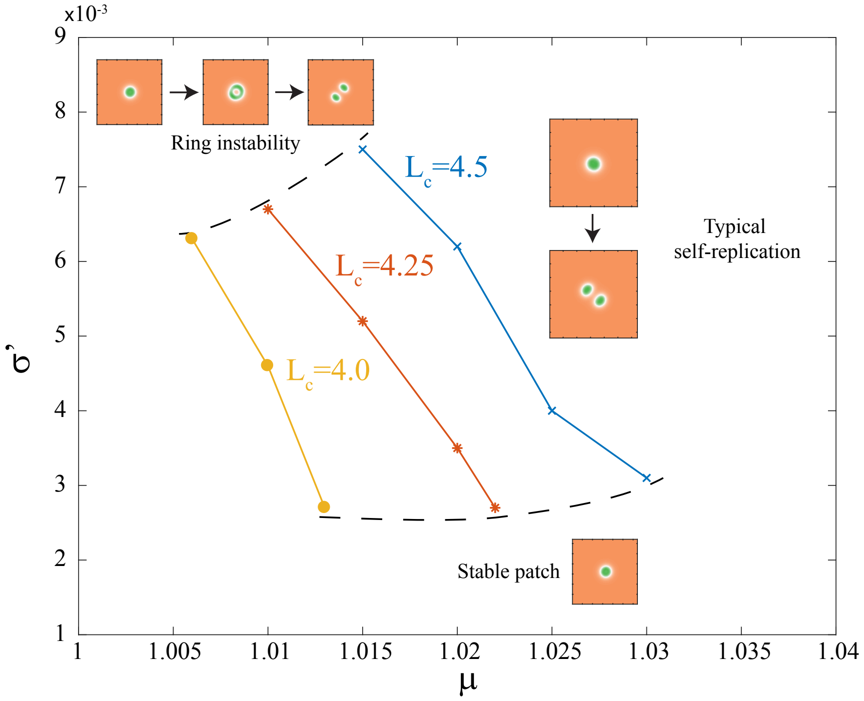


Figure 3.6: **Growth rate σ' of the unstable mode corresponding to an elongation of the vegetation patch, leading to self-replication**, for competition range ($\times$) $L_c = 4.50$, ($*$) $L_c = 4.25$, and ($\bullet$) $L_c = 4.00$, as a function of the aridity parameter μ. Outside the limits indicated by black dashed lines, localised patches may remain stable (bottom), or destabilize through a ring instability (top). Points were obtained through direct simulation of Eq. (3.1), for a system of size 128×128-points, and parameters: $\chi_f = 2, \chi_c = 1, D = 1$. Lines correspond to linear interpolation between the numerical data points for visualisation purposes only. This figure, for which I hold the copyright, has been adapted from [6].

Fig. 3.6) there is a different route to self-replication, where the patch suffers from a ring instability. The dynamics of the ring instability is as follows: the unstable localised patch growing radially, after reaching a critical size the central portion of the structure decays, forming a doughnut-like shape. This is also unstable, and is followed by the decay of two opposite sides of the doughnut, the structure ends up dividing into two new localised patches. However, this instability is observed solely on isolated patches. Once the single initial patch has divided once through the ring instability further self-replicating events take place through the usual mitotic-like divisions.

In real ecosystems vegetation patches are rarely isolated. Interactions between patches may play an important role in shaping the large scale properties of their spatial distribution. To assess the effect of such interactions, large scale simulations containing hundreds of uniformly distributed patches were performed. The main objective of this was to assess the role of the self-replicating process in the patterning of the system, where a characteristic wavelength emerges in the system. For this, a system of 512×512-points was considered. The initial condition was constructed through a spatial Poisson point process with rate 0.001, this generates on the order of 600 points randomly distributed in the two-dimensional plane considered. Each of these points was then considered as the center of a localised patch of 3-points radius, this can cause the merging of two patches, the resulting initial state can be seen in Fig. 3.7a. The random field constructed was then used as the initial condition for simulating (3.1), for a parameter region inside the self-replication windows (see Fig. 3.6, for $L_c = 4.50$). The transient state in Fig. 3.7c

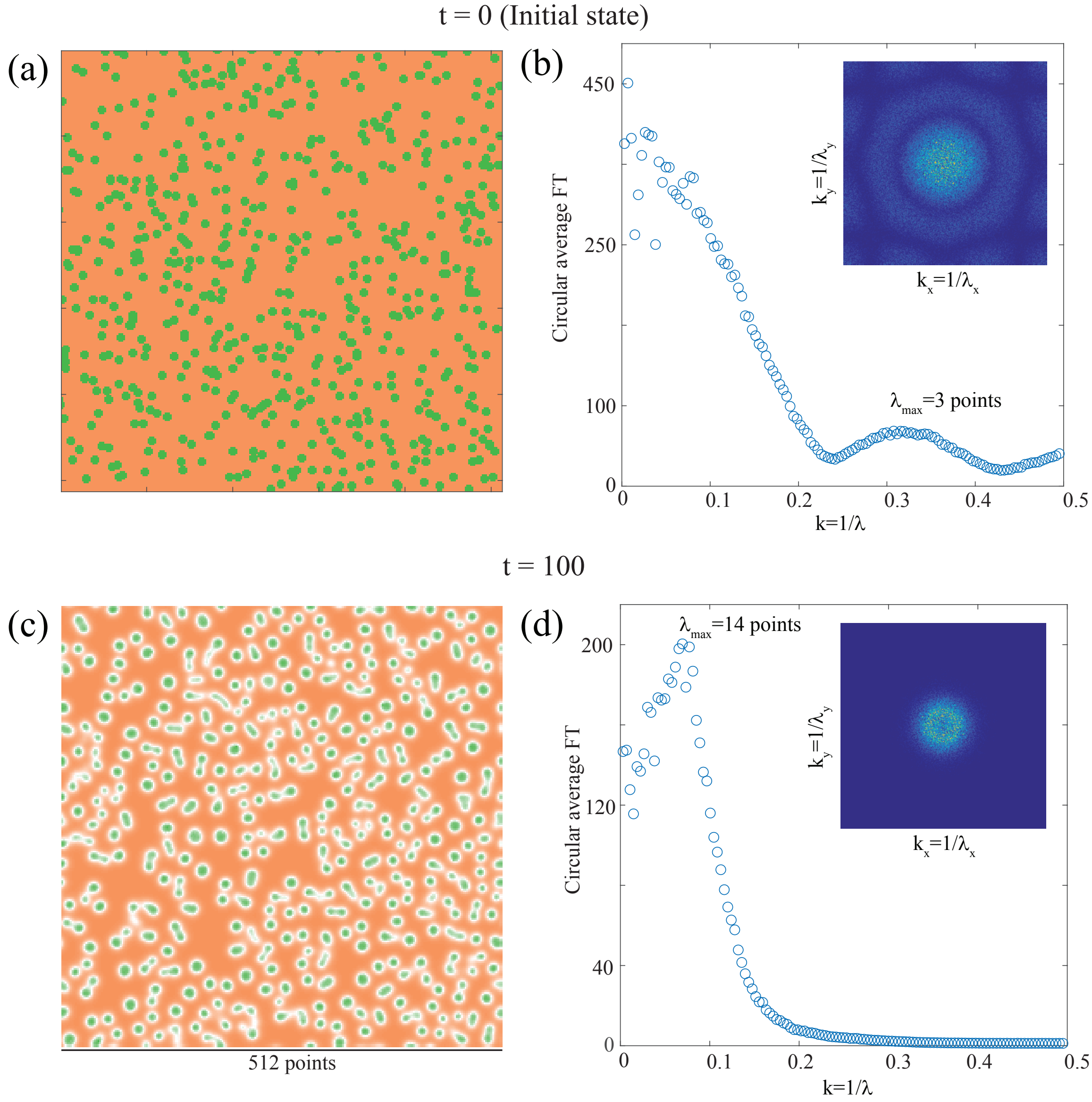


Figure 3.7: **Large scale simulations**, (a,b)-shows initial initial condition generated from a Poisson point process with structures of 3-point radius, and its corresponding Fourier transform, respectively. (c,d)-Shows the evolved state after 1000 iterations ($dt = 0.1$, $dx = 2.0$), and the corresponding Fourier transform, respectively. Parameter: $\mu = 1.02$, $L_c = 4.5$, $\chi_f = 2$, $\chi_c = 1$, and $D = 1$. This figure, for which I hold the copyright, has been adapted from [6].

shows that the clumps quickly break up into distinct patches, while smaller patches grow in size, commencing the self-replication process.

To assess the level of patterning of the initial and transient systems, the spatial Fourier transforms and corresponding power spectrum was computed. For the initial state (Fig. 3.7b), the power spectrum (apart from the central peak) exhibits a characteristic wavelength of the order of the size of the localised structures ($\lambda = 3$ A.U), due to the approximately mono-dispersed sizes of the computer generated structures.

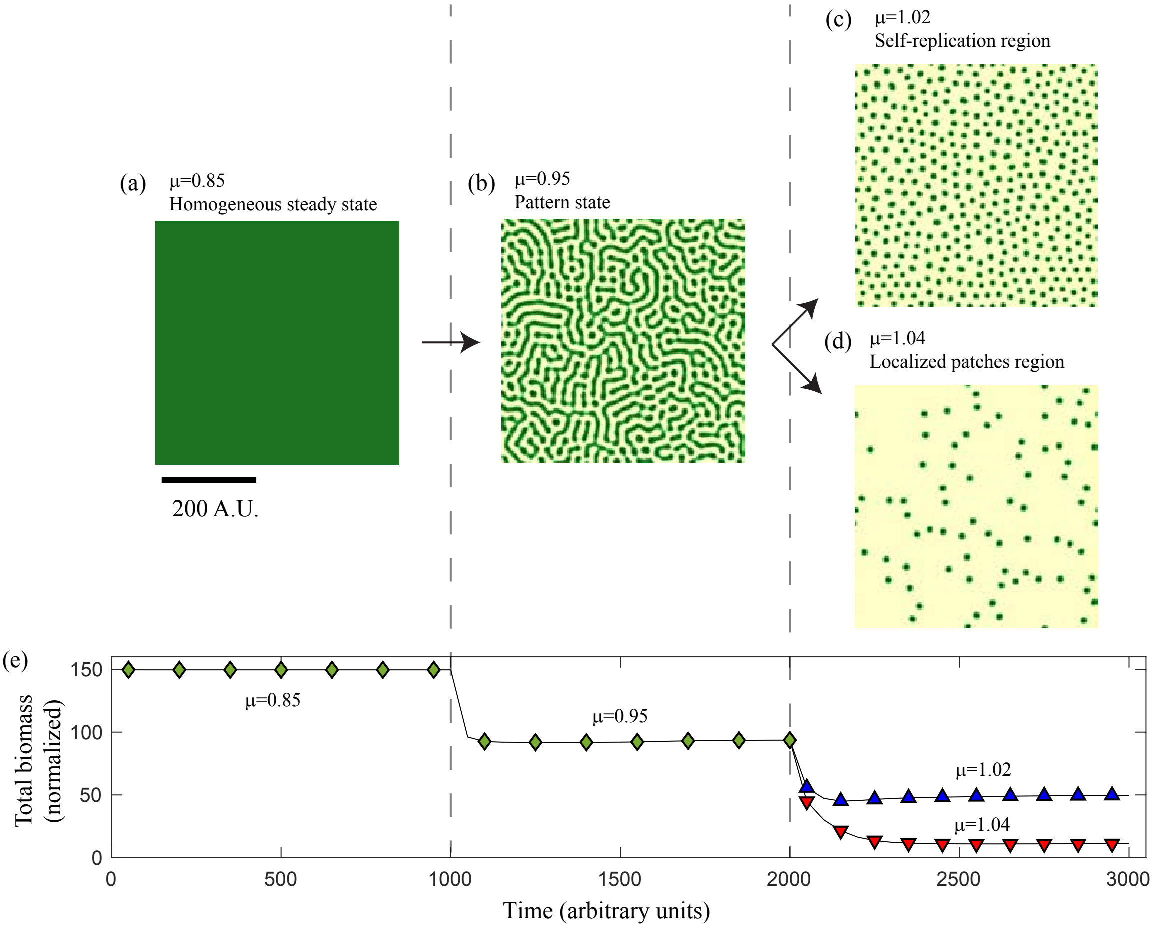


Figure 3.8: **Fragmentation of the homogeneous cover**. (a) Shows the initial condition, where the system is in a completely vegetated state for $\mu = 0.85$. (b) The homogeneous state decays to a labyrinthine pattern by an increase in aridity to $mu = 0.95$. (c) Shows the evolution of the system in the self replicating region ($mu = 1.02$). (d) Shows the evolution of the system in the stable localised patches region ($mu = 1.04$). (e) Shows the corresponding temporal evolution of the total biomass (normalized by the biomass of a the single patch), diamonds show the transition from $\mu = 0.85$ to $\mu = 0.95$, triangles show the transition to $\mu = 1.04$, while down-triangles show the transition to $\mu = 1.02$. Simulation parameters: $L_c = 4.5, \chi_f = 2, \chi_c = 1$, and $D = 1$. This figure, for which I hold the copyright, has been adapted from [6].

On the other hand, when analyzing the transient state ($t = 100$, see Fig. 3.7d), a characteristic wavelength much larger than the size of the structures ($\lambda = 14$ points) is observed. This wavelength emerges as a result of the self-organizing nature of the system (as described by Eq. (3.15)), where clusters split and the localised patched interact to rapidly arrive to a characteristic spatial distribution. There are regions in the parameter space where clumps can elongate without splitting and form labyrinthine patterns, which correspond to finger-like structures that coexist with circular patches in a spatially disordered pattern, this was as also observed by Bordeu [20] on a related model.

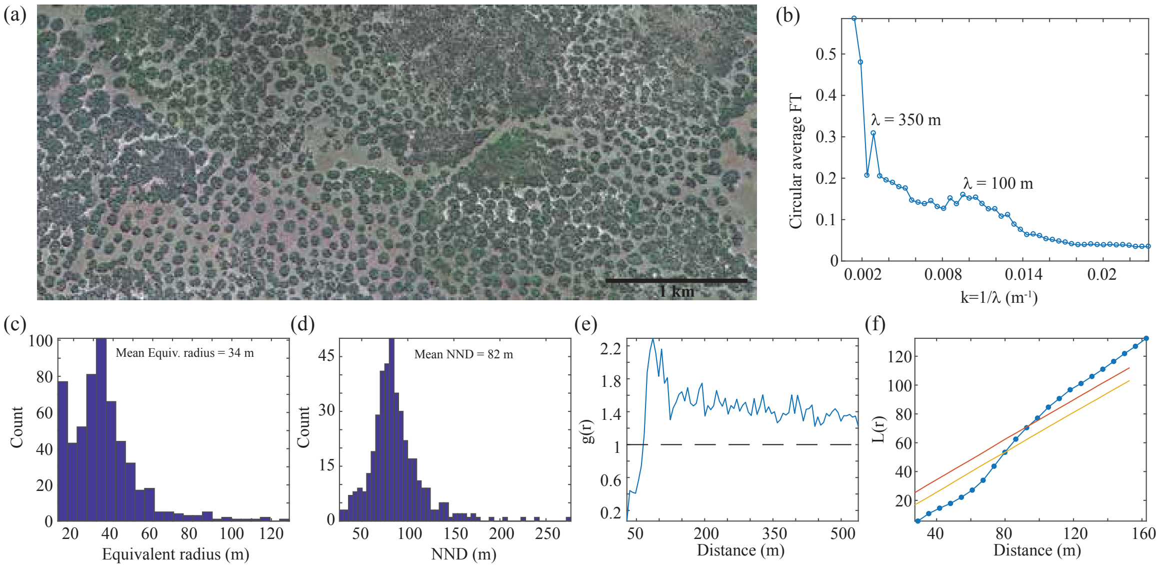


Figure 3.9: **Remote measurements of vegetation patches in Zambia**. (a) Shows the satellite image obtained from Google Earth Pro. (b) Shows the circular average of the 2d Fourier transform of the detected patches. (c) & (d) Show the histograms for the equivalent radii and the nearest-neighbor distances between patches, respectively. (e) & (f) Show the radial distribution function $g(r)$ and Ripley's L-function $L(r)$ (lines show the 95% confidence interval), respectively. This figure, for which I hold the copyright, has been adapted from [6].

3.4.4 Fragmentation: From homogeneous vegetation to patchy landscapes

In the previous section, it was shown that through self-replication a single or multiple localised patches may increase the total biomass of a system, leading ultimately to an extended hexagonal pattern state. However, this is not the only route by which a system can reach such patchy state. A second route is shown in Fig. 3.8, whereby an homogeneously vegetated state (Fig. 3.8a, $\mu = 0.85$) is destabilized by an increase in aridity towards a labyrinthine pattern (Fig. 3.8b, $\mu = 0.95$). Further increasing aridity drives the system towards two qualitatively distinct states, one is reached when the aridity μ is increased to the region where localised patches are unstable and suffer from self-replication (Fig. 3.8c, $\mu = 1.02$), there the labyrinthine pattern breaks down into multiple localised patches that cover the system, generating a patterned state. The second, qualitatively different, state is reached when the aridity is increased to the parameter region where localised patches are stable (from $\mu = 0.95$ to $\mu = 1.04$, Fig. 3.8d), there the system reaches state consisting of a low number of isolated patches distributed in a disordered manner.

It is important to note that a direct transition from the stable homogeneous vegetation region ($\mu = 0.85$) directly into any of the patchy regions ($\mu = 1.02$ or 1.04) generates an overall decay of the homogeneous state to the bare state. Thus, an intermediate level of aridity is necessary for allowing the destabilization of the homogeneous cover into a non-periodic patch pattern.

3.4.5 Field observations through remote-sensing

The self-organisation hypothesis for vegetation dynamics and ecosystem patterning has solid theoretical grounds [76, 104, 70]. However, processes such as ecosystem fragmentation or patch self-replication may, in principle, take decades, this limits the data availability for the comparison between theoretical and numerical observations.

Here, satellite images are used in order to close the gap between the analytic and numeric results presented and the dynamics of real ecosystems. Geographic locations were selected for which vegetation patches were clearly visible, so that their spatial distribution could be studied. Eight distinct regions where selected, four located in the Fombeni region, Mozambique, and other four in the Mufumbwe District, North-Western Province of Zambia (see Fig. C.1), where the vegetation is dominated by medium size trees *Brachystegia spiciformis*. The observed patches appear to be composed of groups of trees forming compact clusters (see Fig. 3.9a). Each patch can cover an area of thousands of square-meters, with an effective radius of tens of meters, as can be observed in Figs. 3.2 and 3.9a.

The typical radius of the patches varies from region to region, ranging from 7 to 36 meters. Most of the regions show a marked characteristic size of the vegetation patches, and nearest-neighbor distance between them, as can be seen in Figs. 3.9c and 3.9d, respectively. This suggests the existence of an underlying mechanism for the spatial organisation of the patches.

To further study the properties of the spatial organization of the patches, both the Ripley's L-function and radial distribution function where measured. The radial distribution function $g(r) = N(r)/N_r(r)$ measures the ratio between the number of patches $N(r)$ that lie within rings of radius r and width dr centred on a patch given, and the expected number $N_r(r)$ from a random distribution of patches. Values of $g(r) < 1$ ($g(r) > 1$) indicate clustering (dispersion) of patches at a given scale. The L-function corresponds to the cumulative $g(r)$, thus measures the level of clustering or dispersion up to a given scale. In all regions analysed, both $g(r)$ and $L(r)$ (Figs. 3.9e and 3.9f) show strong clustering at short scales, of the order of two to three patches, and dispersion for longer scales. The different scales observed is further supported by the spatial Fourier transform of the patches detected (see Fig. 3.9b), which shows two apparent characteristic scales, one of the order of 100 m and a larger one on the order of 350 m.

To study if there was a relationship between the size of the patches and the distance between them, as predicted by Eq. (3.1), the mean equivalent radius was plotted against the mean nearest-neighbour-distance (NND) between structures for each of the regions considered (see Fig. 3.10). A strong correlation was observed between the two quantities, indicating that there might be an underlying mechanism controlling both the structures size and the distance between them, with Zambia exhibiting the largest and more disperse structures. These observations are consistent with the self-organisation hypothesis, where, as described theoretically by Eq. (3.16), there is an inverse relation between the characteristic wavelength (and size of patches) and the level of aridity. Measurements of the rainfall and temperature records of both regions (see Tab. C.1) showed that Mozambique is rainier, however Zambia 3°C colder on average. The way in which these two environmental measures, together with properties of the soil, can be related to the theoretical *aridity* μ remains unclear.

Field measurements of vegetation similar to the one observed in the regions of Zambia and Mozambique studied here [114] have found that the typical competition distances (here L_c) are in the order of 9 meters, and are an order of magnitude larger than above ground canopy. However the space between vegetated pattern can extend

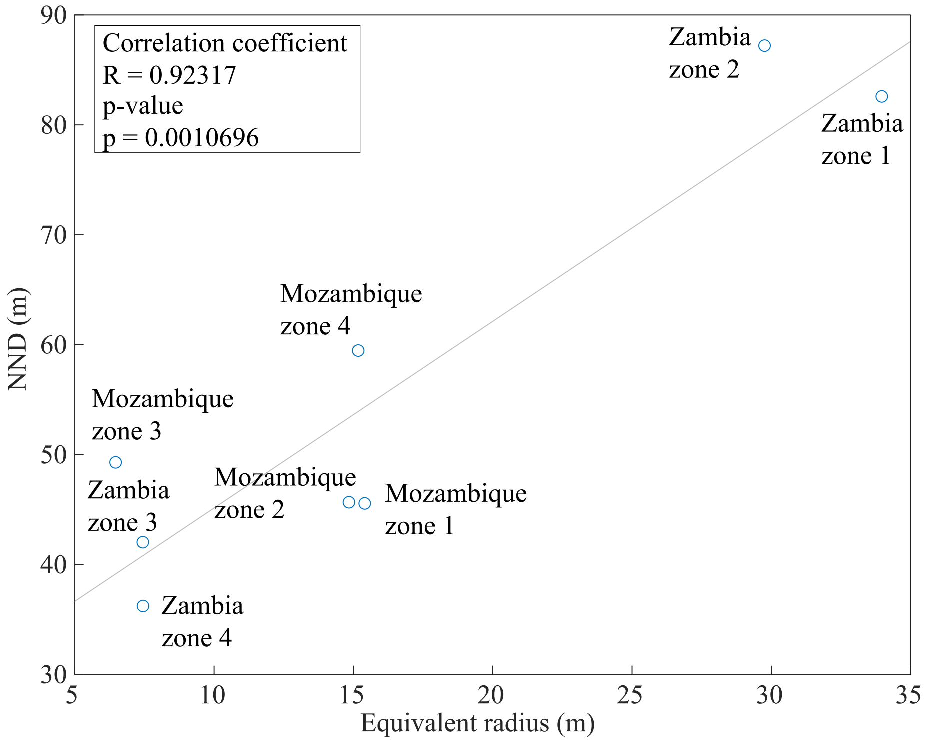


Figure 3.10: **Linear correlation** between the mean equivalent radius of the structures and the nearest-neighbor distance for each of the eight regions analyzed. This figure, for which I hold the copyright, has been adapted from [6].

for up to 50 meters. These values are consistent with the theoretical measurements of the model used here (see Tab. 3.1) where it was observed that the characteristic wavelength λ_c of the patterns ranges from five up to more than ten times the characteristic distance of the competitive interactions L_c.

3.5 Discussion and outlook

In this chapter, it has been shown that self-replication, a phenomenon previously observed in local model of vegetation dynamics (Bordeu et al. [16]), is also present when non-local competitive interactions are considered. Self-replication occurs when a localised patch destabilised and split into two new patches. An unstable patch may undergo multiple self-replication events, to finally cover the whole space available with a regular hexagonal pattern, causing an increase in the total biomass in the system, over time. The use of the non-local model, as compared to the local non-linear model used in earlier work [16], allowed the analysis of the effect of the non-local interaction range in the destabilisation of localised patches. It was shown that the rate of growth of the instability leading to self-replication depends not only on the level of aridity considered, but also on the range of the non-local interactions. A longer characteristic range for the competitive interactions lead to a faster destabilisation of the vegetation patches. Furthermore, it was found that self-replication may occur trough a ring instability, where the patch grows radially followed by the decay of the central region. Ring structures are also observed in real landscapes (see Fig. 3.11) in transition regions from patchy to gaped terrains.

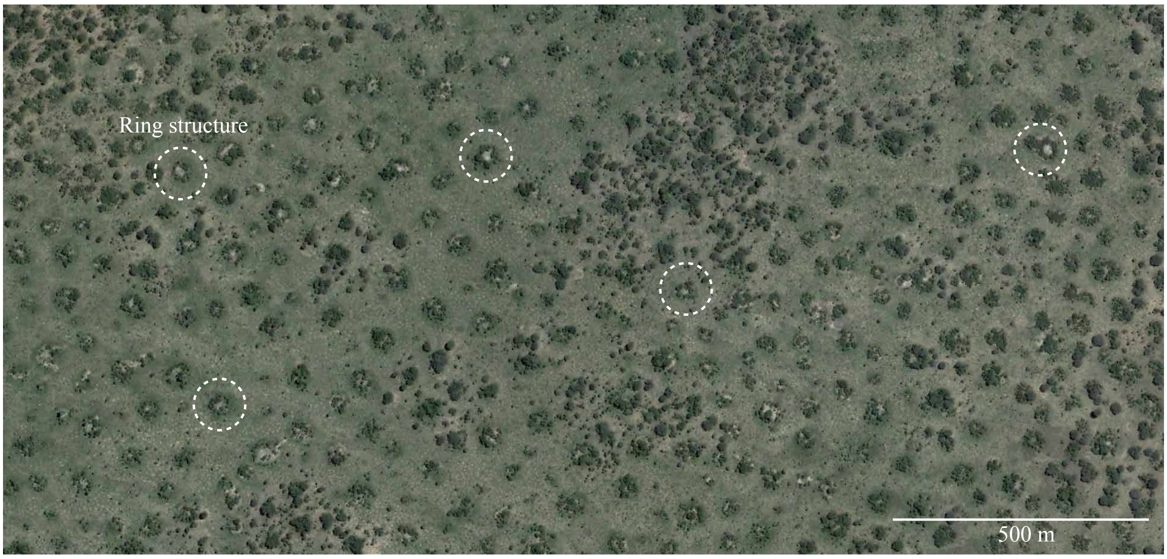


Figure 3.11: **Remote observation of vegetation patches and rings in Zambia.** Some representative rings are circled.

The study of the reverse process, by which a homogeneous cover decays into a patchy state revealed that the self-replication regime may be an important intermediate stage in the decay of homogeneous covers of vegetation. This provides enough time for the system to destabilise into a dense patchy state instead of decaying to the barren state altogether. Alternatively, a decrease of the aridity beyond the self-replication regime leads to a more abrupt decay in biomass, resulting in a low number of sparsely distributed patches. It is hypothesized that these processes are occurring in the observed landscapes of Mozambique and Zambia.

Through the analysis of satellite images, evidence was found to support the self-organization hypothesis. This includes (i) the existence of well defined circular vegetation patched, with a characteristic size and NND, (ii) the relation between the equivalent radius and NND, which is compatible with the theoretical predictions, where a single wavelength controls both the size of the structures and the pattern wavelength and has an inverse relation with aridity, (iii) the existence of ring structures, (iv) large scale observations of the transition from homogeneous vegetation to patchy landscaped gaped vegetation. Despite the limited amount of data, which prevents an evaluation of the underlying processes involved in the organisation of the observed landscapes at the plant level, the results presented here strongly suggest a self-organisation origin for the observed landscapes. More data, especially field measurements, long time-lapse imaging and controlled experiments are necessary to confirm any claim on the nature of the observed patterns. Nevertheless, the measurements presented in this chapter are a first step to understand the distribution of the observed patterns. The state of some of the observed patches suggest that self-replication might be an ongoing process, for example in the "dividing" patches showed in Fig. 3.2.

Similar models to the one presented in this chapter, Eq. (3.1), has been previously confronted quantitatively to experimental measurement for two type of plants (Festuka orthophylla [90] and Combretum micranthum [87, 88, 89]. However, the knowledge of the below-ground structures of the type of vegetation of Zambia and Mozambique is rather limited. From the current literature only rooting depth data and a quantitative index of the vertical distribution of roots are available, which are irrelevant for the models considered here. Moreover, the full lateral roots extension has not yet been measured. For this reasons parameter extraction and quantitative comparisons between theoretical results and field observations are unfeasible with the currently available data.

In the model, the parameters, such as aridity and interaction ranges are homogeneous in space and time. In real ecosystems, variations in the terrain, due to the presence or absence of rivers, valleys, mountains and animals, render the effective aridity, space dependant. Time-dependent cycles, such ad day and night, seasonality, and extreme weather events might also influence the patterning observed. Moreover, the life cycle of vegetation, which are not considered in the model, imply an increase in the competitive range over time, as the rhizosphere matures and enlarges. Future work will focus on addressing these space and time dependent features to study how they affects the stability and pattern selection in the system. Special emphasis will be on obtaining real data to compare the theoretical results with real observation. An additional line of study, will focus on the interaction between multiple species. In real ecosystems different species can compete for nutrient and space, however they might as well have symbiotic relations whereby two species benefit each other, increasing their survability. It is expected, from a neutral dynamics perspective, that a species with an advantage will quickly dominate over others. However, it is yet unclear what the effect of different interaction ranges and different sensitivities to aridity will be on the stationary state of the system.

Chapter 4

Effective potentials in cell-signal interactions

The results presented in this chapter are under review for publication in *Physical Review X* as part of a manuscript entitled *Effective potentials in cell-protein signal interactions*, co-authored by Clare Garcin, Shukry J Habib and Gunnar Pruessner. Large parts of this chapter are quoted verbatim from a manuscript submitted to Phys. Rev. X [7].

I performed and optimized experiments together with Clare Garcin. I developed the methods for data analysis and theoretical model together with Gunnar Pruessner. I carried out analytic calculations and numerical simulations. Shukry Habib designed the experiments in discussion with CG, IB and GP. I wrote the manuscript. All authors discussed and interpreted the data, and commented on the manuscript.

4.1 Summary

Cell signaling is essential for cell fate determination and tissue patterning. Cells can dynamically explore their environment and probe for instructive signals. As signalling molecules are presented to the receiving cell, they are recruited and recognised by its membrane as to elicit a response to organise multicellular tissues. Cell-signal interactions depend on multiple factors, such as the biochemical properties of the signals and the cell dynamics. The accumulation and transport of signalling proteins by cells can generate a redistribution of the signals in the environment, resulting in an emergent organisation of the signals' spatial distribution. The detailed study of these processes at the single cell level lies at the foundation of understanding multicellular organisation. In this chapter, a simplified experimental setup is used, where single mouse embryonic stem cells (mESCs) can interact with immobilized protein signals, in vitro. This system was developed to mimic in vivo situations where the protein signal is presented asymmetrically to the stem cell.

A two species age-dependent correlation function is introduced to describe the spatio-temporal dynamics of cell-signal interactions, which take place as active cells explore an inanimate environment that holds protein-coated micrometer-sized beads. Through the analysis of mESC data and numerical simulations it is shown that cells act as effective force-field generators, perturbing and organising their environment. This organisation, captured in the form of an ageing effective potential, is an emergent property of the population of single cells interacting with randomly distributed localized signals. Although effective potentials are extensively used in the modeling of active particles, little is know of their biological applications. In this chapter, experimental evidence is given of ageing effective potentials in a single cell system, their origin is determined, and the implications for the organisation of cellular systems is discussed. In this context, ageing is a consequence of the active nature of cells, which accumulate and transport signals, thereby re-arranging their environment and influencing the observed effective interactions. By constructing the phase-diagram for the effective potential a previously unobserved transition from a zone of net attraction to a zone of net repulsion is found. This transition is related to the effective cell-protein signal interaction, which can be dominated by a capturing, disregarding or release mechanism, generating different cell-signal interaction profiles. Furthermore, it is shown how these dynamic interactions relate to the organisation of the protein-coated beads by cells. The ageing interactions give rise to temporal changes in the configurational entropy of the system. This production of entropy, a fingerprint of out-of-equilibrium systems, is then cast in terms of the effective potential, to show that self-renewal signals contribute to a rapid organisation of the environment by embryonic stem cells.

4.2 Introduction

Pluripotent stem cells are the building blocks of multicellular organisms. Pluripotency refers to the ability to specialise or differentiate into any of the three germ layers, endoderm, ectoderm and mesoderm [115]. The fate of a pluripotent stem cell is determined by multiple factors. In particular, external factors, such as chemical and mechanical [116] signals can regulate the precise outcome of the differentiation process. Concentration gradients of morphogens are thought to be one of the mechanisms responsible for controlling generic developmental processes and tissue patterning [117, 118, 119]. The producing cells can secrete signaling ligands into the extracellular matrix, which may diffuse to receiving cells or be actively recruited by them via membrane extensions projecting into the extracellular matrix. The producing cells can secrete signaling ligands into the extracellular matrix. These may either diffuse to receiving cells or be actively recruited by them via thin membrane extensions projecting into the extracellular matrix. In some cases, signalling molecules are tethered to the cell membrane, preventing them from diffusing [120, 121, 122]. Biochemical signals are then transmitted by direct cell to cell contact, which might provide a faster and more controllable mechanism for signal propagation in tissues [117, 120, 122]. Cells recruit, accumulate and transport these protein-signals by thin cell projections and movement, thereby organising their environment. In the remainder of this chapter the term "signal" is used synonymously to "protein-signal".

The ways in which the mesoscopic organisation of cells and thus signals emerges from the microscopic cell-signal interactions is still unclear. Understanding the spatio-temporal properties and scales involved in cell-signal interactions at the single cell level are the basis to understand the organisation of cells at the tissue level.

Studying specific stem cell-signal interactions in vivo is challenging due to the inherent complexity of multicellular tissues, where a myriad of signals are being generated and received by cells simultaneously. There,

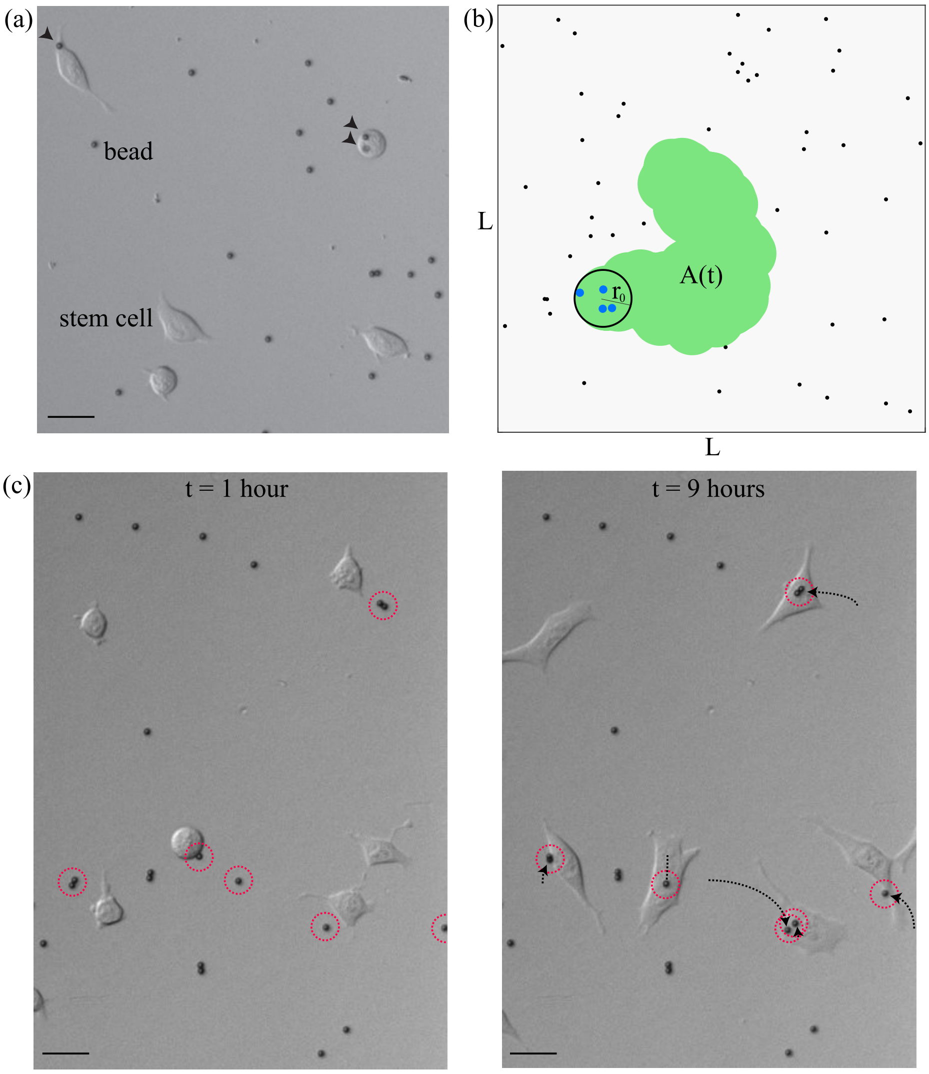


Figure 4.1: **Representative images of cell-bead interactions.** (a) Snapshot of the experimental observation of single mESCs and beads. Arrowheads indicate beads in contact with cells, scale 20 μm. (b) Simulation of a single cell of radius $r_0 = 1/15$ (encircled black) performing a random walk with step length $\ell = r_0/5$, on a periodic system of linear size $L = 1$. The area covered by the cell up to time $t = 500$, $A(t)$, is shown green. Dots correspond to beads randomly distributed in the environment (black) and beads captured by a cell (blue). Total bead density $\rho_b = 50$. (c) Shows two representative images, at times $t = 1$ hour (left) and $t = 9$ hours (right), of the interaction of mESCs and Wnt-beads. Beads that come in contact with cells are encircled, while their net translation between the two frames is depicted by dashed arrows (right). scale bar in (a) and (c) is 20 μm. This figure has been submitted to PRX as part of the manuscript [7]

pinpointing the exact role of each signal and analysing specific cell-signal interaction is impractical.

In recent years, there has been a wide interest in applying techniques from equilibrium and non-equilibrium statistical mechanics to study cell organisation and tissue growth at multicellular level. Mean-field approaches from fluids [123, 124], glasses [125], and nematic liquid crystals [126] have proven useful for modelling tissue growth and ageing. Specifically, correlation lengths [127] and stationary cell-cell effective potentials of interaction [128] have been quantified. However, the dynamic interactions between cells and biochemical signals at single cell level are only partially understood.

In this chapter a previously reported experimental method is used by which Wnt3a protein is covalently immobilised to the surface of micrometer-sized beads [129, 130]. These are inert, non-toxic beads, they have an iron oxide nucleus encased in a thin polymer layer. This layer is coupled to Carboxylic Acid to which the ligand of interest is bound. When mouse embryonic stem cells (mESCs) come into contact with these biochemically active, **signalling** *Wnt-beads*, the beads provide a localised source of a self-renewal signal that can promote asymmetric cell division, leading to the differentiation of only one of the two daughter cells [131, 129]. A more detailed description of this *canonical* Wnt pathway is given in Sec. D.4. The Wnt protein coating can be chemically modified by treatment with dithiothreitol (DTT) [130] to generate biochemically inactive, **non-signalling** *iWnt-beads*. It is likely that both Wnt and iWnt have the capability of binding to a similar range of co-receptors on the cell membrane [129, 132] (this hypothesis has not yet been tested experimentally). However, the inactivated iWnt protein has an altered conformation preventing it from activating the pathways associated with asymmetric cell division and differentiation. For this reason, differences can be expected in the interaction of cells with Wnt-beads with that of cells with iWnt-beads. Uncoated beads, *Unc-beads*, are used as control. These have no protein coating and would otherwise stimulate the interaction between the bead and receptors on the cell membrane. The beads considered are not engulfed into the cytoplasm and remain on the membrane surface for the duration of the experiments.

To quantify the interactions between single cells and localised signals (beads) a suitable correlation function is introduced. This allows to show that cell-signal interactions give rise to, and can be quantified by, ageing effective potentials. These potentials capture cell-bead interactions from the level of the microscopic positioning of beads up to the mesoscopic organisation of the environment. They characterise the interaction strength, range, and reveal how cells position signals on the cell membrane, while also shedding light on the consequences of transport of signals by cell movement. This method reveals differences in the interaction of cells with different types of bead coating (signalling, non-signalling, and uncoated), which affect both the spatial and temporal scales of the interactions. Numerical simulations of simplified models for the cell-bead system reveal different mechanisms in the cell-signal interaction. Even more, these simulations allow the exploration of regimes where the transport of beads by cells controls a smooth transition of the effective potential from being effectively attractive, promoting the accumulation of beads, to effectively repulsive profile, promoting a depletion of beads around the cells. This behaviour turn out to be key in the recapitulation of experimental observations.

The ageing properties of the effective cell signal interactions relate to the continuous and dynamic organisation of the environment by cells (see Fig. 4.1c). This organisation is here quantified based on the relative position of beads with respect to cells, by measuring the excess configurational entropy of the system. This quantifies the driving of the system away from the uniform distribution of beads. In the absence of cells, beads remain static on the bottom of the well, thus thermal contributions to the entropy can be neglected. Any spatio-temporal changes in the configurational entropy of the system are then a result of the active driving of cells. This positive entropy production is a hallmark of out-of-equilibrium systems [133]. In such systems, detailed balance does not hold, implying a preferred direction in phase space for the system to evolve in.

Entropy production measures a ratio between the probabilities of observing *forward* transitions say, from state

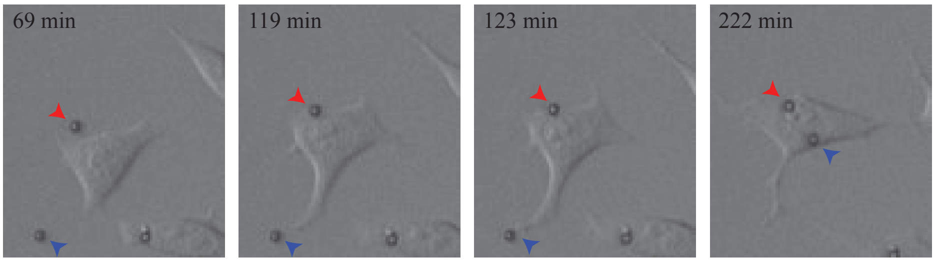


Figure 4.2: **Cell contacting beads.** A mESC recruits two Wnt-beads (time $t = 69, 119, 123, 222$ min, runs from left to right). These images are the raw input data as obtained from the microscope and to be fed to the tracking algorithm. At $t = 69$ min a bead lies by the cell (red arrowhead), yet it only comes in contact with the cell at $t = 119$ (second frame). At $t = 119$ min the cell extends, coming in contact with a second bead at $t = 123$ min (third frame). This cannot be detected reliably by the algorithm until much later, for instance, when the bead is on the cell body, $t = 222$ min (last frame). The same difficulties arise when attempting to measure the potential release of beads.

A to state B, and *backwards* probabilities of observing the transitions from say, state B to state A [134]. Experimental measurements of entropy production have been obtained in carefully prepared systems, where states are well defined, and transition rates or time-dependent distribution functions can be extracted accurately. These include colloidal particle experiments [135], fluidised media [136], ideal gasses [137], and molecular motors [138]. However, little is known about entropy production of living matter at the cellular and tissue level and its relation to the structural order of living systems. There, measuring the observable necessary for computing the entropy production from experimental data is not trivial, as they are not finite states systems, thus the transition rates are not well defined. Formal approaches for computing the entropy production of field theories for active matter have been formulated [139], however their possible application to real data is not clear.

In this chapter it is shown how, by casting the entropy production of the cell-beads system solely in terms of the ageing effective potential, the temporal evolution of the configurational entropy of the bead spatial distribution can be measure both experimentally and numerically. This entropy production rate is characterised as a function of time, and as a function the properties of the signals, to show that relevant biochemical signals contribute to a faster and more efficient organisation of the environment. This might have relevant implications in the formation of specialised micro-environments in vivo, such as the stem-cell niche [121].

The remainder of this chapter is organised as follows: in Section 4.3 the experimental set-up used for the cell-signal interaction analysis is described. In Section 4.4 the age- and time-dependent effective potential is introduced, and in Section 4.5 the experimental measurements are presented. Numerical characterization of the effective potential, and its interpretation in the light of the experimental observations are shown in Secs. 4.6 and 4.7. In Section 4.8 the results on the entropy production based on experimental and numerical data are shown. The discussion and outlook are left for Section 4.9.

4.3 Experimental setup

To measure experimentally the spatio-temporal interactions between single cells and beads live imaging of mouse embryonic stem cells (W4 cells [140]) was performed in the presence of Wnt-, iWnt- or Unc-beads (Dynabeads M-270 Carboxylic acid) at one-minute time resolution for periods of twelve hours. About 2500 mESCs were seeded together with ca. 20000 (0.3 μg) of either Wnt-, iWnt or Unc-beads on each of two wells of a black, glass-bottomed, 96 well plate (Greiner), in serum and LIF containing media. On the bottom of the 136 mm^2 well the bead number density reaches ca. 150 beads/mm^2. Also, the cell number density, which reaches ca. 20 cells/mm^2 is such that cells are observed mostly in a single cell state for the duration of the experiment.

The time-lapse imaging at 1 minute resolution for 12 hours was performed using a Zeiss inverted Axio Imager epifluorescence microscope, equipped with a CoolSNAP HQ2 CCD camera, at 10× amplification. Cells were allowed to settle at 37°C, 5% CO$_2$ for 30 minutes prior to imaging. The area covered by a single position of the microscope was 874.99×653.73 μm^2 containing ca. 11 cells and ca. 84 beads (see Fig. 4.1). About 15 positions were selected per well, and each experiment was repeated at least three times (biological repeats), for a total of more than 1000 cells and more than 10000 beads for each condition. The experiments avoid imagining techniques that might alter the behaviour of cells, such as additional genetic manipulation, molecular markers or phototoxicity. The time-lapse images were analysed with a custom-built multi-particle tracking algorithm (based on MATLAB R2018b), which extracts the trajectories of both cells and beads as function of time (see Sec. D.1.3 for details). Because of the absence of membrane markers, features of a linear extent of less than 0.6 μm are not reliably detected by the algorithm. It is therefore impossible to reliably and reproducibly determine when a bead (such as the highlighted in Fig. 4.2) acquires of loses contact with the cell. However, the algorithm locates the cell body and its centroid reliably, as required in the measurement of spatio-temporal correlations.

Preliminary observations indicate that in the absence of cells, beads distribute, up to effects by excluded volume, uniformly at the bottom of the well. However, in the presence of cells, the spatial distribution of beads is strongly affected by cell-bead interactions, which appear to change in time, and depend on the bead coating. After settling at the bottom of the well, cells move and interact with the surrounding beads. When contacting a bead, the bead might get *captured* by binding to the cells' membrane. Beads might also be *released*, unbinding from the cell membrane and remaining at the bottom of the well. In the next section, a method is described for quantifying cell-signal (bead) interaction and the emergence of spatial order in the bead distribution by measuring the spatio-temporal correlations of the positions of beads due to the presence of cells.

4.4 Ageing effective potential

The spatio-temporal structure of condensed matter has been extensively studied through the measurement of the dynamic or van Hove correlation function [141]. Its Fourier transform in space and time, namely the dynamic structure factor, can be measured directly in X-ray experiments [141]. For mesoscopic objects, such as cells and beads, where positions of every particle can be precisely determined, the dynamic correlation functions can be calculated in real space and time. In the cell-bead system, cell movement and interaction with their surroundings

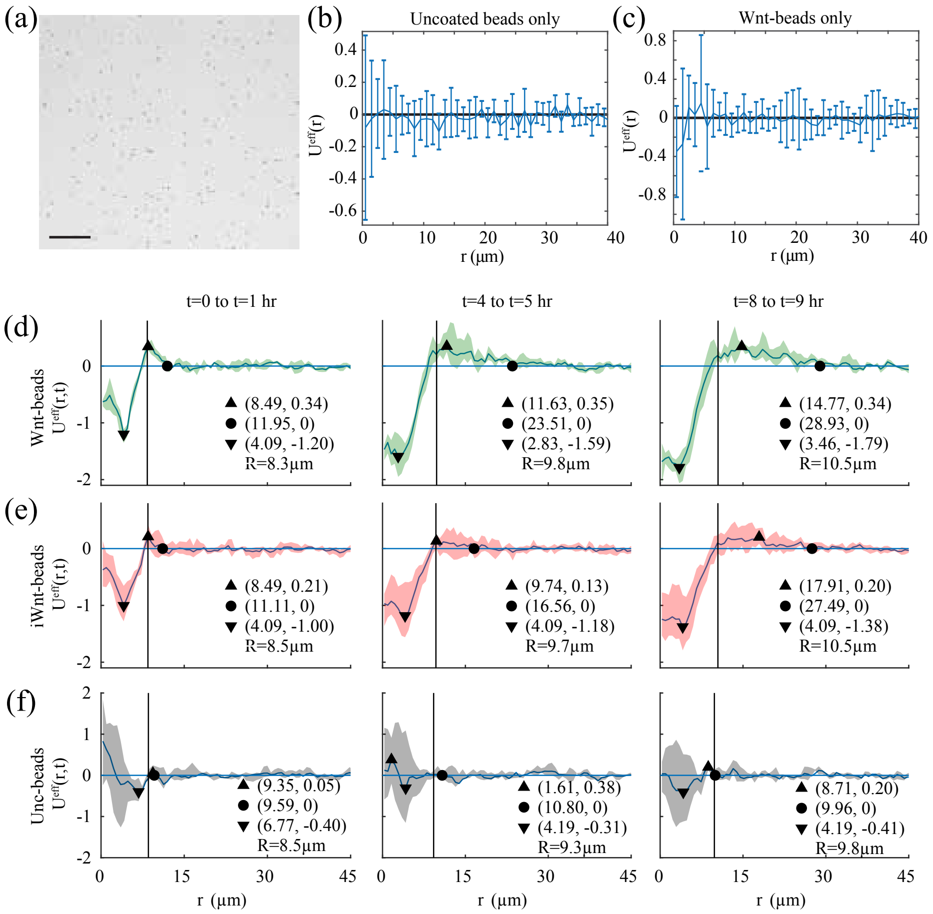


Figure 4.3: **Instantaneous effective potentials.** (a) Snapshot of the experimental distribution of beads, and (b) the instantaneous effective potential ($\tau = 0$) for Unc-beads, and (c) Wnt-beads, in the absence of cells. (d) Effective potential for Wnt-beads, (e) iWnt-beads, and (f) Unc-beads in the presence of mESCs, taken over three different time windows (time increases from left to right, as indicated). The solid curve represents the mean across the repeats, while the shaded area represents the range of the data. The reference line $U^{\text{eff}}(r,t) = 0$ shows the effective potential of a uniform bead distribution. Indicated by (▲), (▼), and (●) are the points of maximum depletion, accumulation, and the range, respectively (see main text). The vertical reference line indicates the mean cell radius R, at each time point. This figure has been submitted to PRX as part of the manuscript [7]

causes temporal changes in the correlations, known as *ageing* [142]. It occurs because the system is initialised in a configuration that is not necessarily its steady state. Moreover, the non-equilibrium nature of living cells might prevent the system from having an steady state configuration in the first place, thus exhibiting permanent dynamics.

The age-dependent properties of cell-bead interaction can be quantified by the correlations in the positions of beads in relation to the positions of cells. To formalise this idea, a modified version of the van Hove correlation function is introduced, namely a *two-species age-dependent correlation function* (ACF). Given a spatial arrangement of N cells and M beads, at positions R_i and $\mathbf{r}_j$, respectively, the ACF, $\hat{G}(\mathbf{r}, t, \tau)$ is defined as function of the position, **r**, age or time, t, and delay, τ, as

$$\hat{G}(\mathbf{r}, t, \tau) = \frac{1}{N} \left\langle \sum_{i=1}^{N} \sum_{j=1}^{M} \delta[\mathbf{r} - \mathbf{r}_j(t+\tau) + R_i(t)] \right\rangle, \tag{4.1}$$

where $\delta(\cdot)$ corresponds to the Dirac-delta function. This correlation function measures the density of beads at position **r** at time $t + \tau$, given that there was a cell at the origin at time t. The angular brackets $\langle \cdot \rangle$ indicate the ensemble average. The expression in Eq. (4.1) corresponds to a two-species version of the van hove correlations function, where self-interactions are negligible, and there is an implicit dependence on time t, which is referred to as *ageing* (see Sec. D.1.2 for details).

In the case of non-ageing systems, i.e. without dependence on absolute time t, Eq. (4.1) depends only on the time-delay τ, so that time t can be set to zero. Hence, the ACF has a fixed value for a given delay τ, independent of the moment t at which the measurement has been taken. As described above, and as will be demonstrated in the next section, this is not the case for cell-bead interactions, which displays a temporal evolution due to active cellular dynamics.

From the ACF, Eq. (4.1), the ageing pair-correlation function $\hat{g}$ can be obtained from the relation

$$\hat{G}(\mathbf{r}, t, \tau) = \rho_b \hat{g}(\mathbf{r}, t, \tau), \tag{4.2}$$

where ρ_b is the density of beads. For isotropic systems, such as the cell-bead system, the pair correlation function depends only on the distance $r = |\mathbf{r}|$. This is usually referred to as the *radial distribution function* (RDF) and is the central observable in the theory of simple liquids [141]. The effective potential of interactions between the two species (cells and beads) is then defined here as

$$U^{\mathrm{eff}}(r, t, \tau) = -\ln[\hat{g}(r, t, \tau)]. \tag{4.3}$$

For systems in thermodynamic equilibrium, the effective potential is stationary in time, whereas in the setup considered in this chapter the active nature of cells renders the potential age-dependent. An equilibrium liquid equipped with the effective potential determined by Eq. (4.3) will reproduce the RDF used in its definition, even when the RDF is obtained in non-equilibrium. Effective potentials have been used before in the context of cellular sheets [128], as they offer some intuition about the evolution of the system. However, similar to statistical potentials [143, 144] the effective potential contains no more information than the RDF it has been derived from, and the link to quantities like the entropy (Sec. 4.8) is convenient, yet equally well expressed in terms of the RDF.

The interpretation of the effective potential (4.3) in the cell-bead system, corresponds to that of other areas in physics. In the present case, the effective potential quantifies attraction (accumulation) and repulsion (depletion) of beads by cells as a function of position r. When the effective potential $U^{\mathrm{eff}}(r, t, \tau)$ is equal to zero for all values of r, the system is (effectively) non-interacting, i.e. beads are uniformly distributed irrespective of the locations of cells, and there are no effective forces that disturb their spatial distribution. Negative values of the potential, $U^{\mathrm{eff}}(r, t, \tau) < 0$, imply an accumulation of beads, caused by an effective attraction of beads towards the given point of interest. In contrast, positive values of the effective potential, $U^{\text{eff}}(r, t, \tau) > 0$, are a signature of bead depletion, which can be caused by cells transporting beads away from that particular region in space.

When extracting the radial distribution function and effective potential from the experimental data, the observables are averaged over the entire population of single cells in the experiment (ensemble average), which is repeated for every experiment (details in Sec. D.1.4 and Fig. D.2).

4.5 Experimental observations

In the absence of cells, beads distribute uniformly at the bottom of the well (Fig. 4.3a), which can be quantified by measuring the stationary effective potential, Eq. (4.3), of beads around randomly selected points in the plane. The measurements were performed 30 min after seeding the beads, allowing them to reach a stationary position at the bottom of the well. This implies that the effective potential is independent of both age and time-delay, i.e. $U^{\text{eff}}(r, t, \tau) = U^{\text{eff}}(r)$. Thus, for this particular setup $U^{\text{eff}}(r)$ was computed at a single time point, and it is observed that the effective potential shows no significant deviations from the null-line, $U^{\text{eff}}(r) = 0$ for all r, as expected for a uniform, independently random distribution of beads (Fig. 4.3b). In this setting, bead-bead interactions occur only through volume exclusion, which can be ignored in the low bead density conditions considered in these experiments. As expected, the protein coated beads by themselves also show a uniform distribution and thus a vanishing effective potential (Fig. 4.3c).

In general, the emergence of a non-zero effective potential is a consequence of the interaction between particles. In the cell-bead system, interactions arise due to the presence of active mESCs, which act on the inanimate micron-sized beads. Beads move only by the effect of external forces, such as direct contact with cells, or by thermal fluctuations when they are in suspension in the liquid medium. Regardless of the type of bead coating considered (uncoated, Wnt or iWnt) cells are always able to interact with beads. However, the properties of such interactions are greatly affected by the coating, which is evident even to the naked eye. Both the number of cells in contact with beads and the duration of the interaction was visibly higher in the case of coated beads compared to uncoated-beads, throughout the 12 hours of imaging. As a first approach to quantify these observations, and to characterize the strength and properties of cell-bead interactions the *instantaneous* effective potential, $U^{\text{eff}}(r, t, \tau = 0)$ was measured. It is considered that the effective interactions are significant if the measured effective potential deviates from the null, $U^{\text{eff}} = 0$, by an amount greater than the range of the experimental measurements (shaded region in Fig. 4.5).

When $\tau = 0$, $r = 0$ corresponds to the center of the cell, thus this observable gives information on the preferred location of beads around the cell center. Measurements of the instantaneous effective potentials in the case of protein-coated beads show important differences in the interaction between mESCs and Wnt-beads or iWnt-beads, as seen in Figs. 4.3d and 4.3e, respectively. Although the qualitative features of the effective potentials are similar between both conditions, their quantitative features differ. The interaction range, that is, the radial distance from the cell that is affected by depletion or accumulation ($U^{\text{eff}} \neq 0$) quantifies the ability of cells to recruit beads from the environment (indicated by (●) in Fig. 4.3), while the position of the minimum (maximum) of the effective potential indicates the most (least) probable position r of a bead that interacts with a cell, indicated by ▼ (▲).

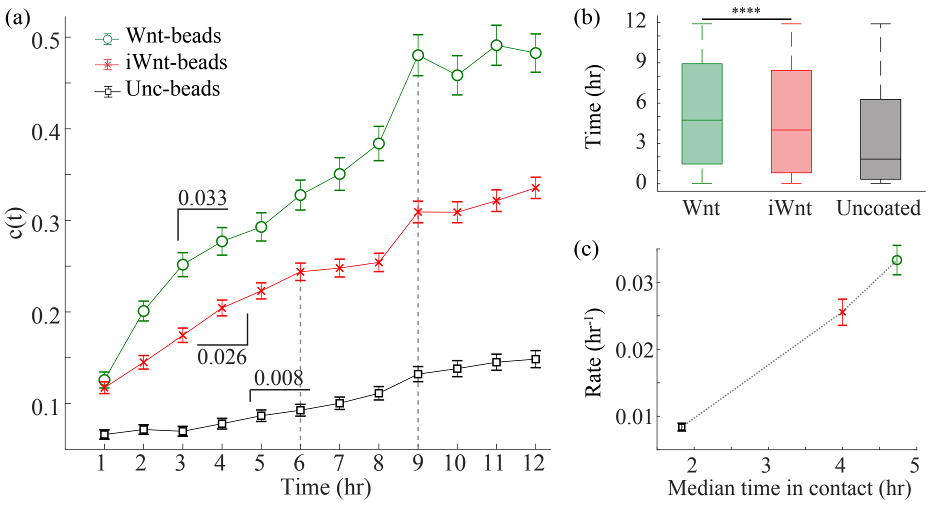


Figure 4.4: **Statistical properties of the interactions.** (a) Coordination number $c(r)$, Eq. (4.4), as a function of time for the three bead types. Dashed lines indicate the time at which the $c(t)$ start plateauing for Wnt- and iWnt-beads. (b) boxplot of the effective time beads stay in contact with cells, and (c) linear relationship between the accumulation rates in (a) and the median contact time in (b). In all panels green, red, and black corresponds to Wnt-, iWnt-, and Unc-beads, respectively. Error bars in (a) and (c) show the 95% confidence intervals. The statistical difference between the data distributions in (b) were calculated via a Kolmogorov–Smirnov test, where **** indicates a p-value<0.0001. This figure has been submitted to PRX as part of the manuscript [7]

The wider interaction range and deeper effective potential in Fig. 4.3d, compared to that in Fig. 4.3e, indicates a stronger interaction of cells with Wnt-beads than with iWnt-beads. When in contact with Wnt-beads, mESCs appear to keep the beads closer to the center of the cell compared to cells in contact with iWnt-beads as seen when comparing the location of the potential minima (▼) in Figs. 4.3d and 4.3e. These observations suggest that when in contact with a signal, cells might control the signal's spatial location depending on the biochemical properties.

When cells are exposed to uncoated beads the interaction properties are drastically different compared to coated (Wnt- or iWnt-) beads. Uncoated beads have no protein ligand bound to their surface, thus provide no biochemical cue to the cells, acting merely as a physical entity in the environment. In this scenario the instantaneous effective potential remains close to zero for the duration of analysis (Fig. 4.3f and Fig. D.5). Interactions with uncoated beads are weak throughout the duration of the experiment as observed by the weak effective repulsion from the cell centre ($r = 0$) at the initial stages and a shallow attraction at later time-points. The wide dispersion in the data suggests that the interaction is quite volatile, possibly due to the lack of a proper biochemical coating preventing effective cell-beads bonds. This is in agreement with previous observations [145], which indicate that the presence of surface ligands favours the creation of durable chemical bonds between cells and external elements, such as artificial substrates, bioengineered matrices, and in the present case, micron-size beads.

In analogy to the theory of simple liquids, a time-dependent coordination number $c(t)$ is measured to quantify the ability of cells to capture and retain beads, [141],

$$c(t) = 2\pi\rho_b \int_0^R \mathrm{d}r \; \exp(-U^{\mathrm{eff}}(r,t,0))r, \tag{4.4}$$

where $\exp(-U^{\text{eff}}(r,t,0)) = \hat{g}(r,t,0)$, and R is the average radius of the cell The coordination number Eq. (4.4) is an approximated measure of the average number of beads around a cell at a given time t (Fig. 4.4a). The slope of the coordination number, indicated in Fig. 4.4a, defines an accumulation rate, which determines the overall efficiency with which cells contact and acquire beads once they start probing the environment. A high rate indicates a deepening in the effective potential, and an increase in bead accumulation, and the formation of new cell-bead pairs.

The accumulation rates (Fig. 4.4a) show an efficient recruitment of signalling Wnt-beads by mESCs, increasing the strength of the effective potential at a faster rate compared to iWnt-beads. Moreover, the coordination number for iWnt-beads reaches a plateau around the 7 hour mark, while Wnt-beads recruitment continues for up to 9 hours. For uncoated bead the coordination number grows at a slow rate for the 12 hour period (Fig. 4.4a). These observations suggest that cells are able to retain Wnt-beads for longer periods of time compared to iWnt- and unc-beads, systematically increasing the overall number of cell-bead bonds in the system. This is confirmed by measuring the time beads stay in contact with cells (Figs. 4.4b and 4.4c), where significant differences between the three types of beads are observed. Importantly, the retention time is an effective quantity obtained from the difference between the first and last time point on which beads where considered *in contact* with a cell by the tracking algorithm (for discussion see Sec. 4.3 and Fig. 4.2). The accumulation rates and retention times are linearly correlated (Fig. 4.4c), as argued above, longer retention times imply more stable cell-bead bonds, in turn increasing the overall accumulation of beads by cells. Here, durable and efficient cell-bead contacts appears to rely strongly on the bead coating but also on the biological relevance of the coating, with cells retaining Wnt- over iWnt-bead for significantly longer periods. The observed dependence of the interaction strength on the bead coating is in line with related studies, where surface modification of beads cause differential uptake of beads by cells [146, 147].

In summary, mESCs actively interact with their surroundings, contacting, attracting and repelling localised signals. These interactions can be captured in the form of an effective potential, a statistical observable that gives information on the population-level organisation of the system. The instantaneous effective potential $U^{\text{eff}}(r,t,\tau = 0)$ reveals the range, strength and efficiency of interactions, yet it does not capture the effects of cell motility, which is a key feature of the cell dynamics. Beads are recruited and transported by direct contact with the cell membrane, which in turn, occurs due to cell movement and changes in cell shape. In order to observe the effect of cell dynamics on the effective potential, positive time-delay, $\tau > 0$, is considered in the following section.

4.5.1 Time-delayed measurements

The effect of a time delay ($\tau > 0$) on the measurement of the effective potential is a reduction of the strength of the attraction, as seen in Fig. 4.5. This flattening of the effective potential is caused by cell motility. For a given time-delay τ, the distribution of beads at time $t + \tau$ is measured with respect to the positions of cells at time t. Any displacement of the cell between t and $t + \tau$ affects the reference point from which the measurement of the radial distribution function Eq. (4.2) and the effective potential Eq. (4.3) are taken. For large values of τ, the effective potential can become flat for all values of r, resembling an effectively non-interacting system (Fig. 4.5).

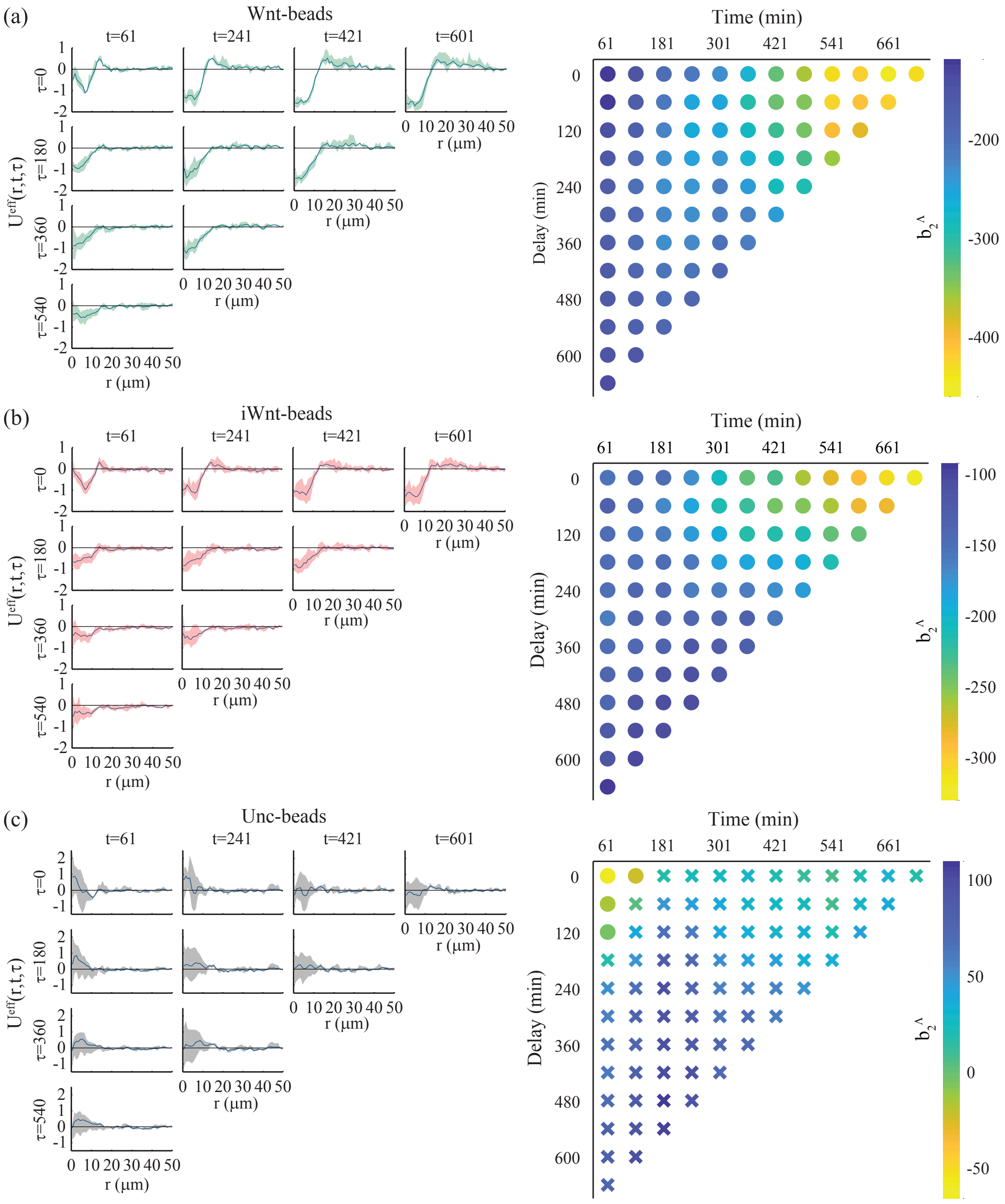


Figure 4.5: **Age- and delay-dependent effective potentials.** Experimental effective potential as function of time t and delay τ (left), and (t, τ)-phase-diagram (right), for (a) Wnt-beads, (b) iWnt-beads, and (c) unc-beads, in the presence of cells. Markers indicate attractive (circles) and repulsive (crosses) regions in the phase diagrams, as measured by b_2^Λ (Eq. (4.5), colorbar) for $\Lambda = 2R$, where R is the cell radius. This figure has been submitted to PRX as part of the manuscript [7]

To characterise this and other changes of the effective potential as a function of both age t and time-delay τ, a phase-diagram in the (t, τ)-plane was constructed (see Fig. 4.5). Each point in the phase diagram if obtained by integrating the Mayer f-function $f(r, t, \tau) = \mathrm{e}^{-U^{\mathrm{eff}}(r,t,\tau)} - 1$ [148, 141]

$$b_2^{\Lambda}(t, \tau) = -\pi \int_0^{\Lambda} \mathrm{d}r\, f(r, t, \tau). \tag{4.5}$$

In equilbium systems and in the limit $\Lambda \to \infty$, Eq. (4.5) corresponds to the second virial coefficient, which accounts for the first order corrections to the pressure in low density systems. In the cell-bead case, the interest is on the behaviour of the effective potential around the cell, thus a finite upper cutoff $\Lambda = 2R$ is considered, where R is the cell radius. This particular value is chosen in order to capture the relevant features of the effective potential at short length scales. Negative values of $b_2^{\Lambda}(t, \tau)$ indicate a deep, attractive effective potential, while positive values indicate effective repulsion around the cell centre [141].

As expected, at fixed time t, the overall effect of increasing the delay τ is to reduce the effect of attractive interactions (see Figs. 4.5a and Fig. 4.5b, right column), as the effective potential approaches the null line. On the other hand, for a fixed delay τ, the overall effect of ageing is to increase the strength of the effective potential. However, for Unc-beads (Fig. 4.5c) the phase diagram appears to transition from attractive at short time and delay, to purely repulsive at longer time and delay. Here the phase diagram appears inverted with respect to that of coated beads, suggesting that a different interaction mechanism might be taking place compared to Wnt- and iWnt-beads. It was already pointed out that interactions are reduced in the case of Unc-beads (see Fig. 4.3), however this is not enough to explain the inversion observed in the phase diagram, where the effective potential appears to become stronger, more repulsive with increasing time t or delay τ.

In the following section, a simple theoretical model for the cell-signal interactions is introduced to obtain a comprehensive understanding of the interplay between the cell dynamics, age, time-delay and the effective potential. Together with numerical simulations, this model allows the identification of the causes of the ageing of the effective potential. Moreover, it explains the differences between the phase-diagrams, as seen in Fig. 4.5.

4.6 Theoretical and numerical analysis

The model of cell-bead interactions considers two species of particles in a continuous, finite, two-dimensional space. Firstly, immobile point-particles, corresponding to beads, were distributed randomly on a unit square with periodic boundary conditions (black particles in Fig. 4.1b), so that their number density was $\rho_b = 1000$. Secondly, a mobile, single cell was placed at the origin of the unit square. This cell corresponds to an active mESC, modeled as a diffusive disk of radius r_0 (black circle in Fig. 4.1b). At each time-step the cell moves a small fixed distance $\delta < r_0$ in a random direction on the torus, which amounts to a diffusion constant $D = \delta^2/4$ (see Sec. D.1.1 for details). A bead is *captured* by a cell if it lies within the cell's disk after a time step. Captured beads (blue particles in Fig. 4.1b) are distributed inside the corresponding cell according to a normal distribution with standard deviation $\sigma = r_0/2$, centred on the cell's centre. The choice of distribution does not affect the quantitative results presented below, it does however modify the precise shape of the effective potential. Initially, release of beads captured by cells is not allowed, such that bead capturing is an irreversible process, and is the only explicit interaction between beads and cells.

In the model described above, the master equation for the probability $P(n,t)$ that a cell has captured n beads up to time t, may be written as

$$\frac{dP(n,t)}{dt} = \rho_b \frac{dA}{dt}[P(n-1,t) - P(n,t)], \tag{4.6}$$

where ρ_b is the number density of beads, and $A = A(D, r_0, t)$ corresponds to the area covered by the diffusive disc, representing the cell, of radius r_0 and diffusion constant D up to time t, sometimes referred to as the area of the *Wiener sausage* [149, 150]. To first order the area explored by the cell grows like

$$A(D, r_0, t) \approx \frac{4\pi Dt}{\log\left(\beta \frac{Dt}{r_0^2}\right)} \tag{4.7}$$

in the long time limit, where $t \gg r_0^2/D$. Here, $\beta = 4\exp(-2C)$, and $C \approx 0.577$ is the Euler-Mascheroni constant [149]. As the upper-critical dimension of a random walk is two, it may be assumed that ρ_b is constant in space and time [151]. Under this assumption, Eq. (4.6) can be solved explicitly to find

$$P(n,t) = \frac{[\rho_b A(D, r_0, t)]^n}{n!} e^{-\rho_b A(D, r_0, t)}. \tag{4.8}$$

The expected number $n(t)$ of beads captured by a cell as a function of time is then proportional to the area explored by the cell up to time t, $n(t) = \rho_b A$ (green area Fig. 4.1b). The excess number of beads within the cell radius $n_{\mathrm{ex}}(t) = n(t) - \pi r_0^2 \rho_b$ determines the level of accumulation of beads, i.e. the depth of the attractive range of the instantaneous effective potential ($U^{\mathrm{eff}}(r, t, \tau = 0) < 0$). At the same time, the accumulation of beads by the cell is proportional to the depletion of beads in the region surrounding it (see Fig. 4.6a, first row). Thus, the excess number of beads $n_{\mathrm{ex}}(t)$ can be regarded as a proxy for strength of the attraction and repulsion parts of the effective potential. The irreversible nature of the capturing process implies the $n_{\mathrm{ex}}(t)$ is a monotonically increasing function, thus the strength of the effective potential is bound to increase with time.

The effective radius of the area explored by a diffusive cell, $r_{\mathrm{e}}(t) = \sqrt{A(D, r_0, t)/\pi}$, (indicated by $*$ in Figure 4.6a) corresponds approximately to the interaction range of the effective potential, as it indicates the extent to which the environment is affected by the action of a cell. The combined evolution of $n_{\mathrm{ex}}(t)$ and $r_{\mathrm{e}}(t)$ explains the experimental observation for mESCs with Wnt- and iWnt-beads, where the effective potentials get stronger and longer-ranged with time, as consequence of capturing and transport of beads by cells.

Considering positive time-delay ($\tau > 0$) in the measurements causes important changes of the effective potential. For increasing values of τ, the accumulation region becomes shallower, projecting further than the radius of the cell (Fig. 4.6a). This is a consequence of the displacement of the cell (and the captured beads) between time t, at which the positions of cells are considered, and the time $t + \tau$ at which the bead distribution is measured, in order to compute the RDF, Eq. (4.2). The mean displacement ℓ, of the cell between these two times, t and $t + \tau$, is related to the diffusion constant by $\ell(\tau) = \sqrt{4D\tau}$.

Interestingly, as the delay τ is increased further the attractive range around $r = 0$ in the effective potential disappears, shifting to a purely repulsive potential (Fig. 4.6a for $\tau > 0$). This happens despite the attractive nature of the *capturing* mechanism, the only explicit interaction in the model. To understand how and when this transition occurs, it is necessary to look at the (t, τ)-phase-diagram for the numerical results with positive delay, $\tau > 0$.

4.6.1 Effective repulsion

As the only characteristic time-scale in the system is given by the diffusive time-scale of cells, it is expected that the transition between the two regimes, attractive and purely repulsive, is controlled by the diffusivity $D = \delta^2/4$, introduced above. In numerical simulations of the model introduced at the beginning of this section, it is observed that an increase in the step length results in a decrease of the offset delay τ_{off} at which a transition of the effective potential from attractive to repulsive occurs (indicated by ► in Fig. 4.6b). The effective repulsion is characterised by the effective potential being positive at the origin and decaying towards zero with increasing distance. The offset delay τ_{off} is related to the time it takes for a cell to displace itself and the beads attached to it, by a distance of the order of the cell size, i.e.

$$\tau_{\text{off}} \approx \frac{r_0^2}{D}, \tag{4.9}$$

as shown in Fig. 4.6b (last panel), effectively transporting all captured beads away from the point of reference considered for measuring the effective potential from. This transport process manifests itself as an effective repulsion.

This offset explains why in the experiments the purely repulsive potential is not observed. The values for the diffusion constant of the cells in the experiments varies slightly between conditions, with a value of $D \approx 0.21$ μ^2/min for cells in the presence of uncoated- and Wnt- beads, and $D \approx 0.23$ μ^2/min for cells in the presence of iWnt-beads (see Sec. D.1.1 for details). Using the average radius of the mESCs, $r_{\text{sc}} \sim 10\,\mu$m, an estimate for the expected offset delay τ_{off} in the experiments can be obtained from Eq. (4.9). Which is $\tau_{\text{off}} = 470$ min for Wnt-beads, and $\tau_{\text{off}} = 430$ min for iWnt-beads. These predictions for the onset of the transition of the effective potential turning repulsive, are compatible with the experimental observations. For the earliest experimental time-window ($t = 61$ min), the smallest delays for which the accumulation region around $r = 0$ crosses the reference $U^{\text{eff}} = 0$ towards positive values are $\tau = \tau_{\text{off}} \sim 420$ min for Wnt-beads (Fig. D.3, first column), and $\tau = \tau_{\text{off}} \sim 360$ min for iWnt-beads (Fig. D.3, first column).

Regimes that are inaccessible in the experimental setup can be studied numerically. For $t \sim 0$, once the offset delay has been exceeded, i.e. $\tau > \tau_{\text{off}}$, the potential transitions to effective repulsion (dashed lines, Fig. 4.6b). For later times $t > 0$, the transition occurs after longer delays, taking place when the displacement $\ell(t) = \sqrt{4Dt}$ of the cell between times t and $t + \tau$, is of the order of the effective radius $\sqrt{A/\pi}$ of the area $A(D, r_0, t + \tau)$ explored by the cell until time $t + \tau$,

$$4\pi D\tau = A(D, r_0, t + \tau). \tag{4.10}$$

The transition occurs therefore at

$$\tau^* = \frac{\exp\left(1 + \frac{t}{\tau^*}\right)}{\left(1 + \frac{t}{\tau^*}\right)} \frac{r_0^2}{\beta D} \tag{4.11}$$

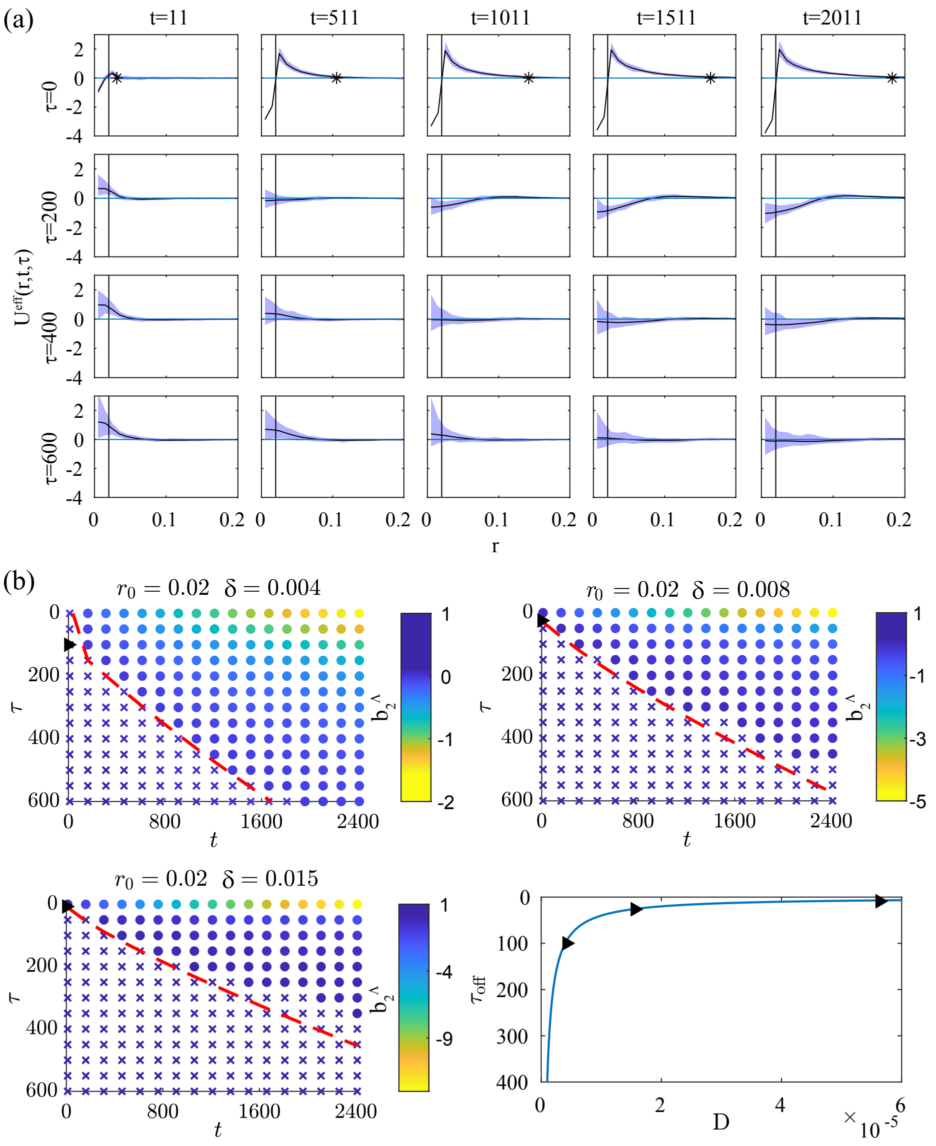


Figure 4.6: **Numerical age- and delay-dependent interactions.** (a) Effective potential as function of age t and delay τ obtained from numerical simulations of the model described at the beginning of Sec. 4.6, indicating the mean (solid curve) and range (shaded) from 10^4 realisations, for step size $\delta = 0.004$ and cell radius $r_0 = 0.02$ (vertical line), the symbol ($*$) indicates the interaction range $r_e(t) = \sqrt{A/pi}$. (b) The (t, τ)-phase-diagram for increasing values of step length δ. Markers indicate attractive (circles) and repulsive (crosses) regions in the phase diagrams, as measured by b_2^Λ (Eq. (4.5), colorbar) for $\Lambda = 2R$, where R is the cell radius. The symbol (▶) indicates the predicted offset delay, Eq. (4.9), the dashed line corresponds to the theoretical transition delay τ^*, obtained from Eq. (4.11). The potential is effectively repulsive below the line (longer delays) and effectively attractive above (shorter delays). The last panel shows the offset delay as function of diffusivity, $D = \delta^2/4$, obtained from Eq. (4.9), the triangles indicate the delay marked on the phase-diagrams. This figure has been submitted to PRX as part of the manuscript [7]

using Eq. (4.7) valid for $t \gg r_0^2/D$. At larger delays the apparent displacement of beads away from the effective origin exceeds the effective range of interaction, $\sqrt{A/\pi}$, so that the potential appears to be repulsive. Solving numerically the transcendental Eq. (4.11) for τ^* produces the dashed transition lines shown in Fig. 4.6b. Below this line (longer delays), the potential appears to be typically repulsive, above this line (shorter delays) typically attractive.

The delayed effective potential serves as a tool to quantify the transport of signals (beads) by cells. In numerical simulations, motility is the only mechanism available to cells to encounter beads. However, in the experimental conditions it was observed that mESCs constantly change their shape, as they elongate and increase their interaction range without net displacement. Moreover, mESCs do not necessarily capture every single bead they encounter, disregarding some of the beads they come in contact with. Furthermore, beads can spontaneously detach from the cell membrane. In the next section, the effect on the effective potential of both of these mechanisms is studied independently, by incorporating an *disregarding* probability and a Poissonian release of beads to the model. Two possible release scenarios are considered, one in which beads are recruited by the cell and subsequently release in a random direction, and a second one, where cells can repeatedly interact with beads. Thus effectively *pushing* them away by releasing them in the same direction they were encountered.

4.7 Disregarding and releasing beads

The effect of mESCs disregarding a fraction of beads encountered can be incorporated into the master equation (4.6) by introducing a *disregarding probability* χ. The density of beads that a cell can interact with is then effectively reduced to $\rho_{\text{eff}} = (1-\chi)\rho_b$, with $\chi \in [0,1]$. This represents the *affinity* of the cell-signal interactions. The master equation in this case reads

$$\frac{dP_\chi(n,t)}{dt} = \rho_{\text{eff}} \frac{dA}{dt}[P_\chi(n-1,t) - P_\chi(n,t)]. \tag{4.12}$$

As χ approaches unity, an increasingly large proportion of beads is disregarded, thus increasing the mean free path of the cell. By introducing rescaled time $t' = (1-\chi)t$ and cell radius $r'_0 = \sqrt{1-\chi}r_0$ in Eq. (4.12), where A is given by Eq. (4.7), the form of Eq. (4.6) is recovered. This shows that larger values of χ imply a slower evolution of the effective potential (see Fig. 4.7). However, its qualitative properties, such as the potential overall shape, and the existence of the transition between attractive and repulsive regimes are not affected by the introduction of the disregarding probability (see Figure 4.8a). Disregarding beads could be responsible for the differences observed in the properties of the effective potentials between Wnt-beads (Fig. 4.5a) iWnt-beads (Fig. 4.5b). As described above, the effective potentials and phase-diagrams evolve in qualitative similar manner, but with the dynamics sped up for the signalling Wnt-beads. This suggests that cells have a lower affinity for non-signalling iWnt-beads. Thus, cells are less likely to recruit iWnt beads.

As suggested by the measurements of the total time beads stay in contact with cells (Fig. 4.4c), cell-bead bonds are reversible, so that beads occasionally detach from the cell membrane getting released to the environment. To account for this process a Poissonian release rate $\nu \in [0,\infty)$ is introduced. In the numerics, a time step $\Delta t = 1$ is considered. Hence, on every iteration, the probability κ that an attached bead will be released by the cell is given by $\kappa = 1 - \mathrm{e}^{-\nu}$. Released beads are considered to simply decouple from the movement of the cell, and their position is held fixed during further evolution of the process, so that the cell disregards released beads without ever

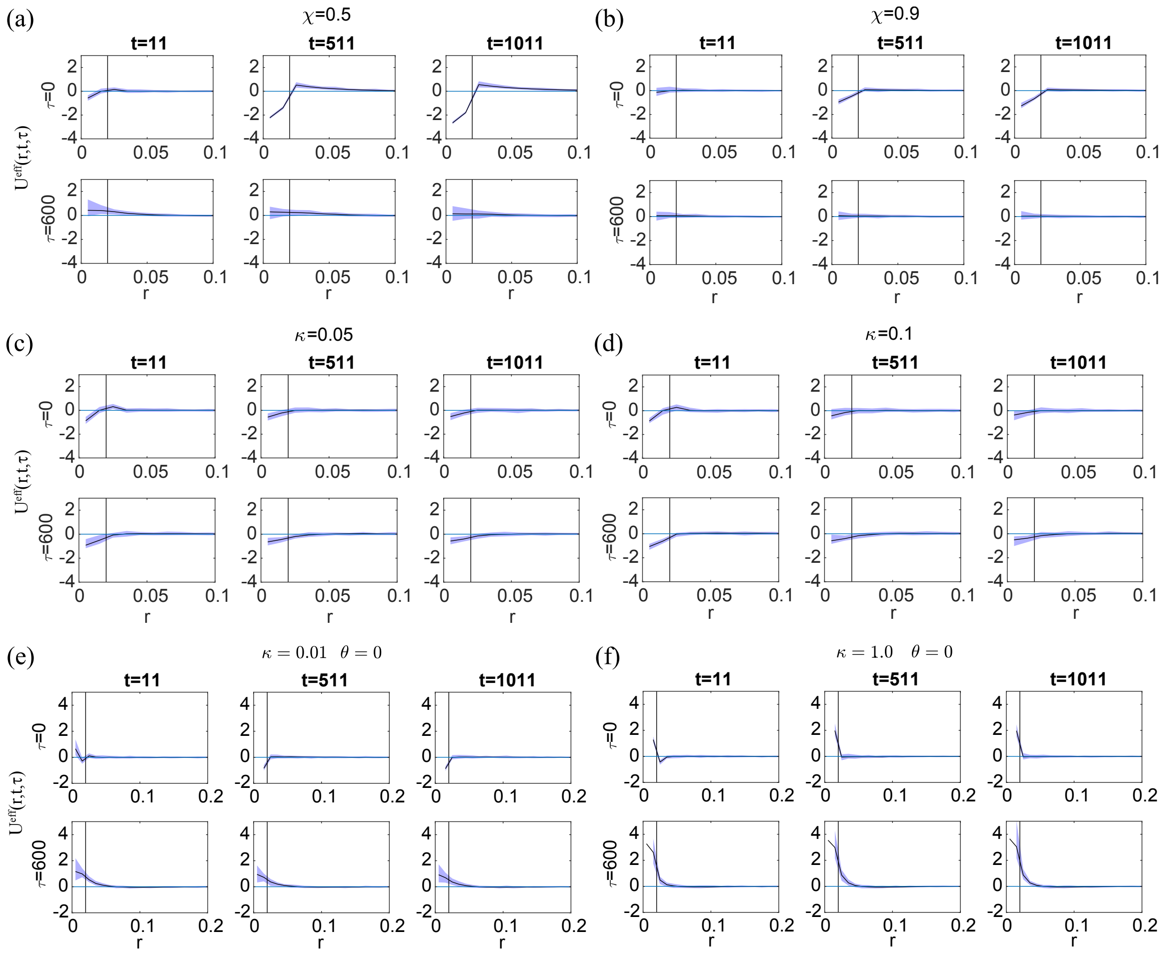


Figure 4.7: **Effective potentials with positive disregarding probability and release rate.** Effective potential as a function of age t and delay τ, as obtained from numerical simulations (Sec. 4.7). The mean and range of 10^4 realisations are indicated by a solid curve and a shaded area, respectively, for disregarding probability (a) $\chi = 0.5$ ($\kappa = 0$), and (b) $\chi = 0.9$ ($\kappa = 0$). For release probability (c) $\kappa = 0.05$ ($\chi = 0$), and (d) $\kappa = 0.9$ ($\chi = 0$), and for refractory time $\theta = 0$, and release to the cell edge with probability (e) $\kappa = 0.01$, and (f) $\kappa = 1.0$. Simulations consider a single cell of radius $r_0 = 0.02$ performing a random walk with step length $\delta = 0.004$, in a periodic system of linear size $L = 1$, and bead density $\rho_b = 10^3$. This figure has been submitted to PRX as part of the manuscript [7].

recapturing them. A diffusing cell will therefore leave a trail of released beads along its path. An advantage of this type of implementation for the release mechanism is that lends itself to a description in terms of a master equation. Given the probability $P(n, t)$ of observing n beads in contact with a cell at time t, the rate of decay of that state into a state with $n - 1$ beads is $\kappa n P(n, t)$. Taking this into account, the master equation for the probability $P_\kappa(n, t)$ reads

$$\begin{aligned}\frac{dP_\kappa(n,t)}{dt} &= \rho_b \frac{dA(t)}{dt}[P_\kappa(n-1,t) - P_\kappa(n,t)] \\ &\quad +\kappa[(n+1)P_\kappa(n+1,t) - nP_\kappa(n,t)].\end{aligned} \tag{4.13}$$

This master equation can be solved via a generating function approach (see Sec. D.1.5 for details) to find the explicit solution

$$P_\kappa(n,t) = \frac{\alpha^n(A,\kappa;t)}{n!} e^{-\alpha(A,\kappa;t)}, \tag{4.14}$$

where $\alpha(A,\kappa;t)$ is a function of the area $A = A(D, r_0, t)$ explored, Eq. (4.7), the release rate κ and time t,

$$\alpha(A,\kappa,t) = \rho_b \left[A(t) - \kappa e^{-\kappa t} \int^t \mathrm{d}s\, A(s) e^{\kappa s} \right]. \tag{4.15}$$

Given the structure of the solution, Eq. (4.14), α is in fact the expected number of beads attached to a cell, $\langle n \rangle = \alpha(A,\kappa,t)$. For $\kappa = 0$, Eq. (4.8) is recovered, where $\langle n \rangle = \rho_b A(t)$. For $\kappa > 0$, α is a concave function (see Fig. D.1), hence there is a time $t = t^*$ for which the Poissonian release overcomes the capturing of beads by cell diffusion. Before $t = t^*$ the average number of beads captured, $n(t)$, and consequently the accumulation strength of the effective potential, increases with time. At $t = t^*$, the parameter $\alpha(A,\kappa;t^*)$ vanishes and $n(t)$ reaches its maximum value. For $t > t^*$ the number of beads decreases and effective potential weakens, asymptotically recovering the non-interacting behaviour $U^{\text{eff}} = 0$.

The impact of the release mechanism on the effective potential (Fig. 4.7) and phase-diagram (Fig. 4.8b) becomes most evident at positive delay, $\tau > 0$. For increasing values of the release rate, the transition curve between attractive and repulsive phases shifts towards shorts times t (Figure 4.8b). For large values of κ the phase-diagram is inverted compared to $\kappa = 0$, as the attraction of the effective potential decreases with age t, and increases with delay τ, while the repulsive regime disappears.

Even though the phase diagram appear inverted in Fig. 4.8b for $\kappa = 0.1$ when compared to $\kappa = 0.001$, the fact that this phase diagram exhibits purely attractive behaviour implies this is not the mechanism controlling the interaction between mESCs and Unc-beads (Fig. 4.5c). A key observation of the experiments is that mESCs do interact with Unc-beads, yet are unable to recruit them, as quantified by the effective potential (Fig. 4.5c). In the experiments, when a cell encounters Unc-beads, beads tend to remain on the cell periphery. As the cell moves, it effectively *pushes* beads in its direction of movement. As the upper-critical dimension of a random walk is two, the pushing dynamics requires cells to interact repeatedly with every bead they encounter. To incorporate this into the numerics, and to take into account possible memory effects, a *refractory time* θ is introduced, which set a minimum time needed by a cell to re-interact with a released bead. In the limiting case of $\nu \to \infty$ (κ=1) and $\theta = 0$, beads are continuously pushed away by the cell. For finite values of ν ($\kappa < 1$), beads can be transported by the cell and released at a later time point. The release in this pushing dynamics is implemented by placing the bead on the point of the cell periphery, where it was first encountered. This accounts for experimental observation that the bead is not transported *on* the cell membrane bulk, but near its boundary.

Interestingly, the phase diagram for the *pushing* mechanism (Figs. 4.8c) for $\theta = 0$ (such that the cell can re-interact with beads on every iteration) shows that with increasing release rates, the attractive region of the effective potential disappears. For $\kappa \approx 1$ the phase diagram transitions from being repulsive for very short time and delay, to purely repulsive for longer time and delay, a feature previously observed in the experiments of mESCs with Unc-beads (Fig. 4.5c). In both the simulations and experiments the effective interactions become stronger, more repulsive with both age t and delay τ, in contrast to was was observed for coated beads (Figs. 4.5a and 4.5b).

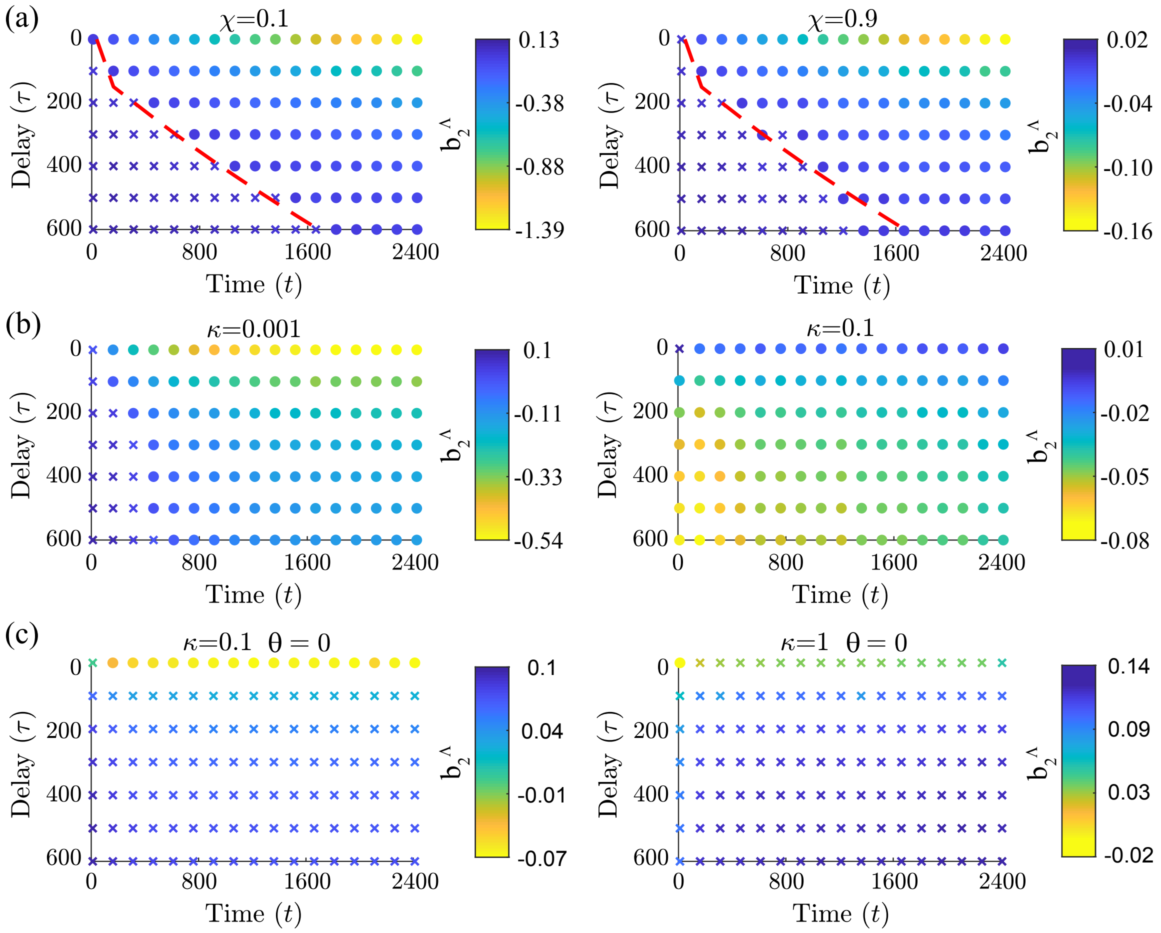


Figure 4.8: **Phase diagrams for the model with disregarding probability, release rate, and recapturing.** (t, τ)-phase diagram of the model with (a) positive disregarding probability χ, (b) positive release probability κ, and (c) positive release probability κ with refractory time $\theta = 0$, for increasing values (left to right). Simulations consider a single cell of radius $r_0 = 0.02$ performing a random walk with step length $\delta = 0.004$, in a periodic system of linear size $L = 1$, and bead density $\rho_b = 10^3$. The dashed lines in (a) represent the predicted transition from attractive to repulsive according to Eq. (4.11). Markers indicate attractive (circles) and repulsive (crosses) regions in the phase diagrams, as measured by b_2^Λ (Eq. (4.5), colorbar) for $\Lambda = 2R$, where R is the cell radius. In (c) the release of beads occurs to the same direction they were encountered by the cell, effectively *pushing* beads away. This figure has been submitted to PRX as part of the manuscript [7]

The results presented in this section, together with the effective retention times (Fig. 4.4c), suggests that the cell signal interactions are strongly affected by particular disregarding and releasing mechanisms. The absence of coating seems to impede the cells' capacity for actively recruiting beads and thus organise their environment.

4.8 Entropy generation

The ability of cells to interact with and organise their environment is evidenced by the emergence of the effective potential of interaction. As observed experimentally, and demonstrated through numerical simulations, different cell-signal interaction mechanisms can contribute to a higher or lower organisation in the spatial distribution of

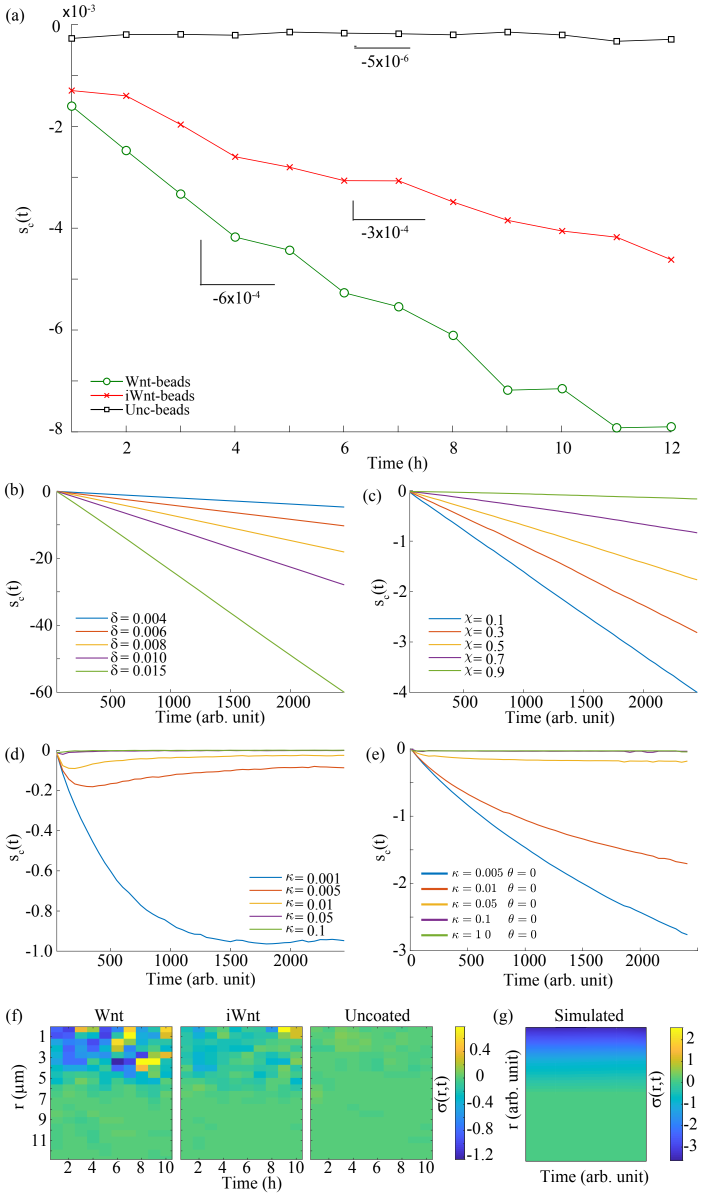


Figure 4.9: **Configurational entropy as a function of time** for (a) the experimental conditions, (b) the numerical simulations of the model the model described at the beginning of Sec. 4.6 for different step lengths δ, (c) the model with positive disregarding probability χ, Eq. (4.12), (d) the model with positive release probability κ, Eq. (4.13), and (e) the model with positive release probability κ and refractory time $\theta = 0$. The dashed line in (a) corresponds to the reference line $s_c = 0$. The rate of change of entropy density $\sigma(r, t)$, Eq. (4.17), is shown for (f) the experimental conditions, and (g) the numerical simulations at $\chi = \kappa = 0$. Simulations in (b)-(f) consider a single cell of radius $r_0 = 0.02$, in a periodic system of linear size $L = 1$, and bead density $\rho_b = 10^3$. The step length for the simulations in (c), (d) and (f) is set to $\delta = 0.004$. In (e) the release of beads occurs to the same direction they were encountered by the cell, effectively *pushing* beads away. This figure has been submitted to PRX as part of the manuscript [7]

beads.

Changes in the distribution of beads affects the excess configurational entropy per cell s_c, a quantitative measure of the level of spatial order. In the experimental setup considered here, where the densities of both cells and beads are low, the configurational entropy per cell s_c can be written perturbatively, to first order, in terms of the radial distribution function [152],

$$s_c = \pi\rho_c \int \mathrm{d}r \left[\hat{g}U^{\mathrm{eff}} + (\hat{g} - 1)\right] r, \tag{4.16}$$

where $U^{\mathrm{eff}} = -\ln(\hat{g})$ and ρ_c is the cell density. The configurational entropy production therefore reads $\dot{s}_c(t) = \pi\rho_c \int \mathrm{d}r\, \sigma(r,t) r$, with entropy production density

$$\sigma(r,t) = \dot{\hat{g}}(r,t,\tau = 0)U^{\mathrm{eff}}(r,t,\tau = 0). \tag{4.17}$$

For a non-interacting homogeneous system, the radial distribution function $\hat{g}$ is unity and the effective potetial U^{eff} vanishes, so that the entropy production $\dot{s}_c$ vanishes identically. The configurational entropy decreases as the system becomes more organised, decreasing the total entropy of the bead distribution. This entropy reduction carries an energy cost, which is provided by the active cells that transport beads and keep the system out of equilibrium.

As expected from the observation of attractive effective potentials for Wnt- and iWnt-beads (Figs. 4.3d and Fig. 4.3e), the dynamics of these systems tends to decrease the configurational entropy s_c, lowering the total entropy of the bead distribution. In both cases the entropy rate is roughly constant (Fig. 4.9a), with the configurational entropy decreasing linearly in time. The entropy for Wnt-beads decreases at a rate almost twice that of non-signalling iWnt-beads, suggesting that the biochemical signalling of Wnt-beads contributes to a faster organisation of the system. The configurational entropy of Unc-beads on the other hand, shows no significant change in time. It remains near the non-interacting reference $s_c = 0$ for the 12 hour period of observation, demonstrating the lack of recruitment and organisation of uncoated beads by stem cells.

As argued in the previous section, the behaviour of mESCs with Wnt-beads is well described by the simplest model, described at the beginning of Sec. 4.6, for which a linear decrease of configurational entropy over time is observed (Fig. 4.9b). The rate depends on the step length and thus the diffusion constant of the cells. However, the diffusivity does not vary significantly between experiments and therefore does not suffice to explain the higher entropy rate observed when cells interact with Wnt-beads compared to non-signalling iWnt-beads. By introducing the *disregarding* probability χ (Fig. 4.9c), the experimental observations could be replicated, where for a fixed diffusion constant and positive probability $\chi > 0$ the entropy generation is reduced compared to $\chi = 0$, but remains constant in time.

When taking into account the release of beads with probability κ (Fig. 4.9d) the entropy reaches a minimum when the mean number of beads captured is maximum (see Fig. D.1), whereafter the entropy generation becomes slightly positive before levelling off, so that $s_c(t)$ reaches a plateau close to 0, asymptotically. When considering the *pushing* mechanism describe at the end of the previous section, the entropy also reaches a plateau (Fig. 4.9e). There, high release probability κ imply the maintenance of lower level of entropy. This behaviour is also observed in the case of Unc-beads (see Fig. 4.9a), where the entropy remains constant and no signs of a minimum are observed. The observations for Unc-beads could also be the result of a combined effect of both disregarding and releasing mechanisms. However only in the cases where a positive release probability κ is considered the configurational entropy will remain constant over time, otherwise it will decrease linearly in time.

The entropy production density $\sigma(r, t)$ shows that, in the case of signalling Wnt-beads (Fig. 4.9e, left panel), the largest contribution to negative entropy generation occurs around $r = 0$, close to the cell centre. This is consistent with numerical simulations (Fig. 4.9f), where, in addition, $\sigma(r, t)$ remains approximately constant over time t. The density σ is significantly reduced for non-signalling iWnt beads and essentially vanishes for Unc-beads (Fig. 4.9e). This is expected, as the temporal changes in their effective potentials are greatly reduced compared to Wnt-beads (Figs. D.3-D.5).

4.9 Discussion and outlook

In this chapter a simplified experimental setup was used, in which single active cells interact with signals that are immobilized on the surface of micron-sized beads. Time-lapse imaging at high spatio-temporal resolution was performed to follow the dynamics of single cells interacting will homogeneously distributed beads. By introducing the age- and time-dependent effective potential the spatio-temporal dynamics of cell-signal interactions were characterised as an ageing system of two interacting species of particles.

Effective potentials have been extensively used in the stochastic and hydrodynamic modeling of active matter [123, 124]. However, until now these had only been measured experimentally in stationary, confluent cell sheets [128]. The results presented in this chapter are the first measurement of ageing effective interaction potentials at the single cell level. They encapsulate information not only of the instantaneous stem cell-ligand interactions, but also of their spatial and temporal evolution, bridging across scales. They specify the location, strength and duration of these interaction on the cell membrane. Furthermore, it was shown that the ageing effective potential provides a previously unexplored route for extracting the change in configurational entropy of a living system. This allows the direct measurement of the organisation of the environment by cells.

Through the comparison of experimental observations for mESCs with numerical simulations it was shown how short and large-scale organisation of beads, as measured by the effective potential and configurational entropy, emerge from the *microscopic* single cell-signal interaction. Analytics and numerical simulations allowed the distinction between three different mechanisms for cell-bead interaction. In the case of signalling Wnt-beads, the interactions are dominated by a *capturing* dynamics, where cells encounter and retain beads in a seemingly irreversible manner, efficiently transporting signals. In the case of non-signalling iWnt-beads, the reduction in the interactions was associated with a reduced affinity of cell-bead interactions, as if cells were *disregarding* a fraction of beads in their environment. Finally, the interactions between cells and Unc-beads were found to be very short-lived, with their lack of recruitment being crucial for reproducing the experimental phase diagrams. For this, in the simulations cells were allowed to re-interact with beads. This, combined with a high release rate, results in an effective pushing dynamics, where cells displace beads as they diffuse. It turns out that the properties of the effective potential obtained for $\theta = 0$ are not affected by moderate increases in the refractory time (see Fig. 4.10), which implies that the results presented here are robust to possible memory effects.

Even though it is known that a suitable biochemical coating is necessary for the formation of durable cell-

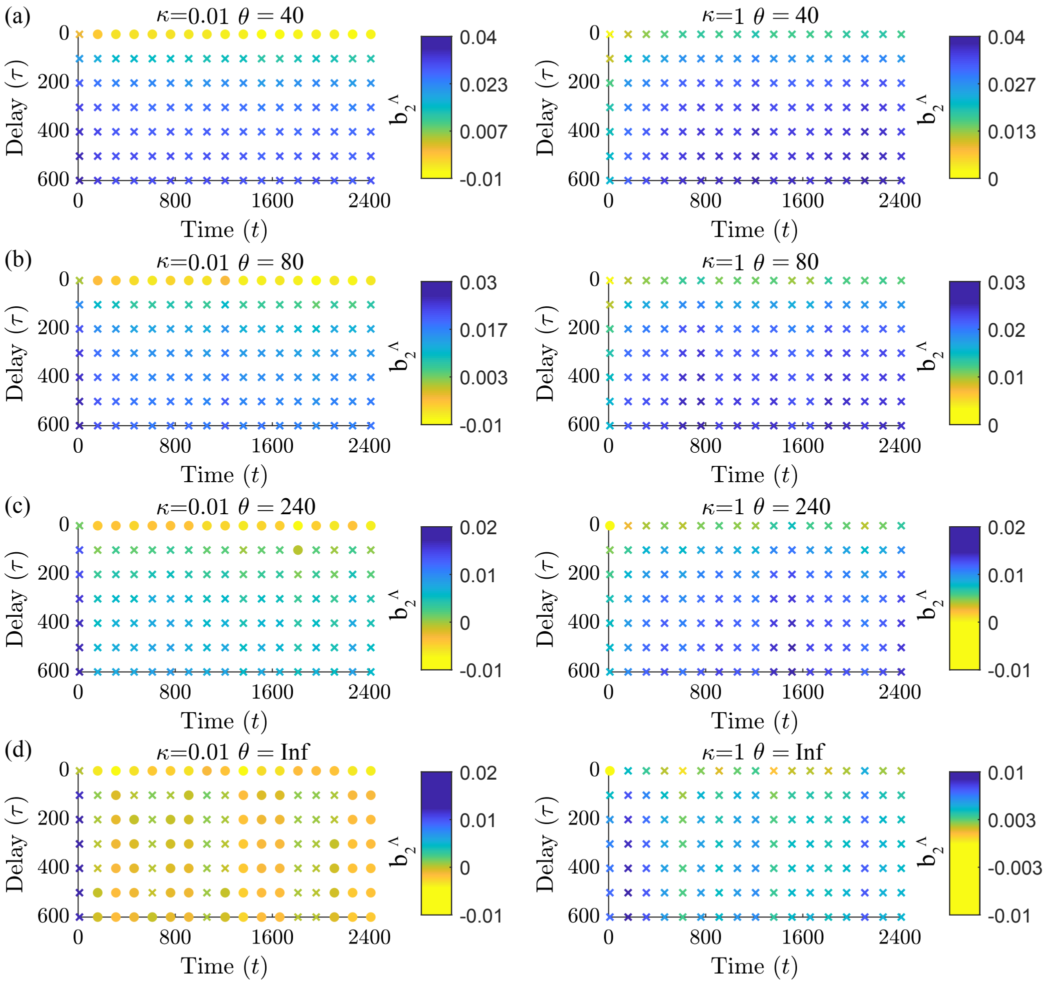


Figure 4.10: **Phase diagram with positive refractory time θ.** Markers indicate attractive (circles) and repulsive (crosses) regions in the phase diagrams, as measured by b_2^Λ (Eq. (4.5), colorbar) for $\Lambda = 2R$, where R is the cell radius, for simulations of the model (Sec. 4.7) with positive refractory time θ and release to the cell edge with probability κ, as indicated. As mentioned in the main text, moderate increases in the refractory time (see (a) and (b)) do not change the qualitative features of the phase diagram. However, larger values of the refractory time, such as (c) $\theta = 240$, or (d) $\theta \to \infty$, which implies no re-interaction, do exhibit some noticeable fluctuations in the phase diagram, especially for increasing time.

bead bonds [146, 147], here it was shown that the signalling nature of the coating plays a more complex role in these interactions. This ranges from the recruiting efficiency to the location of the signal on the cell membrane. In the particular setup considered here, mESCs show a higher affinity for signalling (Wnt3a-coated) Wnt-beads compared to the non-signalling iWnt-beads. Wnt-beads promote more effective interactions and organisation by cells, over time, as seen in Fig. 4.1c. Future research is needed to pinpoint the molecular mechanisms behind the different cell-signal interactions and to explain the observed selectivity. Furthermore, it would be of great interest to determine if this selectivity to particular signals is unique to embryonic stem cells or applies to other cell types.

Numerical simulations allowed the exploration of experimentally inaccessible regimes, which provide deeper insight into the properties of the ageing effective potential and the properties of cell-signal interactions. The existence of the previously unobserved repulsive region in the phase-diagram of the effective potential (such as

Fig. 4.6a) was attributed to the efficient transport of signals by cells. Further, the transition of the potential, from attractive to repulsive, was shown to be directly related to an efficient transport of signals by the cells. Its timing is controlled by the diffusive properties of cells together with the microscopic (capture, disregard and release) mechanisms of cell-signal interactions. It is expected that in experimental setups considering, for example migratory phenotypes, such as cancer or mesenchymal stem cells, the repulsive phase will be observed even when cell signal interactions are dominated by a capturing mechanism. This might have relevant implications on the dynamics and invasive capabilities of these cell types.

By casting the configurational entropy of the system in terms of the effective potential, the efficiency with which cells influence and organise their environment could be quantified, as a function of time, and also as a function of the properties of the beads' coating and thus of the signals. It was observed that relevant biochemical signals contribute to a faster organisation of the environment (Figs. 4.9a and 4.9e). Despite the similarities between the temporal evolution of the coordination number $c(t)$ (Fig. 4.4a) and the configurational entropy s_c (Fig. 4.9a), the nature of both observables is different. The former is a measure only of the *net* accumulation of beads in a specific range of the effective potential, while the latter incorporates the whole spatial information of the effective potential, taking into account the structure and organisation of the bead distribution. The lack of organisation of uncoated beads explains why even when observing an increase in the coordination number for uncoated beads (Figure 4.4a), no significant decrease in the configuration entropy is observed over time (Fig. 4.9a). Moreover, in the case of both Wnt- and iWnt-beads, it is observed that the coordination number $c(t)$ reaches a plateau at the 7 and 10 hour mark, respectively. However, in both cases, the configurational entropy maintains a constant rate for the whole duration of the analysis. The fact that the entropy continues to decrease despite no further bead accumulation, indicates that cells are actively positioning the beads they contact at preferred locations on the cell membrane. These preferred locations correspond to the minima of the effective potentials of interaction. Although the biological implications of this observation remains open for future work, it may be hypothesized that the beads' location might be controlled by Wnt-receptors and downstream effectors, and the dynamics of the cellular membrane.

In the work presented in this chapter, experimental measurements of the motility of cells were performed, which were used to interpret the effective potential observations. It is worth noting that the diffusion constant measured for the cells (~0.2 μm^2/min) is very low, allowing cells to travel an average distance of the order of a cell diameter in the course of 12 hours. Together with the low density of beads, which implies an average distance between beads of the order of 100 μm, one might have though that cells would not be able to interact with more than one bead in their surroundings. However, the experimental observations (see Fig. 4.1) show that long range cell-bead interactions occur though thin membrane projections that recruit beads efficiently. The model presented here, despite not considering this membrane features, does account for them though an increased diffusivity. Future work shall focus on understanding the properties of both types of interactions, one based on changes of body shape and the other one on net displacement.

Other parameters present in the model, such as the release rate and disregarding probability could not be measured directly. This is because the cells' boundaries could not be detected accurately due to the lack of a membrane marker, which were not used to avoid external influence on the cell behaviour (see Fig. 4.2). It is important to remember that Wnt does not diffuse away from the beads [130], thus the interactions, recruitment, release or disregard, rely completely on direct cell bead contact. Distinguishing between disregarding and release would then require high spatial resolution to detect precisely when a cell comes and remains in contact with a bead versus coming very close but not touching the cell. As mentioned earlier, future work shall focus on the molecular interactions at the cell membrane level to determine the origin of the different interaction mechanisms.

An advantage of the effective potential approach is that it does not rely on the precise detection of cell boundaries but rather on the, much easier to determine, location of the cell center. The methods presented in this chapter are applicable to a large number of systems. This includes, but is not limited to other types of stem cells and protein signals both in 2D and 3D environment, such as bio-engineered scaffolds or in vivo conditions. This could provide clues to better understand the formation of specialised micro-environments in multicelullar tissues, such as the stem cell niche, which rely on the precise spatial organisation and localization of cells and signals.

In this and the previous chapters, the focus has been on the study of spatio-temporal dynamics. These ranged from the study of the spatial BRW and the spatial patterning of vegetation, to the interaction of embryonic stem cells with spatially dispersed ligands. In the next chapter, the focus is solely on temporal dynamics. Going deeper into the molecular machinery of stem cells, it is shown how ideas from statistical mechanic and dynamical systems can be used to decipher the intricate process of stem cell fate decision making.

Chapter 5

Pulsatile ERK activity and its role on stem cell fate

The results presented in this chapter are part of a manuscript entitled *Pulse kinetics of ERK MAPK controls epidermal stem cell states*, co-authored by Toru Hiratsuka, Gunnar Pruessner and Fiona M Watt. Some parts of this chapter are quoted verbatim from the manuscript [8].

I developed the detection and analysis methods of ERK pulses, including the construction of phase diagrams and simulations, and interpreted the results. I wrote the manuscript together with Toru Hiratsuka, who performed all experiments. Of the figures presented in this chapter, I prepared Figs. 5.2, 5.3c-5.3e, 5.4 and 5.6a-5.6d, while Toru Hiratsuka provided me with Figs. 5.1, 5.3a and 5.3b, 5.5, 5.6e, 5.7 and 5.8.

5.1 Summary

In the previous chapter, it was discussed how embryonic stem cells probe their environment for instructive signals. Cell signalling can activate pathways that promote differentiation, changes in behaviour, cell division or death. The ways in which the downstream effectors of these pathways interact with each other at the molecular scale are, in many cases, only partially understood. In this chapter, primary neonatal human keratinocytes (NHKs, strain km) are used to study the Extracellular signal-Regulated Kinase (ERK) pathway, which behaves as an on/off switch and is known to control cell proliferation. ERK is involved in the control of a variety of cellular processes including cell proliferation, migration and differentiation. It is known that this pathway exhibits pulses of activity, a behaviour common of multiple signal transduction pathways. However, until now the relation between ERK activity pulses and stem cell fate had not been studied. By using molecular sensors it was possible to track, in real time, the activity or ERK together with the level of differentiation of NHKs, as measured by Involucrin expression. A method was developed to study the co-evolution of these two signals, ERK and Involucrin, which allows the construction of a stem cell state phase diagram. With this, different stem cell states can be distinguished, which are characterised by their mean levels and pusatile activity. Here, it is shown that the transition between states is

mediated by pulsatile ERK activity. It is shown that ERK mean level and pulses are independently regulated by DUSP10 and DUSP6, respectively. These are two components of the unstable autoregulatory protein phosphatase network that controls the transition between the stem and differentiated cell states. The results presented in this chapter show that pulsatile ERK state mediates the transitions between epidermal stem cell state.

5.2 Introduction

The stem cells state is an unstable state of the cell. If left isolated from external stimuli a stem cell will eventually decay into a state where its pluripotency may be partially or totally lost. As mentioned in the previous chapter, this process can be controlled by external biochemical or mechanical cues. At the beginning of the differentiation process, stem cells go though a stage of *commitment*, where the internal biochemistry of cells changes though the activation or deactivation of molecular pathways [153, 154]. These changes determine the ultimate state of the cell, commonly referred to *cell fate*.

The interactions between signalling pathways often leads to rich temporal dynamics of their activity and protein expression. Hysteresis loops, cycles, intermittent behaviour, and oscillatory dynamics [155] have all been observed in the activity of pathways inside the living cell [156, 157]. In this chapter, the focus is on the pulsatile behaviour of a specific molecular pathway, ERK, and its relation to the cell state and role in cell commitment and differentiation. Pulsating signal activity differs from oscillations in that pulses are localised in time, i.e. are not necessarily regular. Pulses are common to several cellular signal transduction pathways [158, 159] and are present in diverse cell types, from microbes to mammalian cells [158]. It is argued that the properties of the pulses, such as frequency, amplitude and duration encodes information [159] that can, for example promote the (de)activation of target genes.

As argued in [160, 161], common steady state experiments where observables are averaged over populations of cells can mask the existence of activity pulses or oscillations, for this reason it is imperative to look at the time evolution of pathways at the single cell level, where the role of pulses in the cellular response to external or internal stimuli can be assessed. Following this single cell approach, Lahav et al. [160] found that pulsatile activation of the tumor suppressor p53 occurs in response to increasing radiation-induced DNA damage. Moreover, they found that the amplitude and duration of the pulses is independent on the level of DNA damage, while the frequency of pulses decreases with the radiation dose.

Pulsatile behaviour has also been observed in ERK activity. ERK is a downstream target of multiple signals, including integrins (proteins related to cell adhesion) and epidermal growth factor (EGF, related to cell proliferation) [162]. Thus, ERK is involved in the control of diverse aspects of the cellular behaviour including cell proliferation, migration and differentiation [162]. ERK activity is regulated by phosphorylation (activation) and de-phosphorylation (de-activation) via upstream kinase activators and phosphatases [162, 163]. Through live imaging approaches, the spatio-temporal dynamics of ERK has been observed in a variety of cell types (in-vitro) and organisms (in-vivo) [161, 164, 165, 166, 167]. This includes the observations of temporal pulses, but also the spatial propagation of two dimensional activity fronts or bursts in cellular tissues. In murine (mouse) epidermis, the outer cover of the skin, ERK pulses and burst have been studied in steady state and during wound healing [165]. The epidermis is a layered structure, stem cells reside in the basal layer (deepest layer), these either self-renew or

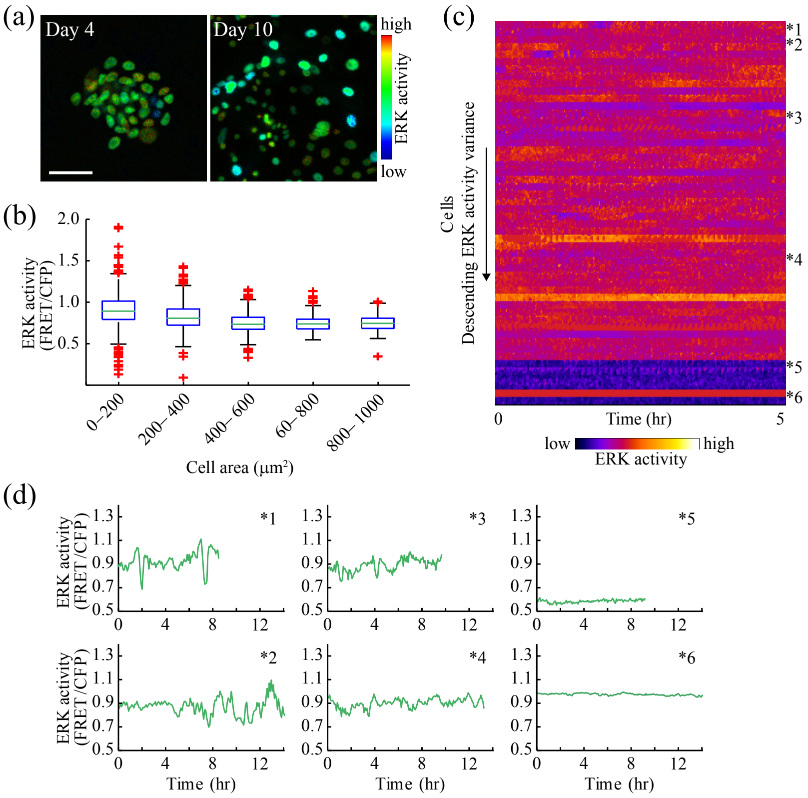


Figure 5.1: **ERK activity and cell size.** (a) Images of NHK colonies expressing EKAR-EVnls. Colours indicate ERK activity. Scale bar 50 μm. (b) Box plots of single cell ERK activity as a function of cell area. Mid-line corresponds to median; box to 25th and 75th percentiles; whiskers to lower and higher 95 percetiles. Red crosses correspond to outliers. (n = 3581 cells). (c) Shows the heat-map of ERK activity as a function of time for 52 cells in descending order of overall ERK activity variance. Colours indicate ERK activity. (d) Representative time-series of ERK activity in cells indicated by asterisks in (c). This show pulsatile (*1-4), stable-low (*5), or stable-high (*6) ERK activity profiles. This figure is part of the manuscript [8].

produce committed cells that move upward into the suprabasal layers as they undergo terminal differentiation.

ERK is required for epidermal stem cell development and maintenance. It has been observed that deregulation of ERK signalling leads to skin abnormalities and proliferation-related disease, such as psoriasis, chronic skin inflammation and cancer in mouse and humans [168, 169, 170, 171, 172, 173]. A variety of data analysis and modeling approaches have contributed to the better understanding on the interaction between ERK and other pathways. The Abstract Boolean Network (ABS) [174] approach considers a binary gene expression network of the cell, where each gene is considered "on" ("off") if the level of expression is above (below) the mean expression among all genes [154]. Following this approach Mishra et al. [154] showed that an autoregulatory network of phosphatases is activated in epidermal cells in response to an extrinsic differentiation signal. This network acts as a switch between two stable cell states, the stem and differentiated states. The simultaneous knockdown of a series of phosphatases (including DUSP6) lead to an upregulation of ERK activity, an increase clonal growth and maintainance of the stem cell commitment. This suggests that DUSP6 contributes to cell commitment. On the other hand, DUSP10 knockdown reduced clonal growth of keratonocytes, increasing commitment, suggesting that DUSP10 antagonises stem cell commitment. As argued earlier, these studied have the drawback of being coarse grained, where detailed features of ERK activity, such as pulses, are averaged out. Thus, the details of the interaction between ERK activity, phosphatases and cell commitment and differentiation remain until now only partially understood.

It has been recently observed in immortalized human mammary epithelial cells, that ERK activity fluctuates in a frequency-modulated manner [161], exhibiting regular pulses of short duration. Aoki et al. [164] found that in normal rat kidney epithelial cells, ERK activity pulses behave like stochastic events with an exponential distribution of waiting times, also referred to as *inter-pulse intervals*. They observed that pulses can occur spontaneously or propagate from cell to cell. Their results further support the observation that ERK activity is frequency-modulated[1] to regulate cellular processes, which in this case corresponds to cell proliferation. As mentioned earlier, Hiratsuka et al. [165] found bursts, waves and pulses of ERK activity both in healthy and wounded mouse skin tissues, where they observed striking differences in the patterns of ERK activation between the two conditions.

It is clear now that at the population level ERK activity pulses play an important role in stem cell maintenance, proliferation, wound healing and development. However, the relation between specific single cell ERK activity patterns and cell fate is still unclear. In this chapter it is shown that mammalian (human and mouse) skin stem cell fate is mediated by ERK pulses. Commitment to differentiation can be triggered by changes in ERK mean and pulse levels. Furthermore, it is show that these two features of ERK activity, mean and pulses, are independently regulated by phosphatases DUSP6 and DUSP10, respectively. These seem to act synergistically to induce and maintain cell commitment and differentiation.

This chapter is organised as follows: in Sec. 5.3 the experimental setup is described. The results are presented in Sec. 5.4, where the relation between ERK activity patterns and differentiation is studied experimentally. The discussion and outlook are left for Sec. 5.5.

5.3 In-vitro experiments

Live imaging of individual primary human neonatal keratinocytes (HNKs) was performed. ERK activity was monitored via lentiviral expression of a nuclear-tagged FRET[2] biosensor for ERK, EKAR-EVnls [175][3]. This biosensor allows the measurement of active ERK and inactive ERK, which are detected as FRET (YFP) or CFP signals, respectively. The ERK activity was then quantified as the ratio FRET/CFP of active over inactive signal. NHKs were transduced with the EKAR-EVnls reporter under feeder-free conditions and then plated on J2-3T3 feeder cells in order to monitor clonal growth and differentiation [176].

To asses the level of differentiation of cells a fluorescent reporter of Involucrin was developed for this study by Toru Hiratsuka. Involucrin is a gene that is upregulated in the suprabasal (more differentiated) epidermal

---

[1] As pulses are exponentially distributed in this case, it would be more accurate to call it a rate-modulation instead of a frequency-modulation.

[2] FRET refers to *Föster resonance energy transfer*. These biosensors have two fluorophores, a donor and an acceptor. By external excitation, usually by a source of light, the donor fluorophore can excited, which produced fluorescence of a given wavelength that can be measured. If the acceptor fluorophore is at a sufficiently small distance from the donor, there is electron transfer between them, allowing both second fluorophore to be excited, this is usually referred as FRET. Thus, allowing the measurement of a second wavelength different from the donors one.

[3] Here, the two fluorophores are YFP (yellow, Venus) for donor and CFP (Cyan, Turquoise) for receptor. When active ERK meets the biosensor EKAREV-NLS in the nucleus, the sensor is phosphorylated and FRET takes place.

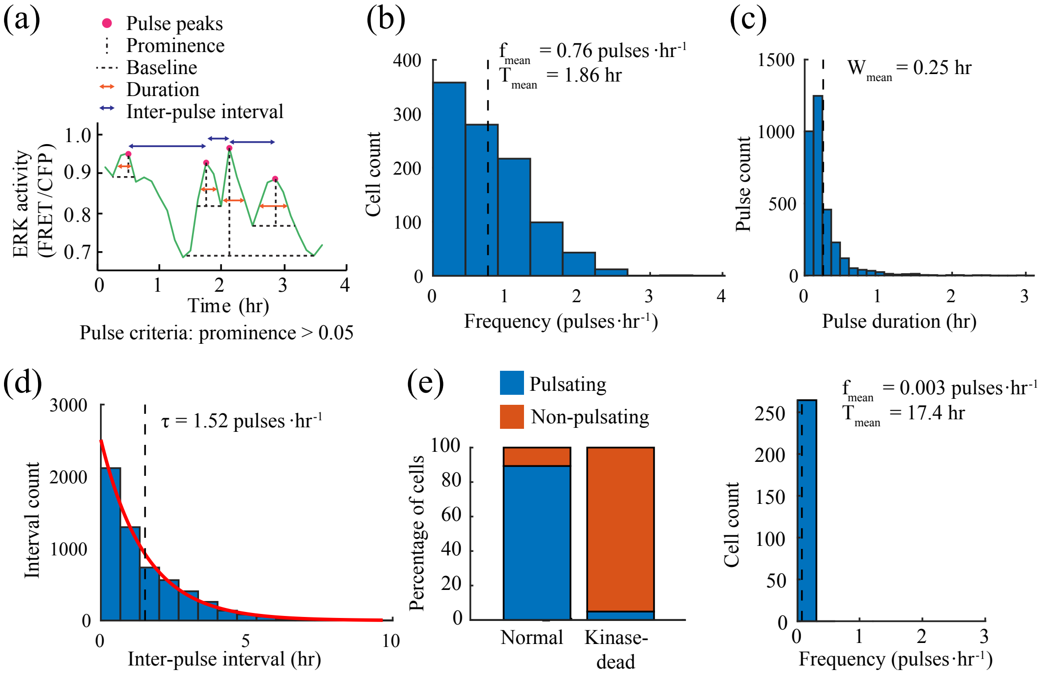


Figure 5.2: **Characterisation of ERK pulses.** (a) Schematic of the ERK pulse detection and quantification method. Pulses are detected as local peaks with prominence larger than 0.05 FRET/CFP value. Pulse duration was determined as the width of pulse at half the prominence of each pulse. The inter-pulse interval is characterised as the latency between consecutive pulses. (b) Shows the histogram of frequencies, indicating the mean frequency f_{mean} and period T_{mean}. (c) Histograms of pulse duration, indicating the mean value. (d) Shows the histogram for inter-pulse intervals. The red curve corresponds to and exponential fit, with decay rate τ. (e) Validation of the quantification methods with kinase-dead EKAR-EVnls biosensor (EKAREV-TA-nls), where FRET does not occur. Left, proportion of pulsatile cells in keratinocytes expressing normal EKAR-EVnls or EKAREV-TA-nls. Right, histogram of ERK pulse frequencies in pulsatile cells detected in keratinocytes expressing EKAREV-TA-nls. Data obtained from NHKs on feeder layers in complete FAD medium. This figure is part of the manuscript [8].

layers. A previously characterized Involucrin promoter and intron sequence [177] were used to drive mCherry (red) expression. The reporter was characterised and optimised, and showed a good response to differentiation stimuli.

5.4 Results

5.4.1 ERK activity and cell size in human epidermal cells

Preliminary measurement of ERK activity (FRET/CFP) in HNK colonies (Fig. 5.1a), revealed that the mean level of ERK activity was lowered in larger keratinocytes (Fig. 5.1b). Keratinocytes are known to enlarge as they undergo terminal differentiation [178], thus this result suggested that ERK activity is related to differentiation in this cell type. It is worth noting that there is a wide spread of mean ERK activity in smaller cells (Fig. 5.1b), which is significantly reduced as cells enlarge. This further suggests a connection between the precise regulation of ERK levels during differentiation.

Monitoring of the temporal behaviour of ERK activity in individual cells revealed important differences from cell to cell on their patterns of ERK activity (see Fig. 5.1c). Signals can be classified according to their mean level and fluctuations, later on it is argued that the level of fluctuations of the activity has a direct relation to its pulsatile behaviour. Some distinctive activity patterns can be seen in Fig. 5.1d, with some cells (1-4) exhibiting high mean-high fluctuations (high-high), other cells (*5) presenting a low mean-low fluctuation (stable-low) or high mean-low fluctuation (stable-high). All signals have quick, small amplitude (noise) fluctuation around the mean (Fig. 5.1d). However some signals present larger deviations of long duration, which are interpreted as pulses of activity. This is consistent with previous observations of ERK activity pulses in human epidermal stem cells [179].

ERK pulse detection

To quantify the pulsatile behaviour of ERK activity a peak detection method was used (MATLAB R2018b, `findpeaks` function). With this method peak-candidates correspond to local maxima in the activity signal (see Fig. 5.2a), this are then screened according to prominence[4], where peaks corresponded to maxima with a prominence larger than the threshold of 0.05 (see Fig. 5.2a)[5]. From the peaks detected, properties such as width (duration), time of occurrence, inter-pulse intervals and frequency, measured as the total number of pulses divided by the total length of the signal, were extracted. In normal NHKs the frequency of pulses ranged from almost zero to approximately 3.5 pulse/hr, with an average of 0.76 pulse/hr amongst the cell population (Fig. 5.2b). The average duration of pulses was 0.25 h, which is consistent with that previously reported in immortalized epithelial cells [161, 164] Fig. 5.2c. The histogram of waiting times between pulses or inter-pulse interval (Fig. 5.2d) was well fitted by an exponential decay curve ($\exp(-\tau t)$), with a rate $\tau = 1.52$ pulses/h. This supports the hypothesis that ERK pulses are stochastic rather than precisely timed events [179].

The possibility that pulses were an imaging artifact or that the peak detection was capturing noise fluctuations as pulses was ruled out by making use of a kinase-dead control, here phosphorilation of the biosensor (EKAREV-TA-nls) by ERK does not take place, thus FRET/CFP levels do not change. In this setup the pulse detection algorithm shows an approximately 90% reduction on cells with at least one pulse and a vast decrease in pulse frequency (Figs. 5.2e), confirming the effectiveness of the method to detect activity pulses.

5.4.2 Co-detection of ERK activity and differentiation

A comprehensive understanding of the relation between ERK activity patterns and differentiation in human keratinocytes, as suggested by the preliminary observations (Fig. 5.1), is obtained by studying the ERK activity and Involucrin, simultaneously. Involucrin expression in human epidermal stem cells can be observed in red in Fig. 5.3a, were differentiation was induced by changing medium from KSFM (low Ca^{2+} serum-free medium) [180] to high Ca^{2+} media (1.6 mM) or serum.

ERK pulses appeared to be downregulated at a time point which coincident with the onset of Involucrin ex-

---

[4] The prominence is the distance from the tip of the peak to the base line. The base line was determined by extending the largest horizontal line, below the corresponding peak, that did not cross the signal nor contains a peak higher than the corresponding one, as shown in Fig. 5.2a.

[5] Different values of the threshold where tested, this value was selected as is maximised the detection of what were interpreted as true peak, by eye inspection.

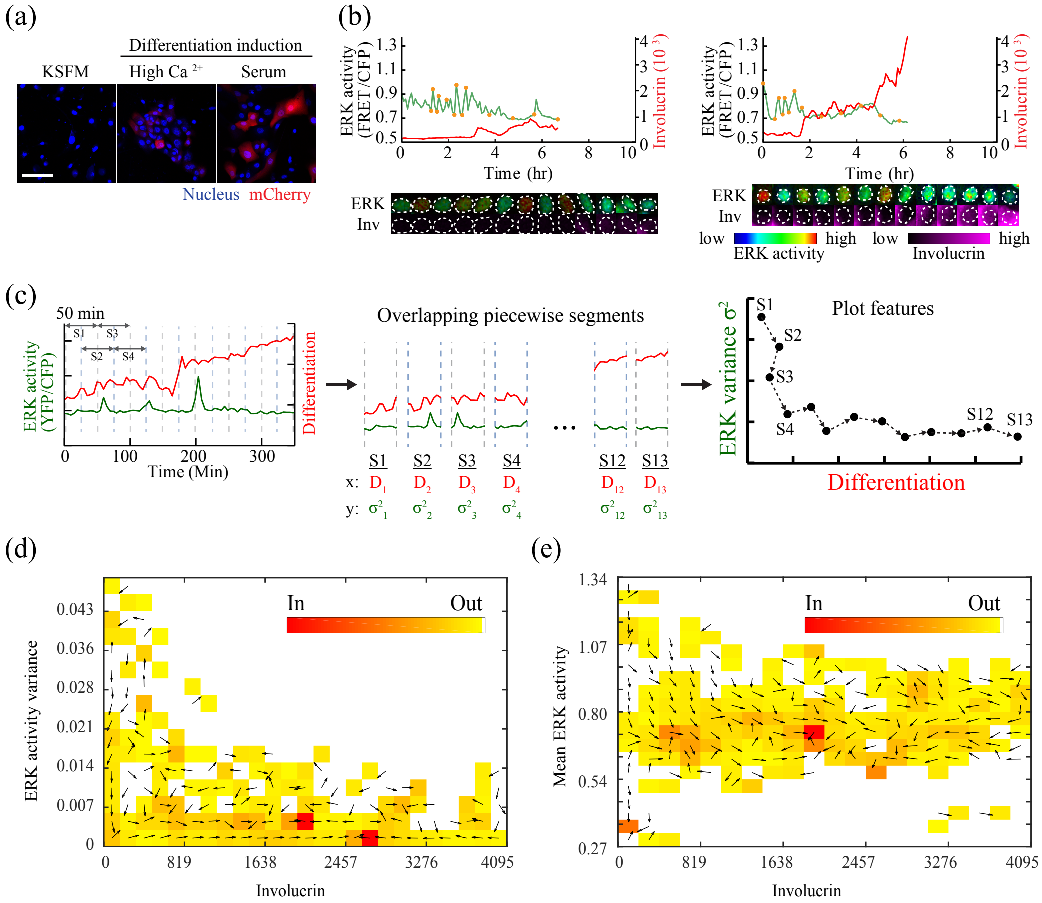


Figure 5.3: **Codetection of ERK activity and Involucrin level.** (a) Involucrin-mCherry reporter expression in cells cultured under the indicated conditions. Scale bar 100 μm. (b) Representative time-series of ERK activity (green) and Involucrin-mCherry expression (red), and images of the corresponding cells at different time points indicated by the orange circles in each time-series. (e) Schematics of the methodology for constructing the phase diagram of ERK activity. ERK moving variance, and ERK and Involucrin moving mean levels are measured for the time series of each cell. By Plotting the $(D_i, \sigma_i^2 var)$-map, a trajectory of the co-evolution of the two factors is obtained. The phase diagram of ERK activity variance (d) and mean ERK activity (e) as a function of the mean Involucrin level is shown ($n = 3397$ cells). Arrows indicate the average direction of transition between compartments. Red colour indicates points of accumulation. This figure is part of the manuscript [8].

pression, example of which can be seen in Fig. 5.3b. Cells that maintained stable-high or stable-low ERK activity profiles, appeared to maintained either a low or high Involucrin expression, respectively.

In the following a method to obtain an integrative view of the temporal dynamics of ERK activity and Involucrin expression, is described.

5.4.3 Construction of the phase diagram of NHKs

Moving variance as a measure of ERK pulse level in a time window

In order to study the change in ERK activity pulses over time in individual cells, overlapping moving time windows of 50 minutes were analysed (see left panel in Fig. 5.3c)[6]. Each time window was small enough for the mean ERK activity, within the window, to be considered fixed, but long enough to accommodate an ERK activity pulse (typical pulse 0.25 h). For each window the variance or mean of ERK activity was computed (see centre panel in Fig. 5.3c). The variance is a measure of dispersion of the measurements in the window, and quantifies the extent of the deviations of the signal from its mean value. These deviations can occur due to pulses, or noise fluctuations (which are short lived and of much smaller amplitude than pulses). The variance captures both the amplitude and number of pulses in the time window, giving a quantitative measure of the pulsing level of ERK activity in the time window. The minimum value for the variance is zero, which corresponds to a signal without pulses or fluctuations; larger values of the variance indicate a higher level of pulsation. This method allows a continuous assessment of the pulse level over time. Other methods, such as peak detection and pulse count, amount to a discrete measurement of pulses, which does not lend itself naturally to a continuous quantitative analysis of the temporal evolution of ERK activity pulses.

Phase diagram

The values of variance of ERK activity and mean Involucrin level, obtained from the procedure described in Sec. 5.4.3, were plotted for every time window, providing a trajectory in the plane spanned by ERK activity variance and mean Involucrin level (see right panel in Fig. 5.3c). This was repeated for each of the 3238 cells considered. The ERK activity variance vs Involucrin mean level plane containing the 3238 overlapping trajectories was then divided into $n \times m$ regular blocks. The trajectories that lay within each block were then averaged to obtain a mean direction for each block. This procedure resulted in the phase diagram of ERK activity variance vs mean Involucrin level (see Fig. 5.3d). Every block in the phase diagram corresponds to a pair of ERK activity variance-Involucrin mean level values, while the arrow in the block indicates the mean direction to which these values changed in time. The same procedure can be followed to construct other phase diagrams, such as mean ERK activity vs mean Involucrin level (see Fig. 5.3e).

Phase diagram normalization and transition probabilities

Arrows in each (i, j) block of the phase diagram was decomposed into its x_i and y_j components, with $i \in [1, n]$ and $j \in [1, m]$. The x_i components were rescaled by the maximum value of the Involucrin mean level, i.e. $x'_i = x_i/\max(x)$, while the y components were rescaled by the maximum value of the ERK activity variance, i.e. $y'_i = y_i/\max(y)$. This rescaling amounted to normalising both axes of the phase diagram to the range [0, 1], and allowed the comparison between the x'_i and y'_i components of the arrows. The transition probabilities between neighbouring blocks corresponded to $rx_i = |x'_i|/(|x'_i| + |y'_i|)$ and $ry_j = |y'_j|/(|x'_j| + |y'_j|)$, where $|\cdot|$ indicates the absolute value. Here,

---

[6] Only time series of more than 90 minutes were analysed.

rx_j accounts for the probability of transitioning to the neighbouring block of Involucrin mean level, while, ry_j corresponds to the probability of transitioning to the neighbouring block of ERK activity variance. The signs of x'_i and y'_j indicate the direction of the transition, a minus (plus) sign signifies a transition towards decreasing (increasing) values of Involucrin and ERK activity variance, respectively. The probabilities are normalised, such that $rx_j + ry_j = 1$ for every pair (i, j).

This procedure allows to turns an ensemble of independent time series for ERK and Involucrin into a single phase diagram that shows the co-evolution of either ERK activity mean or variance versus the mean Involucrin level.

Simulations of the phase space dynamics

The division of the phase space into blocks and the definition of transition probabilities between them (as obtained in Sec. 5.4.3) amounts to turning the continuous phase space of ERK activity and Involucrin into a discrete finite state Markov process. This allow the simulation of the evolution of hypothetical cell states though the phase space of ERK activity variance (or mean) and Involucrin mean level. The stochastic simulations were performed as follows:

1. An initial distribution of N cell states was considered. They were uniformly distributed on the phase space. Their position on the phase diagram corresponded to their initial values of ERK activity variance and Involucrin mean level. Thus, the evolution of each of the N cell state was interpreted as the evolution of the internal state of a given cell.
2. On every iteration of the algorithm one of the N cells was selected at random.
3. The selected cell was allowed to move to one of the four nearest neighbouring blocks according to the transition probabilities, rx_i and ry_j, computed from the experimental phase diagram.

These simulations allows the observation of the relation between ERK activity (variance or mean) and Involucrin mean level over time, visualise the paths cell states take, and detect accumulation regions that can be interpreted as stable cell states.

5.4.4 Observations from the phase diagram

The flow in phase diagrams (Fig. 5.3d) suggests that human epidermal stem cells downregulate ERK activity pulses on the onset of differentiation. In contrast, it is observed that the mean ERK activity gradually converged towards an intermediate value, as the level of Involucrin increased (Fig. 5.3e). This differential behavior suggests that ERK pulses, as measured by the moving variance, and mean levels are subject to distinct regulatory mechanisms, and that the downregulation in ERK pulsatile behaviour is more closely related to stem cell commitment to differentiation.

Simulations of the evolution of stem cell state on the phase diagrams were performed to study how hypothetical cells would flow through the ERK-Involucrin phase diagram, selected time frames of the evolution is shown in Fig. 5.4. The initial condition for the simulations considered 10 cells on each phase diagram compartment (see left panels in Fig. 5.4), this homogeneous distribution of cell states relate the realistic scenario where there is an heterogeneous population of cell sates. Different initial conditions do not affect the observations described here. Simulations on the phase diagram constructed from the whole population of cells (Fig. 5.4a) show that human epidermal stem cells downregulate ERK pulse levels prior to induction of Involucrin expression. This is evidenced by the decrease in variance for cells with low Involucrin level, which afterwards converge to high Involucrin "fixed point" (box in Fig. 5.4a). This behaviour is more evident when performing the analysis only on cells with an overall increase in Involucrin, as seen in Fig. 5.4b. Preliminary observations show that there might be cycles in the high ERK variance-low Involucrin and low ERK variance-low Involucrin, this could be populations that remain in the stem cell state, arguably by regulating their ERK activity. A different interpretation, is that the transition from stem to differentiated cells is a step-wise process, where the cell state transits between metastable state before reaching terminal differentiation.

5.4.5 Pulsing state, stemness maintenance and proliferative capacity

When considering the phase diagram contructed from the subset of cells with overall decrease in Involucrin, simulations showed that some cell transition from high to low Involucrin by increasing ERK variance (Fig. 5.4c). However, it is not clear that an increase in ERK variance is necessary for this transition to take place. Is is interesting to see that, for the phase diagram of cells with decreasing Involucrin, cells converge to a state of stable-high ERK (high mean, low variance), as seen in the last panels of Figs. 5.4c and 5.4d. This indicates that cells that remain or converge to the stem state are characterised by a stable-high ERK actvity. Moreover, the observations of decrease Involucrin suggests that pulsatile ERK activity does not irreversibly commit epidermal stem cells to undergo terminal differentiation, but may mediate the transition between cell states.

It is known that keratinocytes are proliferative, as opposed to non-proliferative, terminally differentiated skin cells. Measurement of ERK mean and variance in dividing and non-dividing cells (Fig. 5.5a) shows that mitotic cells had a slightly higher pulse frequency (114 cells, 1.39 pulse/h) and mean level compared to non-dividing cells (1124 cells, 1.27 pulse/h).

The cells were then grouped according to their pulse frequency to compare the degree of proliferation of the cell with ERK pulse frequency higher or lower than 1.5 pulse/h. Individual cells where followed for 48 h, the ERK frequency was measured only during the first 6 h of observation. The selected were revisited at 48 h to asses the cell number. Of the cells with low ERK pulse frequency (< 1.5 pulse/h) 72% did not proliferate, remained as single cells for the 48 h period (Fig. 5.4b), compared to 41% in cells with high ERK pulse frequency (> 1.5 pulse/h). Of the cells that divided, 71% did it once in the low frequency group compared to a 49% in the high frequency group, where 30% of initially singe cells increased to 4 cells at 48 h (Fig. 5.4b). Thus cells with an initially more pulsatile ERK profile are more likely to divide than cells with a low frequency or stable ERK profile.

Combined with the observations from the phase diagrams (Figs. 5.3 and 5.4) it is inferred that cells remain in undifferentiated state by maintaining a stable-high ERK activity (Figs. 5.4c and 5.4d), pulsatile ERK activity appears to be metastable state showing increase cell proliferation (Figs. 5.5a and 5.5b). This pulsatile state then decays during commitment, leading to differentiation (Fig. 5.4a), as pictorially described in Fig. 5.5c.

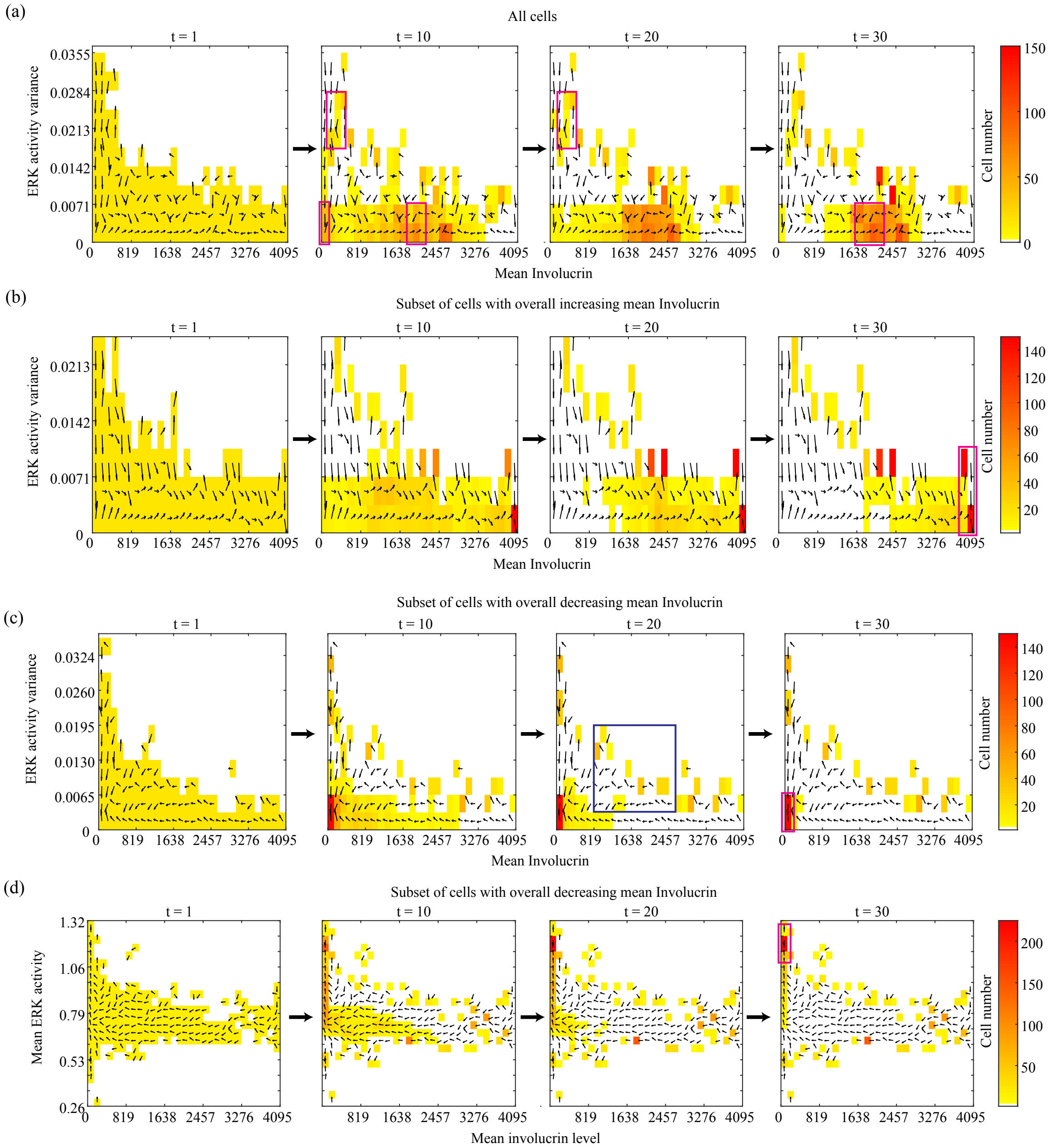


Figure 5.4: **Evolution of simulated cells.** Stochastic simulations of simulated cells in the phase diagram, showing the expected co-evolution of ERK activity variance and Involucrin mean level. The phase diagram was constructed by considering (a) the whole population of human keratinocytes (3397 cells), cultured on feeder layers in complete FAD medium, (b) the subpopulation with overall increase in mean Involucrin level (2175 cells), (c) and (d) the subpopulation with overall decrease in mean Involucrin level (1219 cells). The initial condition for the simulations corresponded to cells uniformly distributed on the phase diagram. As time passes (from left to right), the cells transition according to the rates (represented by arrows), converging to certain regions of the phase diagram (red). This figure is part of the manuscript [8].

In the following, the dose- and time-dependent interplay between differentiation and ERK activity pulses is studied by stimulating the cells with a variety of differentiation cues. The study of pulsatile behaviour of a pop-

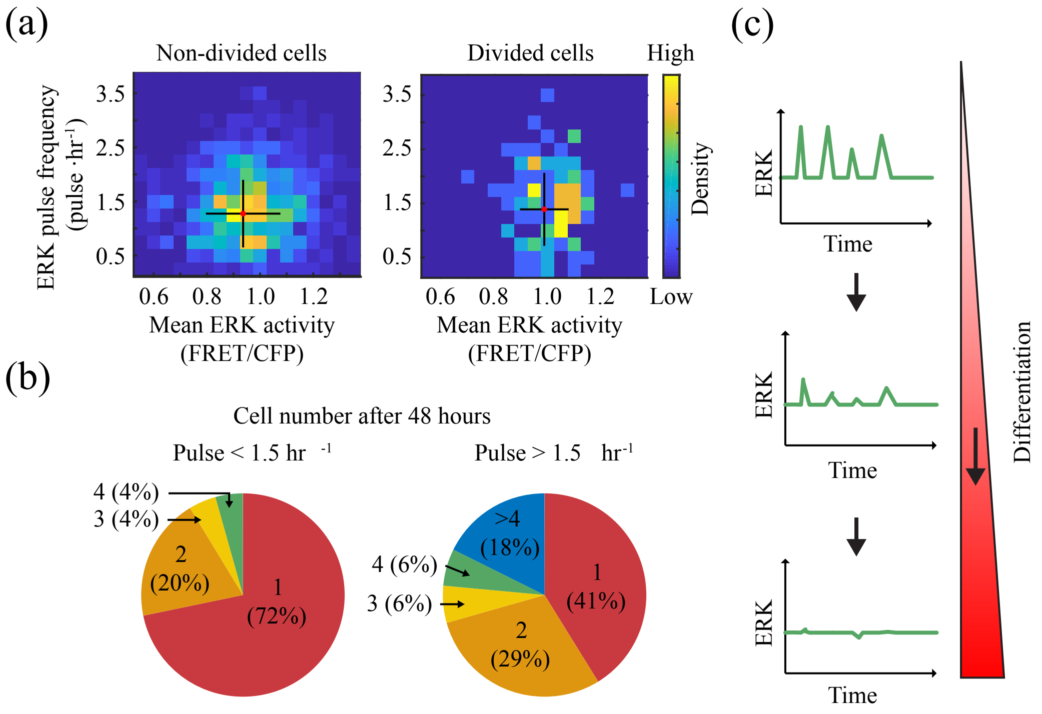


Figure 5.5: **ERK activity and cell proliferation.** (a) 2D histogram of ERK pulse frequency and mean activity in the total cell population (left) and the cell subpopulation that underwent cell division during live imaging (right). The cells expressing EKAR-EVnls were cultured for 7 days on feeder cells. Mean and SD are shown by red dot and black lines, respectively. (b) Cell proliferation assay of single cells with ERK pulses lower (left) or higher (right) than 1.5 pulse/h. Cells were initially imaged at single cell state to measure their pulse levels, and then the same cells were observed after 48 hours. The number of cells at 48 h and the proliferation fractions. (c) Schematic representation of modulations in ERK activation pulses during differentiation. This figure is part of the manuscript [8].

ulation of cells, at a given time point, is not possible through the pulse detection method described earlier, as the measurement of frequency considers the whole temporal signal.

5.4.6 Instantaneous variance as a measure of population-level ERK pulses

To quantify the level of ERK activity pulses of the population of cells at a given time point, the variance of ERK activity (FRET/CFP) was measured at that time point among all cells (instantaneous variance). An increase in the instantaneous variance indicates a higher variability of ERK activity in the population at a specific time point.

Pulses behave as stochastic events, as suggested by the exponential distribution of interpulse intervals (Figs. 5.2d and 5.6). Then, as long as the mean ERK activity does not change significantly between conditions the use of the instantaneous variance as a measure of the level of ERK activity pulses of a population at a specific time point is justified. In the following, when the instantaneous variance remains unchanged between conditions it is said that the level of ERK activity pulses are the same for both populations, regardless of changes in the mean level.

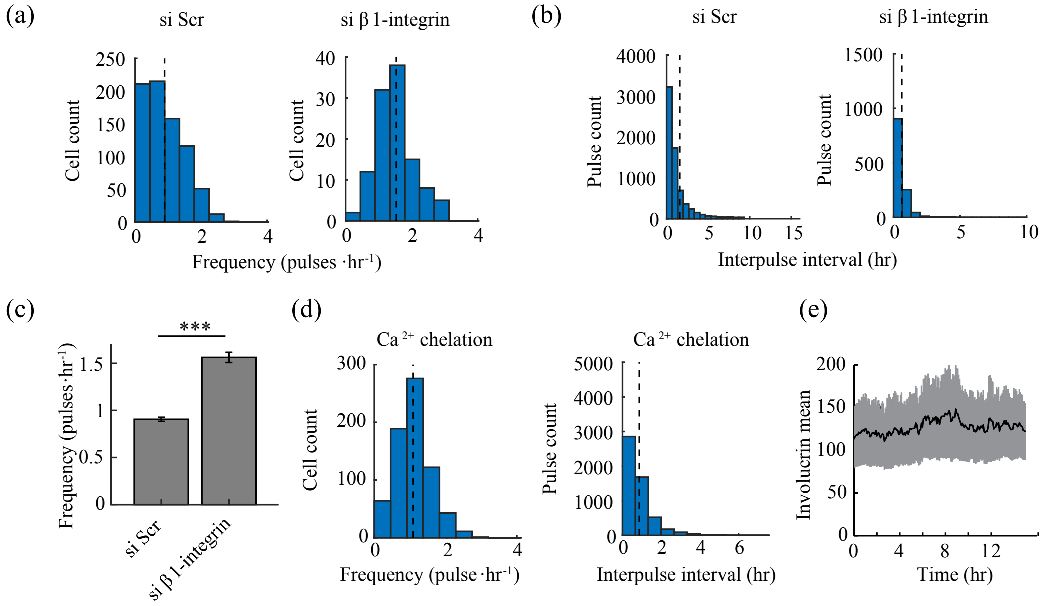


Figure 5.6: **Molecular regulation of ERK activity.** Histograms of (a) frequencies and (b) inter-pulse intervals in cells treated with scrambled control siRNA (left) or β1-integrin-targeted siRNA (right). Black dotted lines indicate the mean. (c) Frequency of ERK pulses for both treated cells, plot shows to mean and ± SEM. (d) Shows the histograms of (left) frequencies and (right) interpulse intervals of cells cultured with Ca^{2+}-chelated medium. Black dotted lines indicate the mean. (e) Mean Involucrin reporter expression as a function of time, indicating the mean ± SD. This figure is part of the manuscript [8].

5.4.7 ERK pulse modulation by terminal differentiation stimuli

Inhibition of cell-substrate interactions by reduction of integrin-mediated adhesion is known to trigger exit from the stem cell compartment via a reduction in ERK signaling [181]. To test the effect of integrins in ERK pulses Human keratinocytes were transduced with an siRNA targeting β1-integrin or a scrambled siRNA control. It was observed that a reduction in β1-integrin expression increased the frequency of ERK pulses (Fig. 5.5a) from 1.54 pulse/hr in si-β1 integrin treated cells compared to 0.89 pulse/hr in the scrambled siRNA treated cells (Fig. 5.6c), implying a reduction in the interpulse intervals (Fig. 5.6b). In contrast to the effect of inhibiting integrin-mediated adhesion, reducing the cell-cell junctions by Ca^{2+} depletion, thus blocking formation of adherens junctions and desmosomes, had little to no effect on ERK pulse frequency (see Figs. 5.6d). Consistent with this, Involucrin expression was not affected by the inhibition of intercellular adhesion (Fig. 5.6e) [182]. These results indicate that modulation of cell-substrate interactions plays a more significant role compared to (Ca^{2+}-mediated) cell-cell interactions.

Direct stimulation of Involucrin expression by phorbol ester (TPA) [183] is also known to increase overall ERK activity [184]. Keratinocytes had a dose-dependent increase in ERK pulse frequency increase after stimulation with TPA (Fig. 5.7a), while mean ERK levels peaked at 10 ng/ml. TPA transiently induced ERK pulses (as measured by the instantaneous variance), which decreased drastically after the peak at 9 h, to stabilise at about 15 h (Fig. 5.7b). The onset of Involucrin upregulation (Fig. 5.7c) coincided with this pulse downregulation. On the other hand,

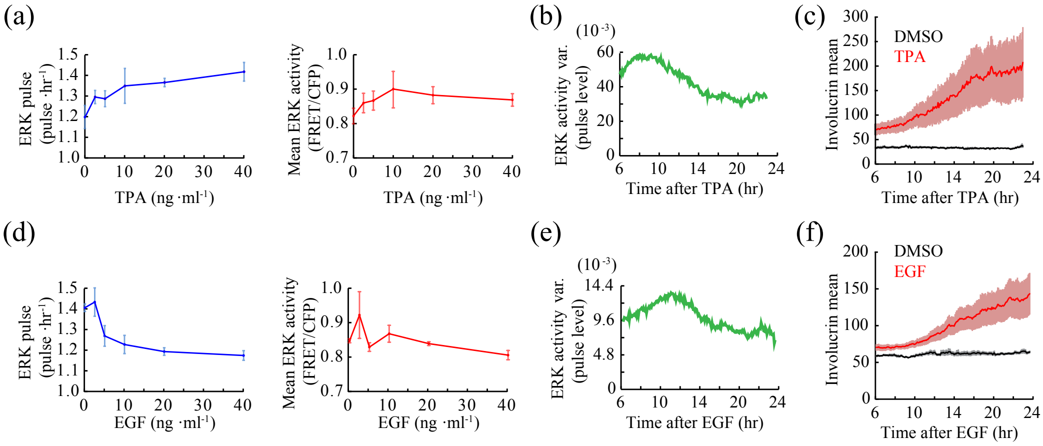


Figure 5.7: **Molecular regulation of Involucrin expression,** (a) and (d) shows the ERK pulse frequency (left) and mean (right) as a function of TPA and EGF dose, respectively. Data shows mean ± SD for about 1000 cells for each condition. (b) and (e) the ERK pulse frequency (left) and mean (right) as a function of time after treatment with TPA and EGF, repectively. Involucrin expression as a function of time after (c) TPA and DMSO control, and (f) MEK inhibitor, EGF and DMSO control, indicating the mean ± SD. This figure is part of the manuscript [8].

EGF[7] stimulation showed a dose-dependent decrease (Fig. 5.7d. Similarly to TPA, EGF transiently increased ERK pulse frequency (Fig. 5.7e), with a peak at 11 h, followed by a sharp decrease in pulse frequency. Again, the onset of Involucrin upregulation (Fig. 5.7f) coincided with the decrease in ERK pulse frequency. The response of keratinocytes to TPA or EGF treatment is a clear indicator of the tight relation between differentiation and ERK pulse regulation.

All three differentiation stimuli considered here—reduced integrin-mediated adhesion, TPA and EGF—triggered an increase ERK pulses prior to their ERK, where the transition coincides with the onset of differentiation, as measured by Involucrin upregulation. It should be noted that it appears like ERK activity pulses do not lead to immediate differentiation, as in the case where cells show a pulsatile ERF profile yet proliferate.

5.4.8 Regulation of ERK activity by protein phosphatases

One likely mechanism for the control of mean ERK activity and pulses frequency is via negative feedback regulation by protein phosphatases [185]. Here the effect of two key members of the protein phosphatase network, DUSP6 and DUSP10, is studied. These phosphatases act as a commitment switch in human epidermal stem cells [154]. To investigate their effects on ERK activity, DUSP overexpression was induced by doxycycline. Analysis though live imaging showed that DUSP6 reduced ERK pulses without changing mean ERK activity (Fig. 5.8a), while DUSP10 downregulated mean ERK levels without changing ERK pulses (Fig. 5.8b). This indicates that mean ERK levels and ERK pulses are independently regulated by different members of the phosphatase network.

DUSP6 overexpression had no effect on Involucrin expression over time compared to the control (Fig. 5.8c). On the other hand, DUSP10 overexpression strongly stimulated Involucrin expression (Fig. 5.8d). Moreover, the histogram of Involucrin expression of cell affected by DUSP6 and DUSP10 show that DUSP6 induction increases

---

[7] Related to cell proliferation [162].

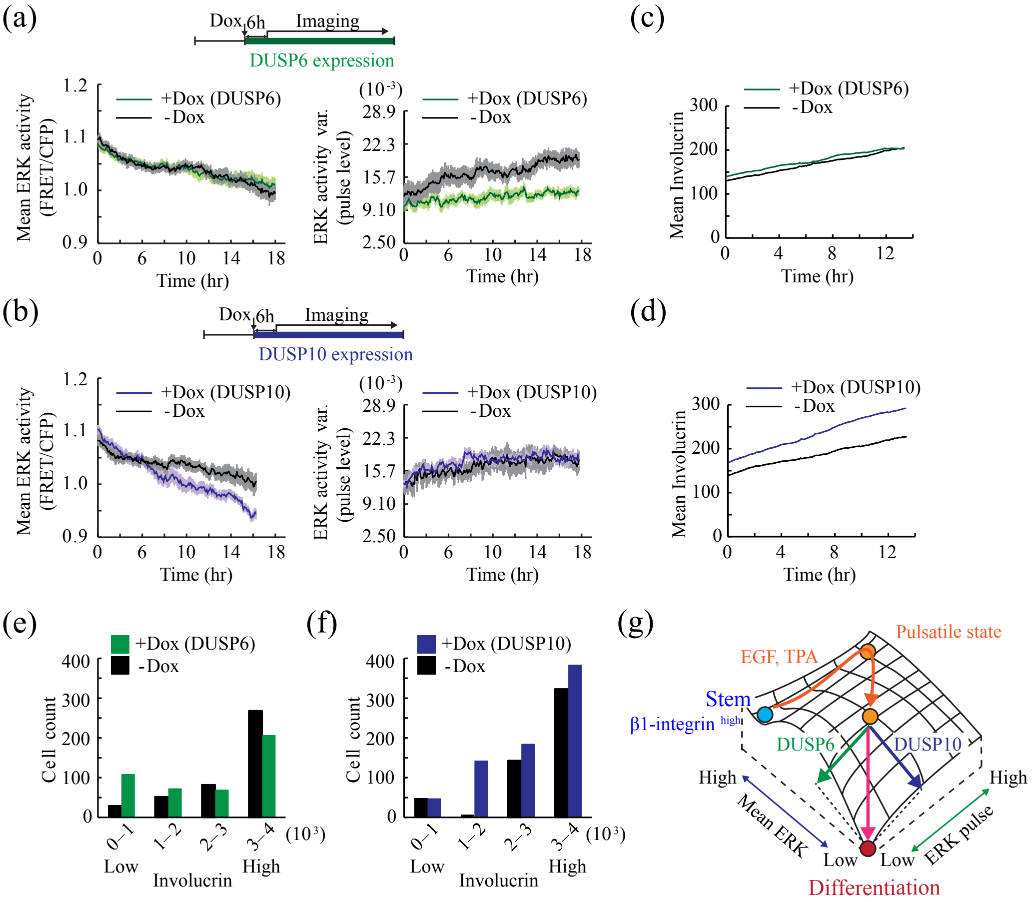


Figure 5.8: **Regulation of mean and variance of ERK activity by DUSP.** (a) and (b) show (left) mean and (right) variance of ERK activity in NHKs treated with 1 μg/ml doxycyclin (green and purple) or vehicle (black). The induction of (a) DUSP 6 and (b) DUSP10 was done by the doxycycline treatment. Data correspond to the mean ± SEM (1220 doxycyclin-treated cells and 1261 vehicle-treated cells for (a), 1224 doxycyclin-treated cells and 1005 vehicle-treated cells for (b)). (c) and (d) show Involucrin expression as a function of time of keratinocytes treated with 1 μg/ml doxycyclin (green or purple) or vehicle (black). DUSP 6 (c) or DUSP10 (d) was induced by the doxycycline treatment. Histograms for involucring expression 18.5 hour after doxycycline-induced (e) DUSP6 and (f) DUSP10 expression (green and purple, respectively) or vehicle (black) treatment. (g) Shows a schematic representation of molecular regulation of ERK mean and pulse level. This figure is part of the manuscript [8].

the proportion of cells with low Involucrin expression (Fig. 5.8e). In contrast, DUSP10 expression increased the proportion of cells with high Involucrin expression (Fig. 5.8f), compared to the control. One possibility is that DUSP6-mediated ERK pulse downregulation promotes the initiation of differentiation, whereas DUSP10-mediated downregulation of mean ERK activity promotes and stabilizes post-commitment differentiation. This is consistent with the finding that DUSP6 is transiently upregulated on commitment, while DUSP10 upregulation during suspension-induced differentiation is more sustained [154]. Altogether, these observations suggest that DUSP6 and DUSP10 regulate ERK pulse levels and ERK mean levels, respectively, and act synergistically to induce differentiation.

5.5 Discussion and outlook

Stem cells are the building block of high level organisms. The balance between the number of stem cells and differentiated cells is key for the long term maintenance of healthy tissues. Keratinocytes play this role in skin, thus understanding the mechanisms for their transition between stem cell and differentiated states is key for understanding developmental processes and adult tissue maintenance, and could contribute to the development of new treatments for growth related diseases.

In this chapter is was shown that the ERK pathway might provide a reliable mechanism for the control of cell commitment. Previous studies have shown that ERK plays a key role in exit of embryonic stem cells from the so-called stem cell compartment and in lineage specification [153, 186]. The role of ERK signaling in cell fate regulation has been extensively studied in multiple cell types [162]. One key feature of ERK is its pulsatile behaviour, which have been reported in multiple cell types in relation to cell proliferation and tissue morphogenesis [161, 164, 167]. However, the role of ERK pulses in stem cell commitment and differentiation was to large extent unknown. With the use of ideas from statistical physics and dynamical systems it was possible to construct a phase diagram for the Keratinocyte state, which allowed the visualisation of the co-evolution of ERK activity and cell differentiation. These observations served as motivation to the more in-depth study of the molecular mechanisms behind the control of ERK activity and its relation to cell commitment and fate determination.

The results presented in this chapter show that epidermal cells present a large variety of temporal patterns of ERK activity, which can be classified according to their mean and pulse levels. Consistent with previous observations [179], the waiting time between pulses was exponentially distributed in all conditions, suggesting that the underlying mechanism for ERK pulses triggering is a Poisson process. Pulses were consistently downregulated on commitment to differentiation, whether spontaneous or induced by external stimuli. ERK pulse levels were regulated independently of mean ERK levels, and thus cells had variable temporal ERK activity patterns, such as stable-high, pulsatile, and stable-low profiles. These results demonstrates that ERK pulses play significant roles in stem cell fate regulation and raises the possibility that pulse-mediated changes in cell fate are conserved in multiple tissues and organisms. A schematics of the Keratinocytes state landscape, as deduced from the results presented in this chapter, is shown in Fig. 5.8g, where the cell state is regulated not only by DUSP but also by integrins, EGF and TPA. The independent regulation of ERK pulse and mean levels by DUSP6 and DUSP10 indicates that DUSP6 and DUSP10 operate as independent feedback loops to achieve different ERK activity profiles and different cellular outcomes. Further investigation will be required to understand how ERK activity pulses are initiated and what downstream targets they induce. Fluctuations of signaling pathways are increasingly recognized as key determinants for tissue development [158]. Multifaceted features of those fluctuations, such as phase, frequency, and amplitude, provide potentially different outputs in terms of cell fate. It will now be of significant interest to explore the generality of ERK pulses in different cells and tissues.

Future work will focus on the theoretical understanding of the mechanisms by which ERK interacts with Involucrin. Some progress has been made in the mathematical modeling or ERK activity [164]. However, the connection between ERK (up-stream) and Involucrin (down-stream), from a dynamical systems perspective, remain unkown.

Chapter 6

Conclusions

In the past century Biology has gone from being uncharted territory to probably the most popular and researched branch of science. Even though sciences are still classified into the three classic branches—physics, chemistry and biology—the notion of independence between branches is not clear anymore, especially in biology. Nowadays biological research involves, in most of the cases, the collaboration between scientists from the other two branches. Chemists develop reagents, and purification techniques for biologists, while physicists and mathematicians bring to the table methods from statistics, inference, modelling and theoretical approaches, which provide descriptive and predictive power.

Initial theoretical approaches were mostly speculative, mainly due to the lack of detailed experimental data, which was the case for example of Turing and his ideas of morphogenesis [71]. In many cases, biological systems were used more as a motivation rather than as the subject of study. Thus, limiting the real predictive power of the models considered. In the past decades new experimental techniques and visualization tools, together with greater computational power and storage capabilities have allowed the observation and recording of ecological and biological processes at high spatial and temporal resolution. These measurements range from the observation of in-vitro microorganisms and cellular tissues to in-vivo recording of embryogenesis and lineage tracing. These observations have been complemented by immunofluorescence techniques and genetic manipulation, which allow the study of perturbations on the biochemical network of the cell and to observe the dynamics of molecular processes occurring inside the cell. This allows the study of specific molecular pathways during the different stages of the cell and tissues.

Starting from a mainly theoretical question in Ch. 2 and moving into a purely experimental work in Ch. 5 this thesis has, hopefully, convinced the reader that physical and mathematical thinking can both contribute to and benefit from attempting to solve fundamental and applied questions in biology and far-from-equilibrium systems. In some cases like Ch. 2 the notion of biology and viral spreading is used as a motivation for the study of the volume explored by a branching random walk. In others, like Ch. 4 is the biological setup that motivates the study of effective interaction between embryonic stem cells and protein signals. These approaches produce results with different levels of abstraction, one contributing to the theoretical understanding of a general class of branching processes, while the other providing insight into the biological nature of cellular processes.

Using tools from non-equilibrium statistical mechanics it was shown in Ch. 2 that relevant observables of the propagation of spatial branching processes can be obtained on regular lattices and networks. This problem lies on the most theoretical side of the spectrum, however due to the relation of the BRW with the propagation of viral processes in animal populations, tissues, and communication chemical networks, these results are of interest for a broad range of scientists. Overall, the results presented here are an important step forward in the understanding of viral propagation in general systems.

Spatial branching processes are ubiquitous in nature. In Ch. 3, the properties of a specific type of spatial branching is studied in the propagation of vegetation in real ecosystems. This process was studied from a theoretical and numerical perspective to describe real observations. A nonlinear model for population dynamics was introduced, which incorporated a nonlocal Kernel for the competitive interactions between plants. It was shown in Ch. 3 that the facilitative and competitive interactions in vegetation give rise to self-replication, whereby localised patches—particle-like solutions of the dynamical system—split to create two identical patches. This process, resembling cellular mitosis, provides a spreading mechanism for vegetation in scarce conditions. Differently from the BRW described in Ch. 2, self replication in this model for vegetation is supercritical, and volume exclusion between vegetation patches contribute to the patterning of landscapes. Satellite images were used in an attempt to close the gap between the theoretical predictions and real observations. This analysis suggests that self replication might be indeed taking place in locations where vegetation patches dominate the landscape, however further on-site investigation are necessary. It is worth noting that even though the model presented in Ch. 3 was motivated by the interactions observed in vegetation, there is no reason to think that the same class of models could not apply for the modeling, for example, of neuron interaction dynamics or cellular tissues and cancer propagation, where nonlocal interactions might also play a key role.

Spatial interactions play an important role in all physical scales. In Ch. 4, the focus shifts from the macro-scale interactions studied in Ch. 3 (on the order of kilometres) to the mesoscopic spatio-temporal interactions taking place at the single cell level (on the order of microns). Cellular processes also occur at much faster time scales, while the life cycle of a tree is in the scale of decades, the life cycle of a cell is in the order of days. In Ch. 4 experiments with embryonic stem cells were combined with theoretical and numerical methods to understand the spatio-temporal interactions of cells and localized sources of protein signals. Inspired by theory of simple liquids, it was shown that cell-signal interactions may be interpreted as effective potentials that emerge at the population level. These effective potentials allow the connection of small and large scales, while also shedding light on the temporal evolution of the system of single cell-signal interactions. Furthermore, the work presented in Ch. 4 shows how simple theoretical methods can be used to understand and predict the non-equilibrium evolution of cellular systems.

In Ch. 5 the focus shift to even smaller length and faster time scales. There, the objective was to connect the temporal behaviour of two specific molecular pathways in skin stem cells and establish a causal relation between them. This was achieved by incorporating single cell and population level measurement and data analysis methods inspired from statistical mechanics. Through the construction of a phase diagram for the cell state a connection was established between the pulsatile behaviour of the ERK pathway activity and the expression of Involucrin, a differentiation marker. The flow in this phase diagram amounted to the temporal co-evolution of both quantities, observations that then motivated new experiments to understand their molecular origins. This work is the perfect example of a synergistic collaboration between physicists and biologists, where the combination of approaches contributed to the development of new experiments and tools to explore fundamental problems in biology. Now that there is a fair understanding of the biological and biochemical interactions of ERK and Involucrin, future work shall focus on modeling and theoretical understanding of the underlying dynamical system.

Lying in the interface between biology, physics and applied mathematics, and between theory and experiments, this thesis describes novel approaches to calculate, model and quantify the spatial and temporal dynamics of active and living matter. These approaches can contribute to find the answers to relevant questions in both mathematics an biology. By making use of techniques and ideas from statistical mechanics the following objectives of this thesis have been fulfilled:

- With the use of tools from equilibrium and non-equilibrium statistical mechanics, phenomena relevant to biological and ecological systems were described and characterised, Chapters 2-5.
- The current knowledge of equilibrium physics was extended to active systems, such as self-replicating (branching) matter and self-organised spatial and temporal processes, Chapters 4 and 5.
- New methodologies were developed for extracting relevant physical and biological observables from experimental data, Chapters 4 and 5.

Findings by chapter

Each chapter of this thesis contributed to the development of a particular field, while also opening avenues for future research. A list of findings and open problems for each chapter are presented below.

Chapter 2

Findings

- In this chapter, the first analytic results, about the long time, large system size behaviour of a branching random walk in arbitrary dimensions were presented, these results were supported by numerical simulations.
- A direct relationship between the dimensionality of the substrate and the rate of propagation of the viral process was found.
- The results obtained for regular lattices were extended to general graphs and applied to the spreading dynamics in simulated and real complex networks.
- These results provide a theoretical framework to model and characterise real-world spreading processes, in particular disease and information propagation.
- All together, these results presents a major step forward in the theoretical and practical understanding of spatial branching phenomena in general environments.

Future work

- To study the limitations of the extension of the results from regular lattices to general graphs. When does it break down and why?
- To study how is the scaling in time and system size affected when considering variations of the BRW, for instance by the inclusion of volume exclusion, self-avoiding walkers or vicious random walks.
- To study further the formulation of the field theory of the BRW, and how different dimensional considerations can lead to other different field theories.

Chapter 3

Findings

- In this chapter a non-local model for vegetation dynamics and its stability analysis was presented.
- Self replication of vegetation patches was found numerically, where two routes were discussed, a mitotic-like division and a ring-like instability.
- The phenomena described analytically and numerically where compared against real ecosystems images, which provided evidence supporting the existence of self-organisation and patch replication.

Future work

- Some models of vegetation incorporate seasonality though periodic fluctuations of aridity, it would be interesting to study how this might affect the self-replication mechanism.
- Related to the previous point is the study of short extreme droughts of floods in the vegetation dynamics. This type of extreme events are now more likely with climate change.
- It would be interesting to apply the class of interaction-redistribution model described here on other systems, such as cellular tissues, where it is know that molecular promoters (facilitation) and repressors (competition) can act at long distances, contributing to the spatial organisation of tissues.

Chapter 4

Findings

- A suitable age and time dependent correlation function was introduced to study stem cell-protein signal interaction.
- This work demonstrates experimentally the emergence of effective potentials for the interaction between single stem cells and protein signals.
- A simple theoretical model combined with numerical simulations allow the distinction between different interaction mechanisms.

- Cells are found to be selective, showing a strong interaction with biologically relevant signals. This has an important effect on the spatial organisation of the system.
- The methods presented in this chapter provide a route to connect the microscopic cell-signal interactions to the mesoscopic organisation of the environment.

Future work

- It would be interesting to carry out measurements of effective potentials in systems with migratory phenotype to see the effect on the effective potentials.
- Preliminary observations show that the proportion of cells in contact with one, two, three... beads as a function of time can be well described by a (non spatial) Markov process. It would be interesting to study why and when this "mean field" approach fails.

Chapter 5

Findings

- The results in this chapter show, for the first time, a connection between the pulsatile behaviour of a molecular pathway (ERK) is directly connected to stem cell commitment and differentiation (Involucrin).
- A method for the construction and visualization of the cell states was developed. The cell state phase diagram allows the observation of the temporal relation between ERK and Involucrin.

Future work

- Future work will focus on building a theoretical model for the interactions between ERK pulses and Involucrin. This is a major challenge as Involucrin is far downstream compared to ERK activation.
- Once a model for the temporal relation between ERK and Involucrin is achieved it will be interesting to look at the spatial propagation of ERK pulses, and their relation to proliferation and differentiation, simultaneously.
- Future work will also focus on tracing ERK activity and Involucrin of individual cells during several cycles of cell divisions.

Appendix A

Field-theory of reaction-diffusion processes

In the following, a field-theoretic formalism to study non-equilibrium reaction-diffusion systems is introduced. This formalism, proposed initially by M. Doi [23] and extended later by L. Peliti [24], provides a route to cast reaction-diffusion systems in a field-theoretic language through the use of creation and annihilation operators for the discrete particles in the system, and the subsequent formulation in path integral form and derivation of a field theoretic action. Once the action has been obtained, one can make use of the well known machinery of perturbative renormalization in order to compute observables of interest.

A.1 The master equation

The first step is to construct the master equation for the (Markovian) reaction-diffusion process of interest on a d-dimensional lattice. Reactions between discrete particles occur locally, while diffusion consider hopping of particle between nearest-neighbour sites. For two species A and B, the *instantaneous* state of the system can be characterised by the sets $\{n_\mathbf{x}\}$ and $\{m_\mathbf{x}\}$ of occupation numbers $n_\mathbf{x}$ and $m_\mathbf{x}$, respectively, for each site $\mathbf{x}$ of the lattice given. Here two particle species have been considered, however the extension to more species is straightforward as every additional particle species is accompanied by an additional set $\{\cdot\}$ of occupation numbers.

Given the independence of these sets of occupation numbers, a master equation can be written for the joint probability as $\mathcal{P}(\{n_\mathbf{x}\}, \{m_\mathbf{x}\}; t)$ to find the particles of species A and B in a certain configuration of occupation numbers $\{n_\mathbf{x}\}$ and $\{m_\mathbf{x}\}$ at time t, respectively. The master equation reads

$$\frac{d}{dt}\mathcal{P} = \mathcal{M}\mathcal{P} \tag{A.1}$$

where $\mathcal{P}$ is the vector of states of the system, and the transition matrix $\mathcal{M}$ contains the transition rates between the states.

The temporal evolution of a reaction-diffusion process as described by the master equation (A.1) can be characterized through the use of a Fock state representation [23]. This representation, borrowed from the second-quantisation formalism for quantum many-body systems, requires the introduction of Fock or number states $|n, m\rangle$, which allows us to write the process in term of ladder operators. The mixed state for a two species system is defined as

$$|\Phi(t)\rangle = \sum_{\{n_\mathbf{x}\}} \sum_{\{m_\mathbf{x}\}} \mathcal{P}(\{n_\mathbf{x}\}, \{m_\mathbf{x}\}; t) a_1^{\dagger n}(\mathbf{x}) a_2^{\dagger m}(\mathbf{x}) |0, 0\rangle \,, \tag{A.2}$$

where $|0, 0\rangle$ corresponds to the vacuum state. The ladder operators $a_1^\dagger(\mathbf{x})$ and $a_2^\dagger(\mathbf{x})$ define the creators, and $a_1(\mathbf{x})$ and $a_2(\mathbf{x})$ the annihilators for species A and B, respectively, and satisfy

$$a_1^\dagger(\mathbf{x}) |n_\mathbf{x}\rangle = |n_\mathbf{x} + 1\rangle \tag{A.3}$$

$$a_1(\mathbf{x}) |n_\mathbf{x}\rangle = n_\mathbf{x} |n_\mathbf{x} - 1\rangle \tag{A.4}$$

and

$$\langle n_\mathbf{x}| a_1^\dagger(\mathbf{x}) = \langle n_\mathbf{x} - 1| \tag{A.5}$$

$$\langle n_\mathbf{x}| a_1(\mathbf{x}) = (n_\mathbf{x} + 1) \langle n_\mathbf{x} + 1| \,, \tag{A.6}$$

and correspondingly for a_2 and $m_\mathbf{x}$, together with the commutation relations

$$[a_i(\mathbf{x}), a_j^\dagger(\mathbf{x}')] = \delta_{i,j} \delta(\mathbf{x} - \mathbf{x}') \tag{A.7}$$

$$[a_i(\mathbf{x}), b_j(\mathbf{x})] = [a_i^\dagger(\mathbf{x}), a_j^\dagger(\mathbf{x})] = 0. \tag{A.8}$$

It is important to note that the creation and annihilation operators act on specific lattice site, denoted by $\mathbf{x}$. Thus, the action of *hitting* a lattice site with a creation (annihilation) operator is to increment (reduce) the number of particles of the corresponding type on the lattice site considered. As it is shown latter, diffusion of particles can be effectively interpreted as the annihilation of a particle a a given site followed by the creation of a *new*, yet indistinguishable from the original, particle of the same species in one of the neighbouring sites.

By writing Eq. (A.1) in terms of the number states (Eq. (A.2)) one obtains an *imaginary time Schrödinger equation*

$$\frac{d}{dt} |\Phi(t)\rangle = \mathcal{A} |\Phi(t)\rangle \,, \tag{A.9}$$

where $\mathcal{A} = \mathcal{A}(a_i, a_i^\dagger)$ corresponds to the action of the process written in terms of creation and annihilation operators in normal order, i.e. creation operators on the left of annihilation operators. There are at least two key difference between the interpretation of the classical Eq. (A.9) and the quantum mechanical Schrödinger equation. First, in general the dynamics of the reaction-diffusion systems is out-of-equilibrium, where detailed balance is broken, thus the action $\mathcal{A}$ in Eq. (A.9) need not be Hermitian. Secondly, in the quantum mechanic interpretation of Eq. (A.9) the state of a system is function of the probability amplitudes, in the classical scenario of reaction-diffusion systems considered here, the state of the system is a linear function of the probabilities [187].

A.2 Path integral formulation

As proposed by L. Peliti [24], Eq. (A.9) can be formulated as a path integral, to them make use of perturbative techniques to evaluate observables of interest. In the following, a brief derivation of the formalism is presented.

Formally, Eq. (A.9) has a solution of the form

$$|\Phi(t)\rangle = \exp(\mathcal{A}t)\,|\Phi(0)\rangle\,, \tag{A.10}$$

where the initial state of the system $|\Phi(0)\rangle$ corresponds to a particular configuration. This state is constructed by acting on the vacuum state with the corresponding set of creation operators, denoted by $\mathcal{I}$. Thus, $|\Phi(0)\rangle = \mathcal{I}\,|0\rangle$. From Eq. (A.10) any observable O can be written int terms of operators

$$\begin{aligned} \langle O\rangle\,(t) &= \langle\bullet|\,O\exp(\mathcal{A}t)\,|\Phi(0)\rangle & \text{(A.11)}\\ &= \langle\bullet|\,O\exp(\mathcal{A}t)\mathcal{I}\,|0\rangle\,, & \text{(A.12)} \end{aligned}$$

where the so called *abyss* [188] has been introduced

$$\langle\bullet| = \sum_{\{n'_{\mathbf{x}}\}} \langle\{n'_{\mathbf{x}}\}|\,, \tag{A.13}$$

which reduces any *ket* to unity, i.e. $\langle\bullet|\{n'_{\mathbf{x}}\}\rangle = 1$. With Eq. (A.11) the expected particle number at position $\mathbf{y}$ with $O = a^\dagger(\mathbf{y})a(\mathbf{y})$, can be, for example, calculated as

$$\langle\bullet|\,a^\dagger(\mathbf{y})a(\mathbf{y})\,|\Phi(t)\rangle = \sum_{\{n'_{\mathbf{x}}\}} \mathcal{P}(\{n_{\mathbf{x}}\};t)n_{\mathbf{y}}\,\langle\bullet|\{n'_{\mathbf{x}}\}\rangle = n_{\mathbf{y}}. \tag{A.14}$$

The solution Eq. (A.10) corresponds to the Shrödinger picture, where the time-dependence is contained in the evolution operator

$$\exp(\mathcal{A}t) = 1 + \mathcal{A}t + \frac{(\mathcal{A}t)^2}{2!} + \cdots = \sum_{n=0}^{\infty} \frac{(\mathcal{A}t)^n}{n!}, \tag{A.15}$$

which has been written in its (arguably convergent) power series form. However, this formal description does not lend itself for the explicit calculation of observables. For this, the exponential $\exp(\mathcal{A}t)$ can be manipulated in order to rewrite it as a path integral. This procedure later reduces to *replacing* the operators by fields, $a^\dagger$ by $\phi^\dagger$ and a by ϕ in the action. To proceed, time t is discretised into small intervals Δt [187], a procedure referred to as *time slicing* [188], then

$$\begin{aligned} \exp(\mathcal{A}t) &= \lim_{\Delta t\to 0}(1+\Delta t\mathcal{A})^{t/\Delta t} & \text{(A.16)}\\ &= \lim_{\Delta t\to 0}\underbrace{(1+\Delta t\mathcal{A})(1+\Delta t\mathcal{A})\ldots(1+\Delta t\mathcal{A})}_{\frac{t}{\Delta t}\text{ terms}}\,. & \text{(A.17)} \end{aligned}$$

Making use of the identity

$$\mathbf{1}_t = \int \frac{\mathrm{d}\phi \mathrm{d}\phi^\dagger}{\pi} e^{-\phi^\dagger(t)\phi(t)} e^{\phi(t)a^\dagger} |0\rangle \langle 0| e^{\phi^\dagger(t)a}, \tag{A.18}$$

where the eigenfunctions

$$|\phi\rangle = e^{\phi a^\dagger} |0\rangle \quad \text{and} \quad \langle\phi| = \langle 0| e^{\phi^\dagger a} \tag{A.19}$$

of the annihilation a, and creator $a^\dagger$ operators were used. This is eigenfunctions satisfy

$$a|\phi\rangle = \phi|\phi\rangle \quad \text{and} \quad \langle\phi| a^\dagger = \langle\phi| \phi^\dagger. \tag{A.20}$$

Inserting the identity Eq. (A.18), with suitable times, around every term of Eq. (A.16)

$$\exp(\mathcal{A}t) = \lim_{\Delta t \to 0} \mathbf{1}_{t+\Delta t}(1 + \Delta t\mathcal{A})\mathbf{1}_t(1 + \Delta t\mathcal{A})\mathbf{1}_{t-\Delta t} \dots \mathbf{1}_{\Delta t}(1 + \Delta t\mathcal{A})\mathbf{1}_0. \tag{A.21}$$

As described in [188], any pair of brackets in Eq. (A.21) contains a term of the form

$$\Xi(t + \Delta t, t) = \langle 0| \exp(\phi^\dagger(t + \Delta t)a)(1 + \Delta t\mathcal{A}) \exp(-\phi^\dagger\phi(t)) \exp(\phi(t)a^\dagger) |0\rangle\,. \tag{A.22}$$

Using the power series representation of the exponential terms and Eqs. (A.5) lead to

$$\Xi(t + \Delta t, t) = \exp(-\phi^\dagger\phi(t)) \sum_{n,m=0}^{\infty} \frac{\phi^\dagger(t + \Delta t)^n \phi(t)^m}{m!} \langle n| (1 + \Delta t\mathcal{A}) |m\rangle\,. \tag{A.23}$$

As long as the action $\mathcal{A}$ is a normal ordered polynomial in $a^\dagger$ and a, such that the action can be written as $\mathcal{A} = a^{\dagger k} a^l$ for $k, l \in \mathbb{N}$, or as a sum of terms thereof, the following holds

$$\begin{aligned}
\langle n| (1 + \Delta t\mathcal{A}) |m\rangle &= \langle n| 1 |m\rangle + \langle n| \mathcal{A} |m\rangle \Delta t && \text{(A.24)}\\
&= \delta_{n,m} + \langle n| a^{\dagger k} a^l |m\rangle \Delta t && \text{(A.25)}\\
&= \delta_{n,m} + \frac{m!}{(m-l)!} \delta_{n-k}\delta_{m-l}\Delta t, && \text{(A.26)}
\end{aligned}$$

where the last terms is valid for $n \geq k$ and $m \geq l$, and 0 otherwise. Inserting this result in Eq. (A.23) one obtains

$$\begin{aligned}
\Xi(t + \Delta t, t) &= \exp(-\phi^\dagger\phi(t)) \sum_{n,m=0}^{\infty} \frac{\phi^\dagger(t + \Delta t)^n \phi(t)^m}{m!} \langle n| a^{\dagger k} a^l |m\rangle && \text{(A.27)}\\
&= \exp(-\phi^\dagger\phi(t)) \sum_{n,m=0}^{\infty} \frac{\phi^\dagger(t + \Delta t)^n \phi(t)^m}{m!} \left(\delta_{n,m} + \frac{m!}{(m-l)!}\delta_{n-k}\delta_{m-l}\Delta t\right) && \text{(A.28)}\\
&= \exp\left[(\phi^\dagger(t + \Delta t) - \phi^\dagger(t))\phi(t)\right]\left[1 + \Delta t\phi^\dagger(t + \Delta t)^\gamma \phi^\sigma(t)\right]. && \text{(A.29)}
\end{aligned}$$

Interestingly, any notion of operators and their commutation properties has disappeared from the expression for

Ξ, this is a consequence of the normal ordering of the operators in the action [188]. Considering Eq. (A.29), Eq. (A.21) may be written as

$$\exp(\mathcal{A}t) = \lim_{\Delta t \to 0} \int \mathcal{D}\Pi e^{-\phi^\dagger(t+\Delta t)\phi(t+\Delta t)} e^{\phi(t+\Delta t)a^\dagger} |0\rangle \prod_{i=0}^{t/\Delta t} e^{(\phi^\dagger((i+1)\Delta t)-\phi^\dagger(i\Delta t))\phi(i\Delta t)}$$
$$\cdot \left(1 + \Delta t \phi^\dagger((i+1)\Delta t)^\gamma \phi(i\Delta t)^\sigma\right) \langle 0| e^{\phi^\dagger(0)a}, \tag{A.30}$$

where the integration measure is $\mathcal{D}\Pi = (2\pi \mathring{i})^{t/\Delta t} \prod_i^{t/\Delta t} \mathrm{d}\phi^\dagger(i\Delta t)\mathrm{d}\phi(i\Delta t)$. Rewriting the products in Eq. (A.30)

$$\exp(\mathcal{A}t) = \lim_{\Delta t \to 0} \int \mathcal{D}\Pi e^{-\phi^\dagger(t+\Delta t)\phi(t+\Delta t)} e^{\phi(t+\Delta t)a^\dagger} |0\rangle e^{\sum_{i=0}^{t/\Delta t}(\phi^\dagger((i+1)\Delta t)-\phi^\dagger(i\Delta t))\phi(i\Delta t)}$$
$$\cdot e^{\sum_{i=0}^{t/\Delta t} \phi^\dagger((i+1)\Delta t)^\gamma \phi(i\Delta t)^\sigma} \langle 0| e^{\phi^\dagger(0)a}. \tag{A.31}$$

With this expression for $\exp(\mathcal{A}t)$ it is now possible to calculate observables using field-theoretic techniques.

A.3 Observables

As discussed previously, the average of observables can be calculated through Eq. (A.11). In general, observables can be written, in normal ordering, as $O = a^{\dagger k} a^\ell$. However, when acting on the abyss Eq. (A.13) the creation operators in the observable are *consumed* by the abyss, thus the observable simplifies to $O = a^\ell$. In the same way, the initialisation operator can be written as $\mathcal{I} = a^{\dagger r} a^s$, which when acting on the vacuum state reduces to $\mathcal{I} = a^{\dagger r}$. With this considerations, and making use of Eq. (A.31), Eq. (A.11) can be written as

$$\langle O \rangle (t) = \langle \cdot | O \exp(\mathcal{A}t) \mathcal{I} |0\rangle \tag{A.32}$$
$$= \lim_{\Delta t \to 0} \int \mathcal{D}\Pi \phi^\ell(t+\Delta t) \exp\Big[(1 - \phi^\dagger(t+\Delta t))\phi(t+\Delta t) \tag{A.33}$$
$$+ \sum_{i=0}^{t/\Delta t} [\phi^\dagger((i+1)\Delta t) - \phi^\dagger(i\Delta t)]\phi(i\Delta t) + \Delta t \phi^{\dagger\gamma}((i+1)\Delta t)\phi^\sigma(i\Delta t)\Big]\phi^{\dagger r}(0). \tag{A.34}$$

In the limit $\Delta t \to 0$ the sum in the exponential turns into an integral in time, at the same time, the term $\phi^\dagger((i+1)\Delta t) - \phi^\dagger(i\Delta t)$ turns into a time derivative. Hence,

$$\langle O \rangle (t) = \int \mathcal{D}\Pi \phi^\ell(t) \exp\left[(1 - \phi^\dagger(t))\phi(t) + \int_0^t \mathrm{d}t' \left(\frac{d\phi^\dagger}{dt'}\phi(t') + \phi^{\dagger\gamma}(t')\phi^\sigma(t')\right)\right]\phi^{\dagger r}(0). \tag{A.35}$$

It is now convenient at this point to perform the *Doi-shift* of the creation fields [188]

$$\tilde{\phi}(t) = \phi^\dagger(t) - 1, \tag{A.36}$$

which allows for the simplification of some of the terms, and eases some of the calculations of the observables. The observable Eq. (A.35) with the Doi-shifted fields reads

$$\langle O\rangle(t) = \int \mathcal{D}\Pi\phi^{\ell}(t)\exp\left[-\tilde{\phi}(t)\phi(t) + \int_0^t \mathrm{d}t'\left(\frac{d\tilde{\phi}}{dt'}\phi(t') + (\tilde{\phi}+1)^{\gamma}(t')\phi^{\sigma}(t')\right)\right](\tilde{\phi}(0)+1)^{r}. \tag{A.37}$$

Integrating the first term in the integral with respect to time by parts

$$\int_0^t \mathrm{d}t'\frac{d\phi^{\dagger}}{dt'}\phi(t') = \tilde{\phi}(t)\phi(t) - \tilde{\phi}(0)\phi(0) - \int_0^t \mathrm{d}t'\tilde{\phi}(t')\frac{d\phi(t')}{dt'}, \tag{A.38}$$

Eq. (A.37) the reads

$$\langle O\rangle(t) = \int \mathcal{D}\Pi\phi^{\ell}(t)\exp\left[-\tilde{\phi}(0)\phi(0) + \int_0^t \mathrm{d}t'\left(\tilde{\phi}(t')\frac{d\phi(t')}{dt'} + (\tilde{\phi}+1)^{\gamma}(t')\phi^{\sigma}(t')\right)\right](\tilde{\phi}(0)+1)^{r}. \tag{A.39}$$

As it turns out, the $\tilde{\phi}(0)\phi(0)$ terms can be dropped out [188], and the lower bound of the integral in time can be extended to negative infinity, which permits the use of Fourier transforms. Finally, the expectation of the observable $O = a^{\ell}(t)$, with initialisation operator $\mathcal{I} = a^{\dagger r}(t_0)$ can be written as

$$\langle O\rangle(t) = \int \mathcal{D}\Pi\phi^{\ell}(t)\exp()(\tilde{\phi}(t_0)+1)^{r}, \tag{A.40}$$

where the field theoretic action

$$\mathcal{A} = \int_{-\infty}^{\infty}\mathrm{d}t'\left(-\tilde{\phi}(t')\frac{d\phi}{dt'} + \phi^{\sigma}(t')(\tilde{\phi}(t')+1)^{\gamma}\right). \tag{A.41}$$

was introduced. In this section, it was shown how though the Doi-Peliti formalism a master equation for a reaction diffusion process can be cast in a field-theoretic language.

Appendix B

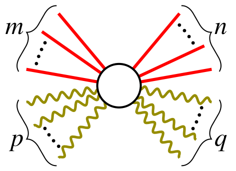
Field theory of the BRW

B.1 Relevant interactions for the BRW

Whether a particular interaction is allowed by the basic process described by the action (2.19) is a matter of some topological constraints, which are discussed in the first part of this section. Whether it is infrared-relevant is determined by its engineering dimension, which is discussed in the second part of this section. Combining topological and engineering constraints will then produce a finite number of interaction vertices to consider. Constraints that avoid certain, otherwise relevant vertices from being generated are preserved under renormalisation.

The general proper vertex

$$\Gamma^{\left[\begin{smallmatrix} m & n \\ p & q \end{smallmatrix}\right]} = \tag{B.1}$$



where time is to be read from right to left, are the one-particle irreducible graphs of the amputated correlation function

$$G^{\left[\begin{smallmatrix} m & n \\ p & q \end{smallmatrix}\right]}\Big(r, D, \tau, s, \sigma, \lambda, \kappa, \chi, \xi; \{\mathbf{k}_1, \ldots, \mathbf{k}_{m+n+p+q}; \omega_1, \ldots, \omega_{m+n+p+q}\}\Big) = \tag{B.2}$$

$$\Big\langle \underbrace{\phi(\mathbf{k}_1, \omega_1) \ldots \phi(\mathbf{k}_m, \omega_m)}_{m \text{ terms}} \underbrace{\psi \ldots \psi}_{p \text{ terms}} \underbrace{\tilde{\phi} \ldots \tilde{\phi}}_{n \text{ terms}} \underbrace{\tilde{\psi} \ldots \tilde{\psi}}_{q \text{ terms}} \Big\rangle . \tag{B.3}$$

The bare couplings are the tree-level contributions to the proper vertices:

$$\tau = \Gamma^{\left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]} + \text{h.o.t.} \qquad s = \Gamma^{\left[\begin{smallmatrix} 2 & 1 \\ 0 & 0 \end{smallmatrix}\right]} \qquad \lambda = \Gamma^{\left[\begin{smallmatrix} 0 & 1 \\ 1 & 1 \end{smallmatrix}\right]} + \text{h.o.t.} \tag{B.4a}$$

$$\sigma = \Gamma^{\left[\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right]} + \text{h.o.t.} \qquad \chi = \Gamma^{\left[\begin{smallmatrix} 0 & 1 \\ 2 & 1 \end{smallmatrix}\right]} + \text{h.o.t.} \qquad \kappa = \Gamma^{\left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right]} + \text{h.o.t.} \qquad \xi = \Gamma^{\left[\begin{smallmatrix} 1 & 1 \\ 2 & 1 \end{smallmatrix}\right]} + \text{h.o.t.} \tag{B.4b}$$

Denoting, where applicable, terms of higher order in non-linear couplings by h.o.t.. Every proper vertex has a number of topological constraints, since any such term needs to arise from the perturbative expansion of the action as a one-particle irreducible (connected, amputated) diagram made from the bare vertices available in the theory. By inspection, the following constraints are found, which will then be used to determine all relevant, possible couplings below: Firstly, all non-linear vertices in the field theory (all diagrams except the propagator of the tracer particles) have at least one straight leg coming in, $n \geq 1$. Secondly, all vertices have at least as many wavy legs coming out, as come in, $p \geq q$. Thirdly, there are at least as many outgoing legs (wavy or straight), as there are incoming straight legs, $m + p \geq n$.

The engineering dimension of the general proper vertex can be determined from the considerations at the beginning of Section 2.4.2, using the fact that each proper vertex may be seen as an effective coupling, which, after integration over real time and space, gives rise to a dimensionless contribution to the action, $\mathtt{L}^d\mathtt{T}[\Gamma^{\left[\begin{smallmatrix} m & n \\ p & q \end{smallmatrix}\right]}\tilde{\phi}^m\tilde{\psi}^p\phi^n\psi^q] = 1$, so that

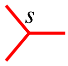
$$[\Gamma^{\left[\begin{smallmatrix} m & n \\ p & q \end{smallmatrix}\right]}] = \mathtt{L}^{d(n+q-1)+2(m-n+2p-2q-1)}\mathtt{A}^{p-q}\mathtt{B}^{m-n+p-q}\mathtt{C}^{n-m-2p+2q+1}. \tag{B.5}$$

Demanding that (effective) transmutation τ, branching s and diffusion D remain relevant at any scale, the independent dimensions $\mathtt{A}$, $\mathtt{B}$ and $\mathtt{C}$, respectively, are set to unity $\mathtt{A} = \mathtt{B} = \mathtt{C} = 1$. The (marginally) infrared-relevant couplings are those whose engineering dimension (in $\mathtt{L}$) is non-positive. At the upper critical dimension $d = d_c = 4$, the inequality $d(n + q - 1) + 2(m - n + 2p - 2q - 1) \leq 0$ gives

$$m + n + 2p \leq 3. \tag{B.6}$$

The field theory needs to include all vertices $\Gamma^{\left[\begin{smallmatrix} m & n \\ p & q \end{smallmatrix}\right]}$ with (non-negative) integers m, n, p and q that fulfill Eq. (B.6) together with the topological constraints $n \geq 1$, $p \geq q$ and $m + p \geq n$ discussed earlier. To find them, two cases can be destinguished for Eq. (B.6):

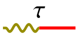
- $p = 0 \overset{p \geq q}{\Longrightarrow} q = 0$, then $m + n \leq 3$. Under the topological constraint $m + p \geq n$ there are only two viable solutions: $m = n = 1$, or $m = 2$ and $n = 1$, that correspond to

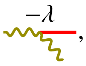

and


, (B.7)

the bare propagator for active walkers, and branching of active walkers, respectively.

- $p = 1 \implies m + n \leq 1$. Only the propagator of the immobile particles allows for $n = 0$. Otherwise, $n \geq 1$ requires $m = 0$. The constraint $p \geq q$ leaves only $q = 0$ and $q = 1$. As a result, there are three viable combinations: Firstly, $m = n = 0$ and $q = 1$, secondly, $m = q = 0$ and $n = 1$, thirdly, $m = 0$ and $n = q = 1$, which correspond to


,


and


, (B.8)

the bare propagator of immobile tracer particles, the transmutation vertex and hindrance of spawning, respectively.

Together with the propagators, the vertices in (B.7) and (B.8) represent all (marginally) relevant couplings at $d = d_c = 4$, consisting of the (bilinear) transmutation, τ, and the interaction vertices s of branching and $-\lambda$ of suppression of spawning.

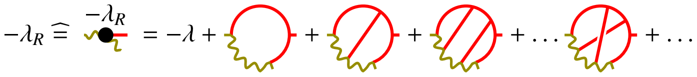
In the following section the field theoretic renormalisation is carried out in dimensions $d = 4 - \varepsilon$ to derive the scaling of the number of distinct sites visited by the BRW.

B.1.1 Renormalisation of the couplings

As far as the observables Eq. (2.33) are concerned, the only couplings to consider are τ and λ. Both are renormalised by the same set of loops

(B.9)

and



(B.10)

where the renormalised coupling (•) is indicated by the subindex R. All diagrams in τ_R and $-\lambda_R$ are amputated (note the short legs). Of all the contributions to the renormalisation of τ and λ the non-crossing diagrams, such as the first three in Eqs. (B.9) and (B.10), are easily calculated (see appendix B.3 for details). All of these non-crossing diagrams can then be summed over by virtue of field-theoretic renormalisation. The last diagram, and further crossing diagrams, in both Eq. (B.9) and Eq. (B.10), on the other hand, require further explicit calculation and subsequent summation. However, the Ward identity [48]

$$\frac{\partial \tau_R}{\partial \tau} = \frac{\lambda_R}{\lambda}, \tag{B.11}$$

which arises from the the symmetry of the action under the shift $\tilde{\psi} \to \tilde{\psi} + \Sigma$, Σ being a constant. This implies that all exponents can be determined without calculating any of the diagrams explicitly. To see this, a dimensionless coupling $g = \lambda s \mathcal{U} \mu^{-\epsilon} D^{-2} \Gamma(\epsilon/2)$ is introduced, with suitable numerical factor $\mathcal{U}$ and arbitrary inverse length scale μ. Both couplings therefore renormalise identically,

$$\tau_R = \tau Z(g) \qquad \text{and} \qquad \lambda_R = \lambda Z(g) \tag{B.12}$$

with $Z(g)$ governing the renormalisation of both λ and τ. To one loop and with suitable $\mathcal{U}$, the Z-factor becomes $Z(g) = 1 - g$, see Eqs. (B.9) and (B.10), and Sec. B.3. However, there is no need to determine the precise dependence of Z on g as far as scaling is concerned. It suffices to know that the renormalised, dimensionless

$$g_R = \lambda_R s \mathcal{U} \mu^{-\epsilon} D^{-2} \Gamma(\epsilon/2) \tag{B.13}$$

$$= Z \lambda s \mathcal{U} \mu^{-\epsilon} D^{-2} \Gamma(\epsilon/2) \tag{B.14}$$

has β-function

$$\beta_g = \frac{\mathrm{d}g_R}{\mathrm{d}\ln\mu} = -\epsilon g_R + g_R \frac{\mathrm{d}\ln Z}{\mathrm{d}\ln\mu} \tag{B.15}$$

in $d < 4$, which implies $\mathrm{d}\ln Z/\mathrm{d}\ln\mu = \epsilon$ at the root $\beta_g(g = g^*) = 0$, irrespective of $\mathcal{U}$ and therefore irrespective of the presence or absence of the crossing diagrams. It follows that $Z \sim \mu^\epsilon$ in $d \leq 4$ and therefore the effective transmutation rate is $\tau_{\text{eff}} \sim \tau Z \sim \mu^\epsilon$. In the limit of $t \to \infty$, for systems of linear size L, the characteristic scale is $\mu \sim L^{-1}$ and thus $\tau_{\text{eff}} \sim L^{-\epsilon}$. With open boundary conditions, the random walkers visit $\sim L^2$ sites during the course of their lifetimes, leaving behind $\sim \tau_{\text{eff}} L^2 \sim L^{2-\epsilon}$ immobile tracer particles in dimensions greater than 2, so that $\langle a \rangle (t, L) \sim L^{d-2}$. This average is bounded from below by a constant, as at least one site is always visited, so that $\langle a \rangle (t, L)$ approaches a constant below 2 dimensions. As for the time-dependence, the characteristic inverse scale μ is proportional to $t^{-1/2}$ because the dynamical exponent $z = 2$ in $\mu \sim t^{-1/z}$ remains unchanged. It follows that $\langle a \rangle (t, L) \sim t^{(d-2)/2}$.

B.2 Calculating scaling of higher-order correlation functions

The scaling of higher-order correlation functions is derived, within the field theory, from the solution of the Callan–Symanzik equation [48] for the general proper vertex Eq. (B.1), from which the scaling of the moments of the total number of distinct sites visited follow, Eq. (2.37). From dimensional analysis (Sec. 2.4.2), and by introducing a bare scale μ_0, related to μ by $\mu = \mu_0 \ell$, the general proper vertex, Eq. (B.1), then satisfies

$$\Gamma^{\left[\begin{smallmatrix} m & n \\ p & q \end{smallmatrix}\right]}(r, D, \tau, s, \sigma, \lambda, \kappa, \chi, \xi; \{\mathbf{k}; \omega\}) \;=\; \ell^{-d(n+q-1)-2(m-n+2p-2q-1)+(p-q)\gamma_\tau} \Gamma^{\left[\begin{smallmatrix} m & n \\ p & q \end{smallmatrix}\right]}\left(\frac{r}{\ell^2}, D, \tau, s, \sigma, \lambda, \kappa, \chi, \xi; \left\{\frac{\mathbf{k}}{\ell}; \frac{\omega}{\ell^2}\right\}\right), \tag{B.16}$$

asymptotically in small ℓ and provided that r is close enough to the critical point, $r_c = 0$. For the transmutation vertex, where $p = n = 1$ and $q = m = 0$,

$$\Gamma^{\left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]}(r, D, \tau, s, \sigma, \lambda, \kappa, \chi, \xi; \{\mathbf{k}; \omega\}) = \ell^{\gamma_\tau} \Gamma^{\left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]}\left(\frac{r}{\ell^2}, D, \tau, s, \sigma, \lambda, \kappa, \chi, \xi; \left\{\frac{\mathbf{k}}{\ell}; \frac{\omega}{\ell^2}\right\}\right), \tag{B.17}$$

is found, with $\gamma_\tau = \varepsilon = 4 - d$. Generally, for observables of the form Eq. (2.33), where $n = 1$ and $q = m = 0$

$$\Gamma^{\left[\begin{smallmatrix} 0 & 1 \\ p & 0 \end{smallmatrix}\right]}(r, D, \tau, s, \sigma, \lambda, \kappa, \chi, \xi; \{\mathbf{k}; \omega\}) = \ell^{4(1-p)+p\gamma_\tau} \Gamma^{\left[\begin{smallmatrix} 0 & 1 \\ p & 0 \end{smallmatrix}\right]}\left(\frac{r}{\ell^2}, D, \tau, s, \sigma, \lambda, \kappa, \chi, \xi; \left\{\frac{\mathbf{k}}{\ell}; \frac{\omega}{\ell^2}\right\}\right). \tag{B.18}$$

The scaling of the first moment of the number of distinct sites visited, $\langle a(t) \rangle$, as function of time, t, can be obtained by analysing the scaling of

$$\langle a(t) \rangle = \int \mathrm{d}^d\mathbf{x} \left\langle \psi(\mathbf{x}, t) \tilde{\phi}(\mathbf{x}_0, 0) \right\rangle \tag{B.19}$$

$$\widehat{=} \int \text{đ}\omega \text{đ}\omega_0 \mathrm{e}^{-\mathring{\imath}\omega t}$$

$$\Big|_{\mathbf{k}=0} \tag{B.20}$$

$$= \int \text{đ}\omega \mathrm{e}^{-\mathring{\imath}\omega t} \frac{1}{-\mathring{\imath}\omega + \epsilon'} \Gamma^{\left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]} \frac{1}{-\mathring{\imath}\omega + r}. \tag{B.21}$$

Note that the diagram shown here corresponds to the full (not amputated) correlation. According to Eq. (B.21), $\Gamma^{\left[\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right]}$ scales like

$$\Gamma^{\left[\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right]}(L^{-2}, D, \tau, s, \sigma, \lambda, \kappa, \chi, \xi; \{\mathbf{k}; \omega\}) = L^{-\gamma_\tau}\Gamma^{\left[\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right]}\left(1, D, \tau, s, \sigma, \lambda, \kappa, \chi, \xi; \left\{\mathbf{k}L; \omega L^2\right\}\right), \tag{B.22}$$

if we identify $r \sim L^{-2}$ and $\ell \sim L^{-1}$, which means that the effective transmutation rate scales like $L^{-\varepsilon}$ in large linear system size L, as $\gamma_\tau = \varepsilon = 4 - d$. In long time t, the integral over ω in Eq. (B.21) has the effect of evaluating $\Gamma^{\left[\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right]}\frac{1}{\mathring{\imath}\omega+r}$ at $\omega = 0$, because

$$\lim_{t\to\infty}\lim_{\epsilon'\to 0}\int_{-\infty}^{\infty} \exp -\mathring{\imath}\omega t \frac{1}{-\mathring{\imath}\omega + \epsilon'} f(\omega) = f(0) \tag{B.23}$$

provided $f(\omega)$ has no pole at 0.

It follows that

$$\lim_{t\to\infty} \langle a(t)\rangle \propto L^{2-\epsilon}. \tag{B.24}$$

For higher moments, on the basis of Eq. (B.18),

$$\lim_{t\to\infty} \langle a^p(t)\rangle \propto L^2 L^{pd-4}\Gamma^{\left[\begin{smallmatrix}0&1\\p&0\end{smallmatrix}\right]}(1, D, \tau, s, \sigma, \lambda, \kappa, \chi, \xi; \{0, 0\}). \tag{B.25}$$

We thus recover the finite-size scaling results Eqs. (2.37b) and (2.38b) of Section 2.5 for the p-th moment of the volume explored by a branching random walk

$$\lim_{t\to\infty} \langle a^p(t)\rangle \propto \begin{cases} L^{dp-2} & \text{if } \varepsilon > 0 \\ L^{4p-2} & \text{if } \varepsilon < 0 \end{cases} \tag{B.26}$$

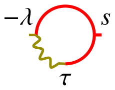
where $\varepsilon > 0$ and $\varepsilon < 0$ separate regions below and above the upper critical dimension, $d_c = 4$, respectively. The dimensionality of the embedding space enters only below the upper critical dimension. Above the upper critical dimension, fluctuations and interactions become asymptotically irrelevant and the process can be considered as free.

The above analysis is easily extended to scaling in time, using $t \propto \mu^{-z}$ with $z = 2$ as the relevant scale, thereby reproducing Eqs. (2.37a) and (2.38a).

B.3 Loop integrals

The non-crossing diagrams, such as the first three in Eqs. (B.9) and (B.10), are calculated through the integral

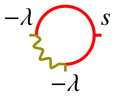
$$I_\tau =$$



$$= \int \text{đ}^d k\, \text{đ}\omega \, \frac{\tau}{-\mathring{\imath}\omega + \epsilon'} \frac{1}{\omega^2 + (Dk^2 + r)^2} = \tau \frac{1}{2} \frac{r^{-\varepsilon/2}}{(4\pi D)^{d/2}} \Gamma(\varepsilon/2)\,, \tag{B.27}$$

and (essentially identical)

$$I_{-\lambda} =$$



$$= \int \text{đ}^d k \text{đ}\omega \, \frac{-\lambda}{-\mathring{\imath}\omega + \epsilon'} \frac{1}{\omega^2 + (Dk^2 + r)^2} = -\lambda \frac{1}{2} \frac{r^{-\varepsilon/2}}{(4\pi D)^{d/2}} \Gamma(\varepsilon/2) \,, \tag{B.28}$$

where the lower part of the loop carries the coupling τ in case of contributing to τ or the coupling $-\lambda$ and an incoming wavy leg in case of contributing to λ. The integration measure is $\text{đ}^d k \text{đ}\omega = \mathrm{d}^d k \mathrm{d}\omega / (2\pi)^{d+1}$.

B.4 Definition of Fourier transform

Throughout this thesis, the Fourier transform $\mathcal{F}[f(\mathbf{x}, t)]$ of a function $f(\mathbf{x}, t)$ in space $\mathbf{x}$ and time t is denoted as $\hat{f}(\mathbf{k}, \omega)$, where the spatial momentum $\mathbf{k}$ is the conjugate of the position $\mathbf{x}$, and the frequency ω is the conjugate of time t. The direct Fourier transform is defined as

$$f(\mathbf{k}, \omega) = \int \exp \mathring{\imath}\omega t - \mathring{\imath}\mathbf{k} \cdot \mathbf{x} f(\mathbf{x}, t) \mathrm{d}^d \mathbf{x} \mathrm{dt}, \tag{B.29}$$

so that the inverse Fourier transform is

$$f(\mathbf{x}, t) = \int \exp -\mathring{\imath}\omega t + \mathring{\imath}\mathbf{k} \cdot \mathbf{x} f(\mathbf{k}, \omega) \text{đ}^d \mathbf{k} \text{đ}\omega, \tag{B.30}$$

where $\text{đ}^d \mathbf{k} = (1/2\pi)^d \mathrm{d}\mathbf{k}$, $\text{đ}\omega = (1/2\pi)\mathrm{d}\omega$, and d is the spatial dimension.

B.5 Extension to general graphs

In this section, further details about the extension of the results to general graphs are provided. The loops integrated over in Eqs. (B.27) and (B.28) are in fact integrals over the spectrum of the Laplacian accounting for the diffusion on the graph considered. Generalising to arbitrary graphs, the Laplacian is to be replaced by a lattice-Laplacian and the integral in Eqs. (B.27) and (B.28) by a suitable sum or, equivalently, an integral with suitable spectral density. In fact, the d-dimensional integral in Eqs. (B.27) and (B.28) can be seen as an integral over all distinct eigenvalues $\mathbf{k}^2$ of the Laplacian entering with weight $w(k)\mathrm{d}k = S_d k^{d-1} \mathrm{d}k$ with $S_d = 2\pi^{d/2}/\Gamma(d/2)$. On regular lattices, their Hausdorff dimension d coincides with the spectral dimension d_s characterising, in particular, the small k asymptote of $w(k) \sim k^{d_s - 1}$. Replacing $\int \mathrm{d}^d k$ by $\int \mathrm{d}k \, w(k)$ suggests that the results derived above remain valid by replacing d by d_s, in order to recover the scaling of the various observables in arbitrary graphs with spectral dimension d_s. The replacement $d \to d_s$ hinges crucially on the fact that d_s characterises the scaling of the spectral density of the Laplacian. If this operator itself renormalises, then a different spectral density may be needed. In other words, d_s may not be the correct dimension if the Laplacian renormalises, *i.e.* if the anomalous dimension does not vanish, $\eta \neq 0$ [50]. This argument relies on the assumption that vertices such as Eq. (B.1) preserve momentum, that is integrals of the form

$$I_n(\mathbf{k}_1, \mathbf{k}_2, \ldots, \mathbf{k}_n) = \int \mathrm{d}^d x \, u_{\mathbf{k}_1}(x) u_{\mathbf{k}_2}(x) \ldots u_{\mathbf{k}_n}(x) \tag{B.31}$$

over eigenfunctions $u_{\mathbf{k}}(x)$ of the Laplacian with eigenvalue $\mathbf{k}\cdot\mathbf{k}$ vanish for off-diagonal terms, i.e. whenever $\mathbf{k}_1+\mathbf{k}_2+\ldots+\mathbf{k}_n\neq 0$. This condition can be further relaxed by demanding merely that off-diagonal terms are sub-leading as observed in the presence of boundaries [189, 31].

Considering only graphs which are translationally invariant such that the indices $\mathbf{j}_m$ of the q neighbours $m=1,\ldots,q$ of any node $\mathbf{i}$ can be determined by adding the same set of translational lattice vectors, $\mathbf{d}_1,\ldots,\mathbf{d}_q$, such that $\mathbf{j}_m=\mathbf{i}+\mathbf{d}_m$, it is easy to show that the Laplacian has exponential eigenfunctions and any of their products are an eigenfunction as well, so that $I_n(\mathbf{k}_1,\mathbf{k}_2,\ldots,\mathbf{k}_n)=I_2(\mathbf{k}_1,\mathbf{k}_2+\ldots+\mathbf{k}_n)$, which vanishes by orthogonality for any $\mathbf{k}_1+\mathbf{k}_2+\ldots+\mathbf{k}_n\neq\mathbf{0}$, i.e. the assumption of momentum conservation mentioned above is fulfilled.

B.6 Numerics for the scaling of moments

The scaling of the moments $\langle a^p\rangle(t,L)$ for $p=1,2,3,\ldots,5$, as function of time t in the limit $L\to\infty$, and as function of the system size L in the limit $t\to\infty$ were obtained from numerical Monte Carlo simulations and fitted against a power-law

$$f(x)=Ax^B \tag{B.32}$$

and a power-law with corrections of the form

$$g(x)=Ax^B+Cx^{B-1/2}. \tag{B.33}$$

The fitting parameter B in Eqs. (B.32) and (B.33) provides the estimates of the exponents that characterise the scaling of the moments in time t and sistem size L (or N, see main text), by fitting the numerical estimates against $f(x)$ and $g(x)$, with x replaced by t and L, respectively. At large times the moments display plateauing due to finite size effects.

For the scaling in system size L, the data was fitted for the latest time point available against Eq. (B.32) and used the estimates of A and B as the initial values for a fit against Eq. (B.33), which gave the final estimates of the finite size scaling exponents.

For the scaling in time t, the data was fitted for the largest system, of size $L=L_{\max}$. The fitting range in t for each moment was determined systematically as follows. To remove the time-point affected by the finite size effects, the upper bound of the fitting range was defined as the time t^{up} for which the lowest moment displaying algebraic divergence ($p=p_{\mathrm{low}}$) reached a value of half the maximum value in the plateau, *i.e.* $\langle a^{p_{\mathrm{low}}}\rangle(t^{\mathrm{up}},L_{\max})=\max_t\left(\langle a^{p_{\mathrm{low}}}\rangle(t,L_{\max})\right)/2$. For the preferential attachment network the plateau was observed to occur at an earlier time point than t^{up}, probably due to the high connectivity of the networks, so the upper bound was set to $t^{\mathrm{up}}_{pa}=(1/5)\max(\left\langle a^k\right\rangle)$, in this case.

To find the lower bound t_{low} of the fitting range in t both equations, (B.32) and (B.33), were fitted to the data for $L_{\max}$. Defining $\hat{f}_{[t^*,t^{\mathrm{up}}]}(t)$ and $\sigma^{\hat{f}}_{[t^*,t^{\mathrm{up}}]}(t)$ as the values and errors, respectively, of fitting Eq. (B.32) to the data in the range $t\in[t^*,t^{\mathrm{up}}]$, and $\hat{g}_{[t^*,t^{\mathrm{up}}]}(t)$ and $\sigma^{\hat{g}}_{[t^*,t^{\mathrm{up}}]}(t)$ as the values and errors, respectively, of fitting Eq. (B.33) to the same data set and range. Further, $N_{[t^*,t^{\mathrm{up}}]}$ is defined as the number of data points within the fitting interval $[t^*,t^{\mathrm{up}}]$.

The lower bound for the time range t_{low} is the earliest time at which both fitting models (B.32) and (B.33) agree within errors, that is

$$t_{\text{low}} = \min\left\{t^* : \left|\hat{f}_{[t^*,t^{\text{up}}]}(t^*) - \hat{g}_{[t^*,t^{\text{up}}]}(t^*)\right| \leq \sqrt{N_{[t^*,t^{\text{up}}]}}\,\max\left(\sigma^f_{[t^*,t^{\text{up}}]}(t^*), \sigma^g_{[t^*,t^{\text{up}}]}(t^*)\right)\right\}. \tag{B.34}$$

Where correlations between estimates of moments were accounted for by rescaling the error by the square root of the number of data points in the fitting range, $N_{[t^*,t^{\text{up}}]}$. The exponents characterising the time depenence of the moments are determined by fitting the data in the range $[t^*, t^{\text{up}}]$ against Eq. (B.33).

The fitting of the power laws, Eqs. (B.32) and (B.33), was done by means of the Levenberg-Marquardt algorithm [190]. The numerical results for the asymptotic scaling, in time $\langle a^p\rangle(t) \sim t^{\alpha_p}$ and in system size $\langle a^p\rangle(t) \sim L^{\beta_p}$, are presented in table 2.3 and 2.3, provided these observables display an algebraic divergence.

Numerical implementation

In the numerical implementation, an active particle is allowed to diffuse by hopping from the site it resides on to a nearest neighbouring site with rate H, branch with rate s by placing an identical offspring at the present site or become extinct with rate e. Each distinct site visited is recorded, equivalent to taking the limit $\gamma \to \infty$ in the theory. The instantaneous number $a(t, L)$ of distinct sites visited up until any time t is therefore the number of sites recorded. Parameters were chosen such that $H + s + e = 1$, H was set to 0.1, and $e = s = 0.45$. If M walkers are present in the system at a given time point the waiting time for the next event (hopping, branching or extinction) is determined by $-\ln(1-u)/M$ where u is a uniformly distributed random variable in the interval $[0, 1)$. For every lattice size $10^6 - 10^9$ realisations of the process were performed.

Regular, integer-dimensional lattices

The regular lattices studied here are hypercubic d-dimensional lattices, characterised by their linear size $L = 2^m - 1$, $m \geq 4$, which is chosen to be odd so that it contains a well-defined centre site, on which the single active walker is initially placed. To study finite size scaling, absorbing boundary conditions were applied. However, it was observed that the boundary conditions have no effect in the scalings (data not shown). The numerical results were fitted to a power-law as described in Sec. B.6, to obtain the values in Tabs. 2.3 and 2.2.

Sierpinski carpet

The Sierpinski carpets were constructed from two dimensional lattices of linear dimension 3^m, $m \geq 2$. The lattice was divided into 3^2 equal sub-squares each of size 3^{m-1}, the central square was removed, leaving $3^2 - 1$ sub-squares. The procedure is iterated over the remaining sub-squares. The spectral dimension of the Sierpinski carpet has been estimated to be $d_s = 1.86$ [56, 191]. A random point around the central hole of the fractal was used as the initial location of the walker in every realisation.

Random trees

The critical random tree networks [192] were constructed as a critical Galton-Watson process, where every node has either 0, 1, or 2 descendants, such that the mean degree of the network is 2. Networks were generated with $2^6 - 2^{12}$ nodes. These graphs have no closed loops. The spectral dimension of the random tree ensemble is $d_s = 4/3$ [57]. For every realisation of the process, a new random tree was generated, and a node was selected at random as the starting location of the initial walker.

Preferential-attachment network

A preferential attachment (PA) network is a class of scale-free networks, characterised by a power-law degree distribution. The Barabási–Albert model of preferential attachment [58] was used. It was initialized with a single node to generate networks with $2^{12} - 2^{19}$ nodes. The networks have power-law degree distribution with exponent -2.9 and mean degree $\langle k \rangle = 6.3$ (Fig. B.1). For every realisation of the process, a new network was constructed, and a node was selected at random as the starting location of the initial walker.

B.6.1 Degree distribution of the PA network

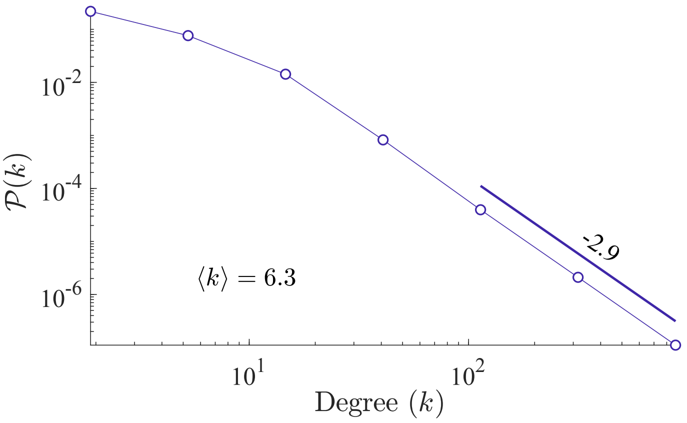


Figure B.1: **Degree distribution of the preferential attachment networks** used for the simulations presented in Sec. 2.6

Appendix C

Self-replication of vegetation patches

C.1 Data analysis

Through the use of satellite images obtained from *Google Earth*, regions in Zambia and Mozambique were located in which vegetation patches dominate the landscape.

C.1.1 Patch detection

In order to detect the vegetation and extract properties of interest a simple segmentation algorithm was implemented in Matlab (R2016b). The boundaries of every detected patch were used to extract the patch geometrical features, such as area, perimeter, equivalent diameter, and centroid positions, which were then used in the spatial analysis.

C.1.2 Equivalent radius and nearest-neighbor distance

The equivalent radius of each structure r_{eq} was calculated as

$$r_{eq} = \sqrt{\frac{A_s}{\pi}} \tag{C.1}$$

where A_s corresponds to the area of the structure. The nearest neighbor distances was obtained by finding the minimum distance between every patch and every other patch.

C.1.3 Discrete Fourier transform

The discrete Fourier transform $\hat{U}$ of a binary image $U(i, j)$ of size m-by-n was computed numerically is MATLAB as

$$\hat{U}(i+1, j+1) = \sum_{j=0}^{m-1} \int_{k=0}^{n-1} U(i+1, j+1) \exp(-2\pi(i/n + j/m)) \tag{C.2}$$

Where (i, j) are the indices of the image pixels. Usually $\hat{U}$ is referred to as the spectrum of U. Peaks in the spectrum indicate dominant wavelengths or typical sizes of the patches detected in the satellite images (see [193, 194]).

C.2 Satellite images

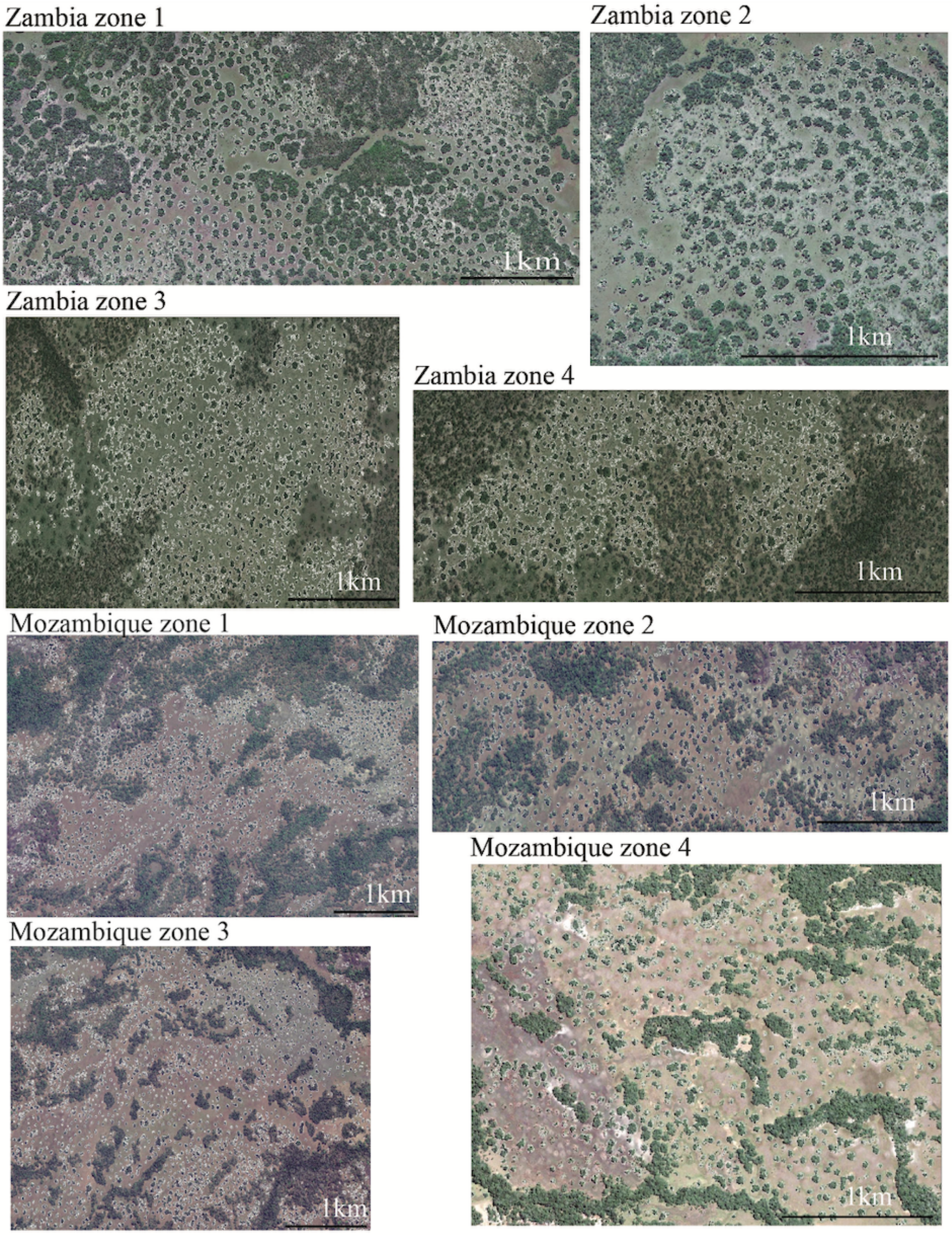


Figure C.1: **Zones considered for the analysis of vegetation patches**, Coordinates: Zambia zone 1 [13°47′23.99″S, 25°17′11.18″E], Zambia zone 2 [13°51′17.88″S, 25°22′33.86″E], Zambia zone 3 [14°39′16.49″S, 25°49′54.86″E], Zambia zone 4 [14°40′21.72″S, 25°49′38.56″E], Mozambique zone 1 [18°41′19.32″S, 35°30′37.62″E], Mozambique zone 2 [18°40′23.96″S, 35°33′32.58″E], Mozambique zone 3 [18°41′55.52″S, 35°38′14.45″E], Mozambique zone 4 [18°49′48.31″S, 35°36′22.37″E]. The white boundaries enclosing each structure corresponds to results of the automatic detection of objects. This figure, for which I hold the copyright, has been published as part of [6].

Location	Rainfall mm/year	Mean temp. 1901-2006	Mean temp. 2006-2016
Zambia zone 1	903	21.99°	22.72°
Zambia zone 2	903	21.99°	22.72°
Zambia zone 3	903	21.99°	22.72°
Zambia zone 4	903	21.99°	22.72°
Mozambique zone 1	1348	24.89°	25.49°
Mozambique zone 2	1348	24.89°	25.49°
Mozambique zone 3	1348	24.89°	25.49°
Mozambique zone 4	1348	24.89°	25.49°

Table C.1: **Information for analyzed regions**. Mean annual rainfall, and mean temperature for the periods 1901-2006 and 2006-2016 (Obtained from CRU TS3 [14]) are shown.

Appendix D

Effective potentials in cell-signal interactions

D.1 Methods

D.1.1 Cell diffusion

In the simulations, at each time step cells move a fixed step length, ℓ, in a random direction on the two-dimensional plane. In the continuum limit, the two-dimensional diffusion equation for the probability, $p(\mathbf{r}, t)$, of finding a cell at position $\mathbf{r}$, at time t, is given by

$$\partial_t p(\mathbf{r}, t) = D\nabla^2 p(\mathbf{r}, t), \tag{D.1}$$

where D is the diffusion constant, and ∇^2 is the two-dimensional Laplacian. The Green function of Eq. (D.1) on the infinite plane $\mathbb{R}^2$ is given by

$$p(\mathbf{r}, t) = \frac{1}{4\pi Dt} \exp\left(-\frac{|\mathbf{r}|^2}{4Dt}\right), \tag{D.2}$$

from which the expected mean square displacement of a cell can be calculated,

$$\left\langle |\mathbf{r}(t)|^2 \mid = \right\rangle \int |r|^2 \, p(\mathbf{r}, t) = 4Dt. \tag{D.3}$$

On the other hand, the mean square displacement of a cell performing t steps of fixed length ℓ in random directions is given by

$$\left\langle |\mathbf{r}(t)|^2 \mid = \right\rangle \delta^2 t \tag{D.4}$$

Comparing Eqs. (D.3) and (D.4) the diffusion constant for simulated cells can be written as

$$D = \frac{\delta^2}{4}. \tag{D.5}$$

It follows that Eq. (D.1) with a diffusion constant given by Eq. (D.5) is the asymptotically Gaussian distribution of the simulated cells' positions in the limit of large times t.

Experimental diffusion coefficient

From the experimental trajectories of single cells, the diffusion constant D was calculated as

$$D(\Delta t) = \frac{\left\langle r^2 \right\rangle}{4\Delta t}, \tag{D.6}$$

where Δt is the averaging time-window considered for computing the mean-square displacement. Given the trajectory, $\boldsymbol{r}_i(t)$, of the ith cell, measured with temporal resolution δ (one minute in the experimental data presented in Ch. 4) at times $t = 0, \delta, 2\delta, ..., T$, the mean-square displacement, as a function of Δt, a multiple of δ, for the cell ensemble can be calculated as

$$\left\langle r^2 \right\rangle = \frac{1}{N(T-\Delta t)/\delta} \sum_{i=0}^{N} \sum_{n=0}^{(T-\Delta t)/\delta} ||\boldsymbol{r}_i(n\delta + \Delta t) - \boldsymbol{r}_i(n\delta)||^2, \tag{D.7}$$

where N is the total number of cells in the experiment [195].

D.1.2 van Hove correlation function

Given a collection of K particles of a given species, with positions $\mathbf{r}_i(t)$, the van Hove dynamic correlation function reads

$$G(\mathbf{r}, \tau) = \frac{1}{K} \left\langle \sum_{i=1}^{K} \delta[\boldsymbol{r} - \boldsymbol{r}_j(\tau) + \boldsymbol{r}_i(0)] \right\rangle \tag{D.8}$$

which measures the average density of particles at position $\mathbf{r}$ at time τ, given that there was a particle at the origin at time 0 [141]. The brakets $\langle \cdot \rangle$ indicate the ensemble average. The van Hove correlation function can be naturally split into two terms

$$G(\boldsymbol{r}, \tau) = G_s(\boldsymbol{r}, \tau) + G_d(\boldsymbol{r}, \tau) \tag{D.9}$$

where, G_s and G_d are the self and distinct terms, respectively. The self part

$$G_s(\mathbf{r}, \tau) = \frac{1}{K} \left\langle \int d\mathbf{r}' \sum_{i=j=1}^{K} \delta[\boldsymbol{r}' + \boldsymbol{r} - \boldsymbol{r}_j(\tau)] \delta[\boldsymbol{r}' - \boldsymbol{r}_i(0)] \right\rangle \tag{D.10}$$

describes the contributions to the correlations from measuring the same particle at both time points, zero, and τ. The distinct part reads

$$G_d(\mathbf{r}, \tau) = \frac{1}{K} \left\langle \int d\mathbf{r}' \sum_{\substack{i=1 \\ j\neq i}}^{K} \delta[\boldsymbol{r}' + \boldsymbol{r} - \boldsymbol{r}_j(\tau)] \delta[\boldsymbol{r}' - \boldsymbol{r}_i(0)] \right\rangle, \tag{D.11}$$

which measures the contributions to the correlations from distinct particles ($j \neq i$).

Extension to two species: cells and bead

To extend the van Hove correlation function to account for a second species, consider that of the K particles, N correspond to cells, and M to beads, such that $K = N + M$. Thus,

$$\hat{G}(\mathbf{r}, t, \tau) = \frac{1}{N} \left\langle \sum_{i=1}^{N} \sum_{j=1}^{M} \delta[\mathbf{r} - \mathbf{r}_j(t + \tau) + R_i(t)] \right\rangle \tag{D.12}$$

is defined, where R_i and $\mathbf{r}_j$ correspond to the positions of cells and beads, respectively. This correlation function, as previously defined in Eq. (4.1), considers only distinct particles of different species. As a result, it possesses no self part. As discussed in Ch. 4, aging is introduced by allowing an explicit dependence on time t of the positions of cells and beads.

D.1.3 Multiple particle tracking

To extract the positional information of cells and beads from the bright-field time-lapse images over time, a custom multi-particle tracking algorithm was implemented implemented (MATLAB R2018b, The MathWorks, Inc., Natick, Massachusetts, United States). This allowed the high-throughput analysis of the interaction of thousands of cells and beads at one minute resolution, and the calculation of the effective potential of interaction, Eq. (4.3). Cells were identified using a Sobel [196] edge-detection algorithm (`sobel` method in MATLAB's `edge` function) and screened according to size to exclude small size objects, e.g. beads and debris, and large objects, e.g. cell doublets and clusters. Beads appear as nearly perfectly black objects, and were detected by applying a Laplacian-of-Gaussian filter (`log` method in MATLAB's `edge` function), and an intensity threshold to identify lower intensity (dark) objects. The detected candidate objects were screened according to eccentricity and size. Frame-to-frame nearest neighbour association was then performed to construct the trajectories of cells and beads. Each trajectory was stored with the corresponding object properties: boundary, area, major and minor axis lengths, and eccentricity.

D.1.4 Experimental effective potential

The age-dependent radial distribution function $\hat{g}(r, t, \tau)$ and effective potential, Eq. (4.3), were computed by making use of the position of the cells and beads obtained from the tracking algorithm (methods Sec. D.1.3) as follows: (i) the distances from the position of each cell at time t, to the position of every beads at time $t + \tau$ were measured, and the density of distances as a function of r was computed by pooling the data obtained for all cell-bead distances, from every repeat (see Fig. D.2a). (ii) A null model, corresponding to the non-interacting system, was constructed by considering a uniform, independently random distribution of beads, and constructing the corresponding density profile, as done in (i), based on the positions of cells obtained experimentally (see Fig. D.2b). The uniform distribution of beads was constructed numerically, and has the advantage of taking into account the finite size of the field of view, nevertheless these finite size effects are noticeable only for large values of r. (iii) The radial distribution function (RDF) $\hat{g}(r, t, \tau)$ (Fig. D.2c) was obtained by computing the ratio between the experimental densities obtained in (i) and the reference densities in (ii). (iv) Finally, the effective potential of interaction was obtained by taking the negative natural logarithm of the RDF, $U^{\text{eff}} = -\ln \hat{g}(r, t, \tau)$ (see schematic Fig. D.2c).

Points of interest in the effective potential

Three points of interest were defined in the effective potentials, (i) The maximum accumulation point (MAP), defined as the minimum of the effective potential, where its spatial derivative vanishes, i.e. $dU^{\text{eff}}/dr = 0$. This point marks the most probable location of the beads along a radial line starting at $r = 0$. This point is indicated by the symbol ▼ in Fig. 4.3. (ii) The maximum depletion point (MDP), defined as the maximum of the potential where $d\text{U}^{\text{eff}}/dr = 0$. It marks the least probable location of beads along the radial coordinate, and is indicated by the symbol ▲ in Fig. 4.3. (iii) The range of interaction, is determined based on the MAP and MDP and is defined as the maximum of $P1$ and $P2$, where $P1$ is the first point beyond the MDP where $U^{\text{eff}}(r, t) = 0$, and $P2$ is the second point after the MAP where $U^{\text{eff}}(r, t) = 0$. This point measures the effective distance up to which cells deplete the environment of beads. The range is indicated by (●) in Fig. 4.3.

D.1.5 Solution to the model with release

The master equation of the probability $P_\kappa(n, t)$ to find n beads in contact with a cell at time t, as beads can be released by the cell at constant rate κ, is given by Eq. (4.13). To solve it, a generating function is introduced

$$\phi_\kappa(z, t) \quad = \quad \sum_{n=0}^{\infty} z^n P_\kappa(n, t), \tag{D.13}$$

with normalization

$$\phi_\kappa(1, t) = \sum_{n=0}^{\infty} P(n, t) = 1. \tag{D.14}$$

Using Eq. (D.13), Eq. (4.13) produces a first order, linear partial differential equation for $\phi_\kappa(z, t)$,

$$\frac{1}{1-z}\frac{\partial \phi_\kappa(z, t)}{\partial t} - \kappa \frac{\partial \phi_\kappa(z, t)}{\partial z} = -\rho_b \frac{dA(t)}{dt} \phi_\kappa(z, t), \tag{D.15}$$

which can be solved by the method of characteristics [197]. Introducing

$$\eta = \eta(z, t; \kappa), \tag{D.16}$$

demanding that $\phi_\kappa = w(\eta(z, t; \kappa), t)$ with a hitherto unknown function w of two variables. Eq. (D.15) can be now written in terms of w,

$$\frac{\partial w}{\partial t} + \frac{\partial w}{\partial \eta}\left(\frac{\partial \eta}{\partial t} - (1-z)\kappa \frac{\partial \eta}{\partial z}\right) + (1-z)\rho_b \dot{A}(t) w = 0. \tag{D.17}$$

Demanding

$$\frac{\partial \eta}{\partial t} - (1-z)\kappa \frac{\partial \eta}{\partial z} = 0. \tag{D.18}$$

implies

$$\eta(z,t;\kappa) = \log(1-z) - \kappa t, \tag{D.19}$$

and Eq. (D.17) reduces to

$$\frac{\partial w}{\partial t} + (1-z)\rho_b \dot{A}(\xi) w = 0, \tag{D.20}$$

which can be directly solved to produce

$$w(\eta,t) = C(\eta)\exp\left(-\int^t \mathrm{d}s\,(1-z)\rho_b\dot{A}(s)\right). \tag{D.21}$$

Noticing that $1-z$ is given in terms of t by Eq. (D.19), then

$$\begin{aligned} w(\eta,t) &= C(\eta)\exp\left(-\rho_b \mathrm{e}^{-\eta}\int^t \mathrm{d}s\,\mathrm{e}^{\kappa s}\dot{A}(s)\right) \\ &= C(\eta)\exp\left(-\rho_b \mathrm{e}^{-\eta}\left[A(t)\mathrm{e}^{\kappa t} - \kappa\int^t \mathrm{d}s\,A(s)\mathrm{e}^{\kappa s}\right]\right), \end{aligned} \tag{D.22}$$

where an integration by part was performed to go from the first to the second line. Writing Eq. (D.22) in terms of z and t, and imposing the normalisation Eq. (D.14), which implies $C = 1$, then

$$\phi_\kappa(z,t) = \exp\left(-\rho_b(1-z)\left[A(t) - \kappa e^{-\kappa t}\int^t \mathrm{d}s\,A(s)e^{\kappa s}\right]\right). \tag{D.23}$$

Finally, the probability $P_\kappa(n,t)$ of finding n beads attached to a cell at time t correspond to the coefficients of nth order term in the Taylor expansion of Eq. (D.23) around $z = 0$. This produces

$$P_\kappa(n,t) = \frac{\alpha^n(A,\kappa,t)}{n!}e^{-\alpha(A,\kappa,t)}, \tag{D.24}$$

where $\alpha(A,\kappa,t) = \rho_b\left[A(t) - \kappa e^{-\kappa t}\int^t \mathrm{d}s\,A(s)e^{\kappa s}\right]$ is the time-dependent rate, as shown in Ch. 4.

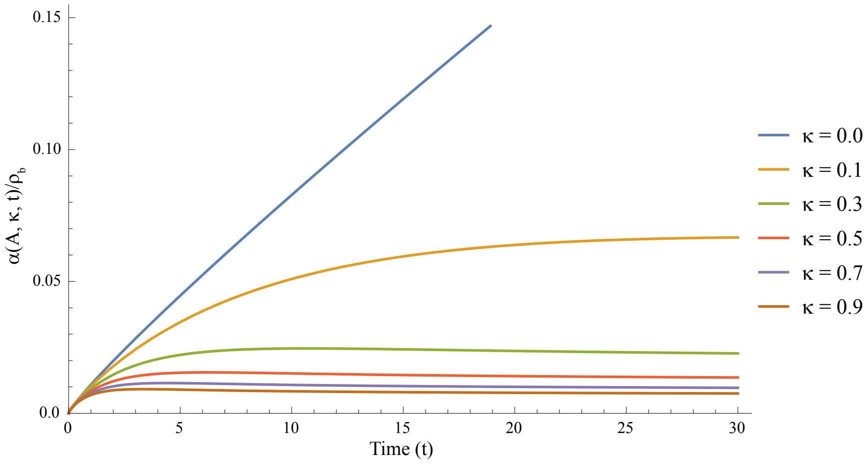


Figure D.1: **Effective rate with release**. The parameter $\alpha(A,\kappa,t)$, given by Eq. (4.15), for different values of the release rate κ. Here $r_0 = 0.02$, and $\delta = 0.004$.

D.2 Experimental measurement of the effective potential

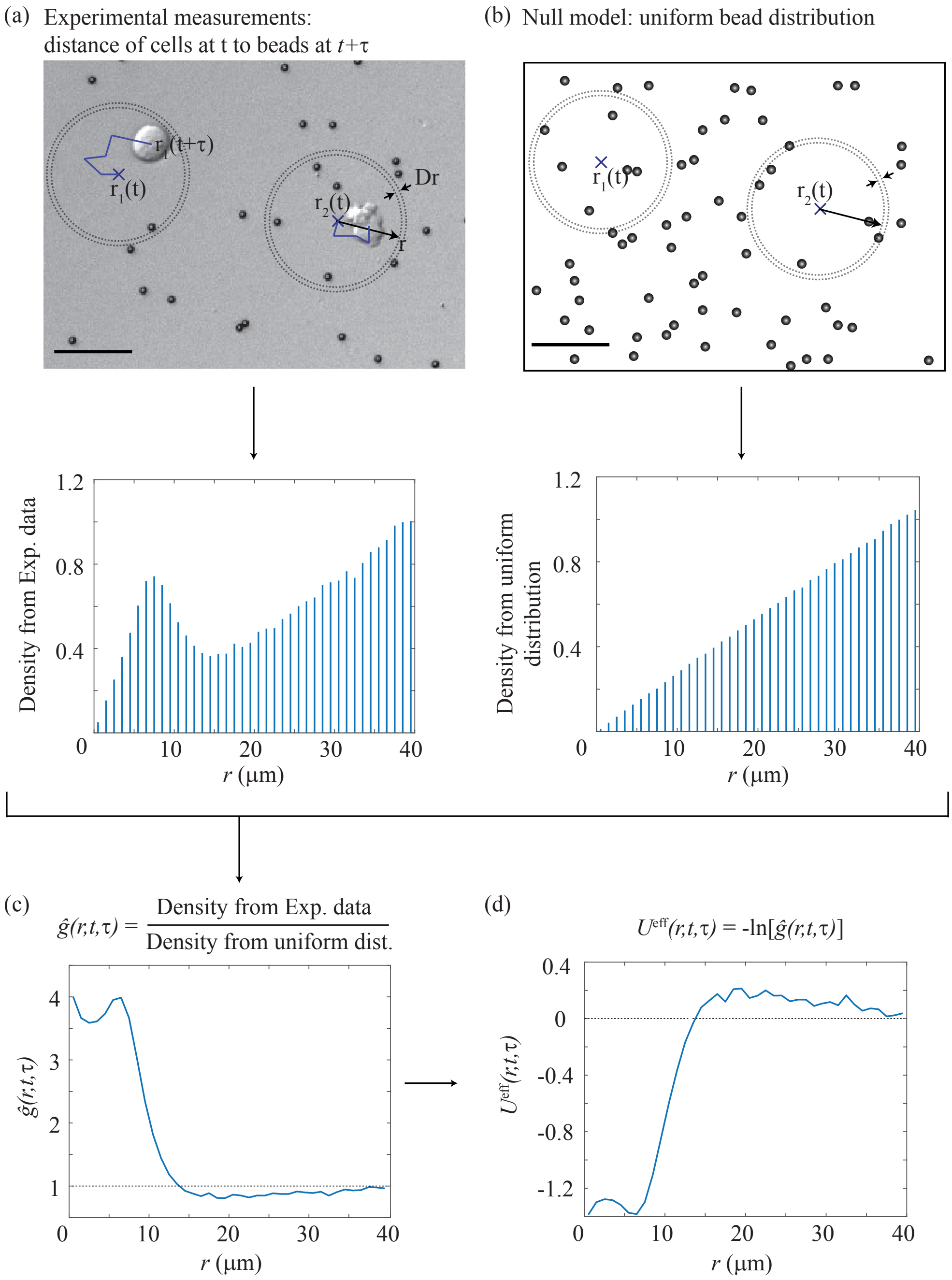


Figure D.2: **Experimental measurement of the effective potential**. Measurement of the distance from the position of every cell at time t, to the position of every beads at time $t + \tau$ (top) and the corresponding density of distances (bottom) for (a) the experimental positions of cells and beads, and (b) the experimental positions of cells and uniformly distributed (numerically drawn) beads. Scale bars 25 μm. (c) Radial distribution function measured as the ratio of the densities in (a) and (b). (d) Effective potential calculated as the negative logarithm of the radial distribution function in (c). Both $\hat{g}$ and U^{eff} are dimensionless.

D.3 Effective potentials as a function of age and delay

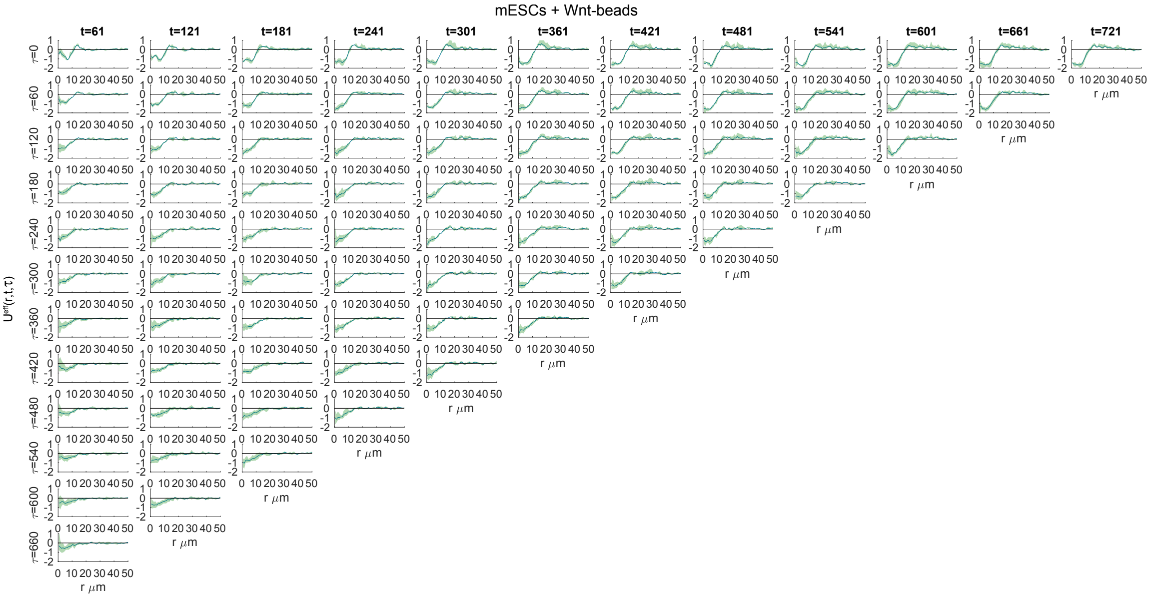


Figure D.3: **Experimental effective potential**, as defined in Eq. (4.3), as function of age t and delay τ for Wnt-beads as mESCs interact with them. The solid curve corresponds to the overall effective potential obtained from pooling the data from three repeats, while the green shaded area corresponds to the range of the measurements from the three independent repeats. Total cell number $\sim 10^3$, total bead number $\sim 10^4$.

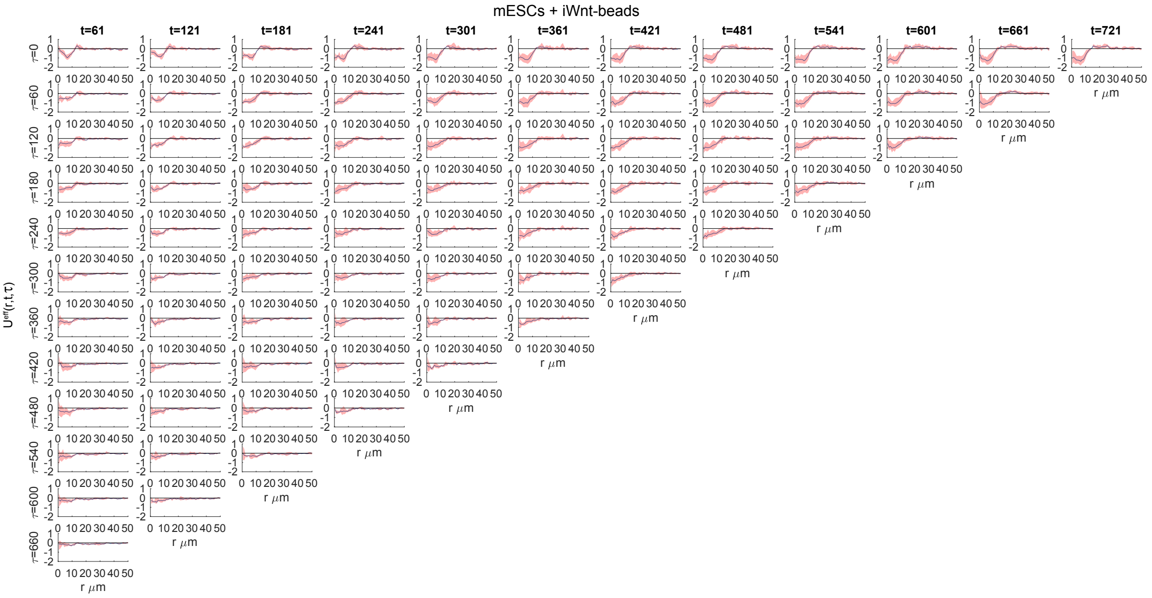


Figure D.4: **Experimental effective potential**, as defined in Eq. (4.3), as function of age t and delay τ for iWnt-beads as mESCs interact with them. The solid curve corresponds to the overall effective potential obtained from pooling the data from three repeats, while the red shaded area corresponds to the range of the measurements from the three independent repeats. Total cell number $\sim 10^3$, total bead number $\sim 10^4$.

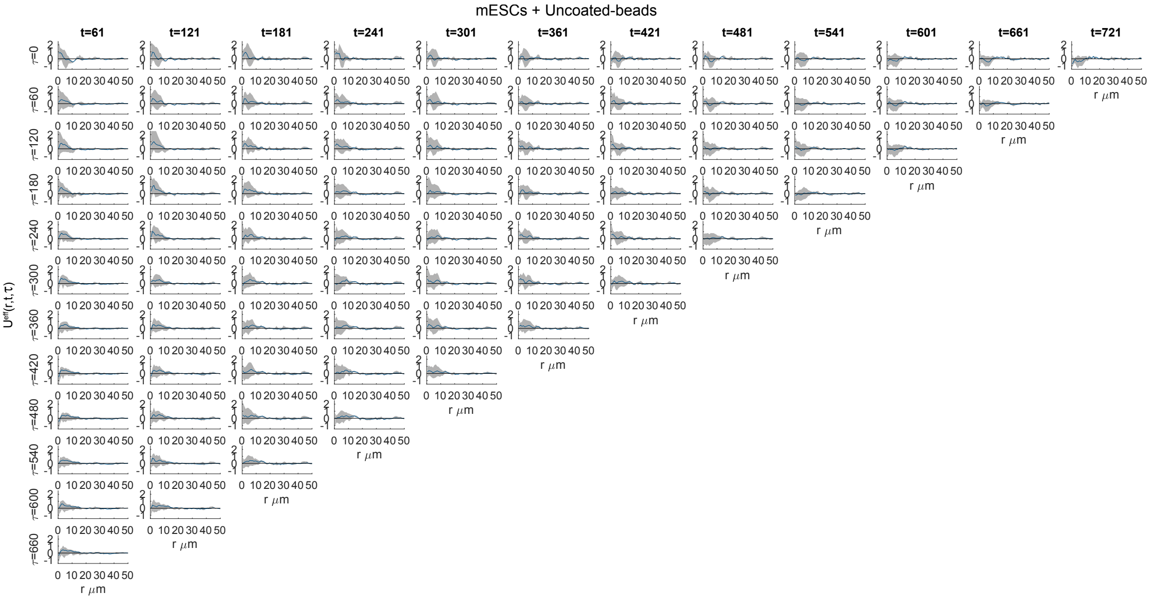


Figure D.5: **Experimental effective potential**, as defined in Eq. (4.3), as function of age *t* and delay τ for uncoated-beads as mESCs interact with them. The solid curve corresponds to the overall effective potential obtained from pooling the data from three repeats, while the black shaded area corresponds to the range of the measurements from the three independent repeats. Total cell number $\sim 10^3$, total bead number $\sim 10^4$.

D.4 Wnt canonical pathway

The Wnt family of proteins has important roles in stem cell fate decision making, and tissue development and maintenance. Deregulation of Wnt-related pathways is associated with cancer and other growth-related pathologies. Wnt3a, a member of the Wnt-family of proteins, plays a key role in cell-polarity and asymmetric cell division via the activation of the Wnt/β-catenin canonical pathway.

The Wnt canonical pathway activation is mediated by cytoplasmic and nuclear β-catenin, a protein responsible for activation of nuclear activation of Wnt target-genes, TCF and LEF. In the absence of environmental Wnt3a (Fig. D.6a), the *destruction complex* [11], a molecular complex inside the cell, tags β-catenin for phosphorilation (degradation) by the proteosome. This prevents β-catenin from being transported to the nucleus and initiating transduction. In this scenario the Wnt/β-catenin canonical pathways is said to be *inactive* (Fig. D.6a). On the other hand, when Wnt3a is present in the extacellular environment, it can bind to its transmembrane co-receptors LRP5/6 and Frizzled (Fzd) [9, 10]. When this happens, the destruction complex is sequestered [12], preventing it from acting on β-catenin. The cytoplasmic β-catenin can then be transported into the nucleus and activate the TCF/LEF transcription factors [13], and promote differentiation (see Fig. D.6b).

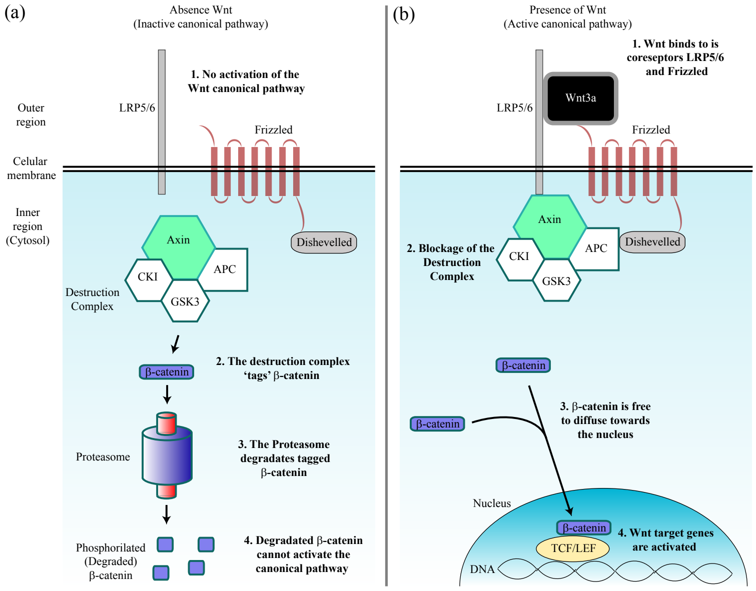
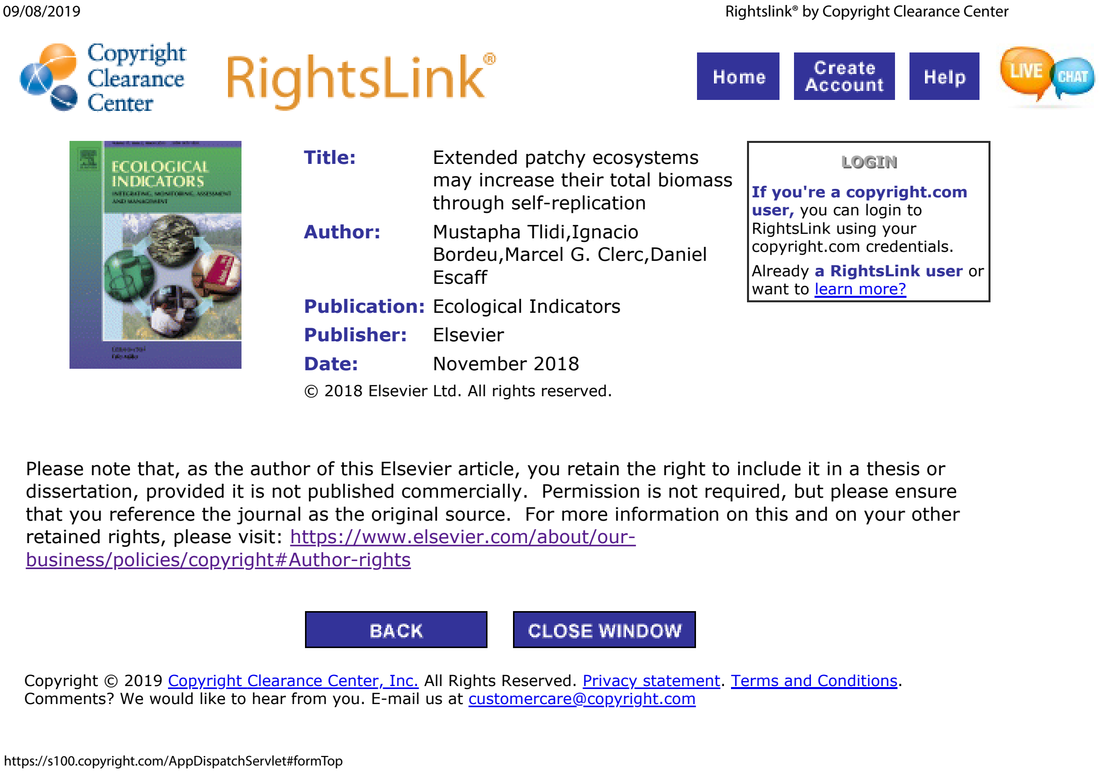


Figure D.6: **Canonical Wnt/β-Catenin pathway.** (a) Shows the unbound Wnt coreceptors LRP5/6 and Frizzled [9, 10], this allows the destruction complex to act on β-catenin, tagging it for degradation (non-activated canonical pathway) [11]. (b) Shows the activated canonical pathway, where by the action of Wnt, the destruction complex is disabled [12], allowing β-catenin to fulfil its transcription role in the nucleus [13].

Appendix E

Copyright permissions



09/08/2019

Rightslink® by Copyright Clearance Center

Title: Extended patchy ecosystems may increase their total biomass through self-replication

Author: Mustapha Tlidi,Ignacio Bordeu,Marcel G. Clerc,Daniel Escaff

Publication: Ecological Indicators

Publisher: Elsevier

Date: November 2018

© 2018 Elsevier Ltd. All rights reserved.

If you're a copyright.com user, you can login to RightsLink using your copyright.com credentials.

Already **a RightsLink user** or want to learn more?

Please note that, as the author of this Elsevier article, you retain the right to include it in a thesis or dissertation, provided it is not published commercially. Permission is not required, but please ensure that you reference the journal as the original source. For more information on this and on your other retained rights, please visit: https://www.elsevier.com/about/our-business/policies/copyright#Author-rights

Copyright © 2019 Copyright Clearance Center, Inc. All Rights Reserved. Privacy statement. Terms and Conditions.
Comments? We would like to hear from you. E-mail us at customercare@copyright.com

https://s100.copyright.com/AppDispatchServlet#formTop

Permissions for the use of figures and text from [16] in Chapter 3.

Bibliography

[1] Fiona M Watt. The reward of great collaborations, 2018.

[2] Ignacio Bordeu, Saoirse Amarteifio, Rosalba Garcia-Millan, Benjamin Walter, Nanxin Wei, and Gunnar Pruessner. Volume explored by a branching random walk on general graphs. *Sci. Rep.*, 9(1):15590, 2019.

[3] Bimal Viswanath, Alan Mislove, Meeyoung Cha, and Krishna P Gummadi. On the evolution of user interaction in facebook. In *Proceedings of the 2nd ACM workshop on Online social networks*, pages 37–42. ACM, August 2009.

[4] Hawoong Jeong, Sean P Mason, A-L Barabási, and Zoltan N Oltvai. Lethality and centrality in protein networks. *Nature*, 411(6833):41–42, 2001.

[5] Lazaros K Gallos, Chaoming Song, Shlomo Havlin, and Hernán A Makse. Scaling theory of transport in complex biological networks. *Proc. Natl. Acad. Sci. USA*, 104(19):7746–7751, 2007.

[6] Mustapha Tlidi, Ignacio Bordeu, Marcel G Clerc, and Daniel Escaff. Extended patchy ecosystems may increase their total biomass through self-replication. *Ecol. Indic.*, 94:534, 2018.

[7] Ignacio Bordeu, Clare Garcin, Shukry J Habib, and Gunnar Pruessner. Effective potentials in cell-protein signal interactions. In review.

[8] Toru Hiratsuka, Ignacio Bordeu, Gunnar Pruessner, and Fiona M Watt. Erk mapk pulses mediate epidermal stem cell fate. In preparation.

[9] Matthew Ling-Hon Chu, Victoria E Ahn, Hee-Jung Choi, Danette L Daniels, Roel Nusse, and William I Weis. Structural studies of wnts and identification of an lrp6 binding site. *Structure*, 21(7):1235, 2013.

[10] Claudia Y Janda, Deepa Waghray, Aron M Levin, Christoph Thomas, and K Christopher Garcia. Structural basis of wnt recognition by frizzled. *Science*, 337(6090):59–64, 2012.

[11] Bonnee Rubinfeld, Iris Albert, Emilio Porfiri, Carol Fiol, et al. Binding of gsk3beta to teh apc-beta-catenin complex and regulation of complex assembly. *Science*, 272(5264):1023, 1996.

[12] Bryan T MacDonald and Xi He. A finger on the pulse of wnt receptor signaling. *Cell Res.*, 22(10):1410, 2012.

[13] Jürgen Behrens, Jens P von Kries, Michael Kühl, Laurakay Bruhn, Doris Wedlich, Rudolf Grosschedl, and Walter Birchmeier. Functional interaction of β-catenin with the transcription factor lef-1. *Nature*, 382(6592):638, 1996.

[14] IPDJ Harris, Philip D Jones, Timothy J Osborn, and David H Lister. Updated high-resolution grids of monthly climatic observations–the cru ts3. 10 dataset. *Int. J. Climatol.*, 34(3):623, 2014.

[15] Anne-Mieke Reijne, Ignacio Bordeu, Gunnar Pruessner, and Giovanni Sena. Linear stability analysis of morphodynamics during tissue regeneration in plants. *J. Phys. D*, 52(8):084002, 2018.

[16] Ignacio Bordeu, Marcel G Clerc, Piere Couteron, René Lefever, and Mustapha Tlidi. Self-replication of localized vegetation patches in scarce environments. *Sci. Rep.*, 6:33703, 2016.

[17] Ignacio Bordeu, MG Clerc, René Lefever, and Mustapha Tlidi. From localized spots to the formation of invaginated labyrinthine structures in a swift–hohenberg model. *Commun. Nonlinear Sci. Numer. Simul.*, 29(1-3):482, 2015.

[18] Ignacio Bordeu and Marcel G Clerc. Rodlike localized structure in isotropic pattern-forming systems. *Phys. Rev. E*, 92(4):042915, 2015.

[19] Ignacio Bordeu, Marcel G Clerc, René Lefever, and Mustapha Tlidi. Finger dynamics in pattern forming systems. In *Nonlinear Dynamics: Materials, Theory and Experiments*, page 275. Springer, 2016.

[20] I Bordeu. Labyrinthine dissipative patterns. In *J. Phys. Conf. Ser.*, volume 720, page 012004, 2016.

[21] Theodore E. Harris. *The Theory of Branching Processes*. Springer-Verlag, Berlin, Germany, 1963.

[22] Kabir Ramola, Satya N. Majumdar, and Grégory Schehr. Spatial extent of branching brownian motion. *Phys. Rev. E*, 91:042131, 2015.

[23] M. Doi. Second quantization representation for classical many-particle system. *J. Phys. A: Math. Gen.*, 9(9):1465, 1976.

[24] L. Peliti. Path integral approach to birth-death processes on a lattice. *J. Phys. (Paris)*, 46:1469, 1985.

[25] Norman TJ Bailey. *The mathematical theory of infectious diseases and its applications*. Charles Griffin & Company Ltd. High Wycombe, Bucks, UK., 2nd edition, 1975.

[26] Zhen Wang, Chris T Bauch, Samit Bhattacharyya, Alberto d'Onofrio, Piero Manfredi, Matjaž Perc, Nicola Perra, Marcel Salathé, and Dawei Zhao. Statistical physics of vaccination. *Phys. Rep.*, 664:1, 2016.

[27] Stephen Eubank, Hasan Guclu, VS Anil Kumar, Madhav V Marathe, Aravind Srinivasan, Zoltan Toroczkai, and Nan Wang. Modelling disease outbreaks in realistic urban social networks. *Nature*, 429(6988):180, 2004.

[28] Romualdo Pastor-Satorras, Claudio Castellano, Piet Van Mieghem, and Alessandro Vespignani. Epidemic processes in complex networks. *Rev. Mod. Phys.*, 87(3):925, 2015.

[29] Quentin Sattentau. Avoiding the void: cell-to-cell spread of human viruses. *Nat. Rev. Microbiol.*, 6(11):815, 2008.

[30] Eric Dumonteil, Satya N. Majumdar, Alberto Rosso, and Andrea Zoia. Spatial extent of an outbreak in animal epidemics. *Proc. Natl. Acad. Sci. USA*, 110(11):4239, 2013.

[31] Stefan Nekovar and Gunnar Pruessner. A field-theoretic approach to the wiener sausage. *J. Stat. Phys.*, 163(3):604, 2016.

[32] A. M. Berezhkovskii, Yu. A. Makhnovskii, and R. A. Suris. Wiener sausage volume moments. *J. Stat. Phys.*, 57(1-2):333–346, 1989.

[33] Paul Elliott, Jon C Wakefield, Nicola G Best, David John Briggs, et al. *Spatial epidemiology: methods and applications*, volume 15. Oxford University Press, Oxford, 2000.

[34] Ronald Aylmer Fisher. The wave of advance of advantageous genes. *Ann. Eugen.*, 7(4):355, 1937.

[35] Kolmogorov A.N, I.G. Petrovsky, and N.S Piscunov. A study of the diffusion equation with increase in the amount of substance, and its application to a biological problem. *Bull. de l'Univ. d'État a Moscou (sér intern)*, 1(6):1–25, 1937.

[36] Maurice S Bartlett. Deterministic and stochastic models for recurrent epidemics. In *Proc. Third Berkeley Symposium on Mathematical Statistics and Probability*, volume IV, pages 81–109. University of California Press, 1965.

[37] Maurice S Bartlett. Measles periodicity and community size. *Journal of the Royal Statistical Society. Series A (General)*, 120(1):48–70, 1957.

[38] J Radcliffe. The convergence of a position-dependent branching process used as an approximation to a model describing the spread of an epidemic. *J. Appl. Probab.*, 13(2):338, 1976.

[39] Nobuyuki Ikeda, Masao Nagasawa, and Shinzo Watanabe. Branching markov processes ii. *J. Math. Kyoto U.*, 8(3):365–410, 1968.

[40] Éric Brunet. *Some aspects of the Fisher-KPP equation and the branching Brownian motion*. PhD thesis, 2016.

[41] Henry William Watson and Francis Galton. On the probability of the extinction of families. *The Journal of the Anthropological Institute of Great Britain and Ireland*, 4:138, 1875.

[42] Stanley Sawyer and Joseph Fleischman. Maximum geographic range of a mutant allele considered as a subtype of a brownian branching random field. *Proc. Natl. Acad. Sci. USA*, 76(2):872, 1979.

[43] Jean-François Le Gall and Shen Lin. The range of tree-indexed random walk in low dimensions. *Ann. Probab.*, 43(5):2701, 2015.

[44] Jean-François Le Gall and Shen Lin. The range of tree-indexed random walk. *J. Inst. Math. Jussieu*, 15(2):271, 2016.

[45] Jacob Grimm and Wilhelm Grimm. *Kinder-und hausmärchen*. JG Cotta Nachf., 1906.

[46] Rosalba Garcia-Millan, Johannes Pausch, Benjamin Walter, and Gunnar Pruessner. Field-theoretic approach to the universality of branching processes. *Phys. Rev. E*, 98(6):062107, 2018.

[47] Uwe C. Täuber, Martin Howard, and Benjamin P. Vollmayr-Lee. Applications of field-theoretic renormalization group methods to reaction-diffusion problems. *J. Phys. A: Math. Gen.*, 38(17):R79, Apr 2005.

[48] Uwe Claus Täuber. *Critical Dynamics A Field Theory Approach to Equilibrium and Non-Equilibrium Scaling Behavior*. Cambridge University Press, Cambridge, England, 2014.

[49] Pierre Pfeuty and Gérard Toulouse. *Introduction to the Renormalization Group and to Critical Phenomena*. John Wiley & Sons, Chichester, West Sussex, UK, 1977.

[50] Raffaella Burioni and Davide Cassi. Random walks on graphs: ideas, techniques and results. *J. Phys. A: Math. Gen.*, 38(8):R45, 2005.

[51] Albert-László Barabási, Réka Albert, and Hawoong Jeong. Scale-free characteristics of random networks: the topology of the world-wide web. *Physica A*, 281(1-4):69, 2000.

[52] Réka Albert. Scale-free networks in cell biology. *J. Cell Sci.*, 118(21):4947, 2005.

[53] Wei Wang, Ming Tang, H Eugene Stanley, and Lidia A Braunstein. Unification of theoretical approaches for epidemic spreading on complex networks. *Rep. Prog. Phys.*, 80(3):036603, 2017.

[54] Boming Yu and Jianhua Li. Some fractal characters of porous media. *Fractals*, 9(03):365–372, 2001.

[55] Russell Lyons. Random walks and percolation on trees. *Ann. Probab.*, pages 931–958, 1990.

[56] Hiroshi Watanabe. Spectral dimension of a wire network. *J. Phys. A: Math. Gen.*, 18(14):2807–2823, 1985.

[57] Claudio Destri and Luca Donetti. The spectral dimension of random trees. *J. Phys. A: Math. Gen.*, 35(45):9499, 2002.

[58] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *Science*, 286:509–512, 1999.

[59] Albert-László Barabási, Réka Albert, and Hawoong Jeong. Scale-free characteristics of random networks: the topology of the world-wide web. *Physica A*, 281(1-4):69–77, 2000.

[60] Roger Guimera, Stefano Mossa, Adrian Turtschi, and LA Nunes Amaral. The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles. *Proc. Natl. Acad. Sci. USA*, 102(22):7794–7799, 2005.

[61] Hawoong Jeong, Bálint Tombor, Réka Albert, Zoltan N Oltvai, and A-L Barabási. The large-scale organization of metabolic networks. *Nature*, 407(6804):651, 2000.

[62] A. N. Samukhin, S. N. Dorogovtsev, and J. F. F. Mendes. Laplacian spectra of, and random walks on, complex networks: Are scale-free architectures really important? *Phys. Rev. E*, 77(3):036115, 2008.

[63] Naoki Masuda, Mason A Porter, and Renaud Lambiotte. Random walks and diffusion on networks. *Phys. Rep.*, 716:1–58, 2017.

[64] Jing-Dong J Han, Denis Dupuy, Nicolas Bertin, Michael E Cusick, and Marc Vidal. Effect of sampling on topology predictions of protein-protein interaction networks. *Nat. Biotechnol.*, 23(7):839–844, 2005.

[65] Michael PH Stumpf, Carsten Wiuf, and Robert M May. Subnets of scale-free networks are not scale-free: sampling properties of networks. *Proc. Natl. Acad. Sci. USA*, 102(12):4221–4224, 2005.

[66] David Aldous. The continuum random tree iii. *Ann. Probab.*, 21(1):248–289, 1993.

[67] Jean-François Le Gall. Random trees and applications. *Probab. Surv.*, 2:245–311, 2005.

[68] Ingve Simonsen, Kasper Astrup Eriksen, Sergei Maslov, and Kim Sneppen. Diffusion on complex networks: a way to probe their large-scale topological structures. *Physica A*, 336(1-2):163–173, 2004.

[69] Donald R Mills, RL Peterson, and Sol Spiegelman. An extracellular darwinian experiment with a self-duplicating nucleic acid molecule. *Proc. Natl. Acad. Sci. USA*, 58(1):217, 1967.

[70] E Meron. *Nonlinear Physics of Ecosystems*. CRC Press, Boca Raton, Florida, 2015.

[71] Alan Mathison Turing. The chemical basis of morphogenesis. *Philos. Trans. R. Soc. B*, 237(641):37, 1952.

[72] Mark C Cross and Pierre C Hohenberg. Pattern formation outside of equilibrium. *Rev. Mod. Phys.*, 65(3):851, 1993.

[73] Shigeru Kondo and Rihito Asai. A reaction–diffusion wave on the skin of the marine angelfish pomacanthus. *Nature*, 376(6543):765, 1995.

[74] Philip Ball and Neil R Borley. *The self-made tapestry: pattern formation in nature*, volume 198. Oxford University Press, Oxford, 1999.

[75] N Verschueren, U Bortolozzo, MG Clerc, and S Residori. Chaoticon: localized pattern with permanent dynamics. *Philos. Trans. R. Soc. A*, 372(2027):20140011, 2014.

[76] René Lefever and Olivier Lejeune. On the origin of tiger bush. *Bull. Math. Biol.*, 59(2):263, 1997.

[77] Mustapha Tlidi, René Lefever, and Andrei Vladimirov. On vegetation clustering, localized bare soil spots and fairy circles. In *Dissipative Solitons: From Optics to Biology and Medicine. Lecture Notes in Physics*, volume 751. Springer, Berlin, Heidelberg, 2008.

[78] Karna Gowda, Hermann Riecke, and Mary Silber. Transitions between patterned states in vegetation models for semiarid ecosystems. *Phys. Rev. E*, 89(2):022701, 2014.

[79] Christopher A Klausmeier. Regular and irregular patterns in semiarid vegetation. *Science*, 284(5421):1826, 1999.

[80] Reinier HilleRisLambers, Max Rietkerk, Frank van den Bosch, Herbert HT Prins, and Hans de Kroon. Vegetation pattern formation in semi-arid grazing systems. *Ecology*, 82(1):50, 2001.

[81] Jost von Hardenberg, Ehud Meron, Moshe Shachak, and Yair Zarmi. Diversity of vegetation patterns and desertification. *Phys. Rev. Lett.*, 87(19):198101, 2001.

[82] Riccardo Vesipa, Carlo Camporeale, and Luca Ridolfi. Noise-driven cooperative dynamics between vegetation and topography in riparian zones. *GeoResLett*, 42(19):8021, 2015.

[83] Sonia Kéfi, Max Rietkerk, Concepción L Alados, Yolanda Pueyo, Vasilios P Papanastasis, Ahmed ElAich, and Peter C De Ruiter. Spatial vegetation patterns and imminent desertification in mediterranean arid ecosystems. *Nature*, 449(7159):213, 2007.

[84] Sonia Kéfi, Max Rietkerk, Minus Van Baalen, and Michel Loreau. Local facilitation, bistability and transitions in arid ecosystems. *TheoPopBio*, 71(3):367, 2007.

[85] Vasilis Dakos, Sonia Kéfi, Max Rietkerk, Egbert H Van Nes, and Marten Scheffer. Slowing down in spatially patterned ecosystems at the brink of collapse. *Am. Nat.*, 177(6):E153, 2011.

[86] Vincent Deblauwe, Pierre Couteron, Jan Bogaert, and Nicolas Barbier. Determinants and dynamics of banded vegetation pattern migration in arid climates. *Ecol. Monogr.*, 82(1):3, 2012.

[87] Nicolas Barbier, Pierre Couteron, Jean Lejoly, Vincent Deblauwe, and Olivier Lejeune. Self-organized vegetation patterning as a fingerprint of climate and human impact on semi-arid ecosystems. *J. Ecol.*, 94(3):537, 2006.

[88] Nicolas Barbier, Pierre Couteron, René Lefever, Vincent Deblauwe, and Olivier Lejeune. Spatial decoupling of facilitation and competition at the origin of gapped vegetation patterns. *Ecology*, 89(6):1521, 2008.

[89] René Lefever, Nicolas Barbier, Pierre Couteron, and Olivier Lejeune. Deeply gapped vegetation patterns: on crown/root allometry, criticality and desertification. *JTheoBio*, 261(2):194, 2009.

[90] Pierre Couteron, Fabien Anthelme, Marcel Clerc, Daniel Escaff, Cristian Fernandez-Oto, and Mustapha Tlidi. Plant clonal morphologies and spatial patterns as self-organized responses to resource-limited environments. *Philos. Trans. R. Soc. A*, 372(2027):20140102, 2014.

[91] John E Pearson. Complex patterns in a simple system. *Science*, 261(5118):189, 1993.

[92] Kyoung-Jin Lee, William D McCormick, John E Pearson, and Harry L Swinney. Experimental observation of self-replicating spots in a reaction–diffusion system. *Nature*, 369(6477):215, 1994.

[93] Akiko Kaminaga, Vladimir K Vanag, and Irving R Epstein. “black spots” in a surfactant-rich belousov–zhabotinsky reaction dispersed in a water-in-oil microemulsion system. *J. Chem. Phys.*, 122(17):174706, 2005.

[94] Theodore Kolokolnikov and Mustapha Tlidi. Spot deformation and replication in the two-dimensional belousov-zhabotinski reaction in a water-in-oil microemulsion. *Phys. Rev. Lett.*, 98(18):188303, 2007.

[95] Mustapha Tlidi, AG Vladimirov, and Paul Mandel. Curvature instability in passive diffractive resonators. *Phys. Rev. Lett.*, 89(23):233901, 2002.

[96] Ehud Meron, Erez Gilad, Jost von Hardenberg, Moshe Shachak, and Yair Zarmi. Vegetation patterns along a rainfall gradient. *Chaos Solitons Fractals*, 19(2):367, 2004.

[97] AG Vladimirov, René Lefever, and Mustapha Tlidi. Relative stability of multipeak localized patterns of cavity solitons. *Phys. Rev. A*, 84(4):043848, 2011.

[98] René Lefever and John W Turner. A quantitative theory of vegetation patterns based on plant structure and the non-local f–kpp equation. *Comptes Rendus Mécanique*, 340(11-12):818, 2012.

[99] Rob W Brooker, Fernando T Maestre, Ragan M Callaway, Christopher L Lortie, Lohengrin A Cavieres, Georges Kunstler, Pierre Liancourt, Katja Tielbörger, Justin MJ Travis, Fabien Anthelme, C Armas, L Coll, E Corcket, S Delzon, E Forey, Z Kikvidze, J Olofsson, F Pugnaire, C Quiroz, P Saccone, K Schiffers, M Seifan, B Touzard, and R Michalet. Facilitation in plant communities: the past, the present, and the future. *J. Ecol.*, 96(1):18, 2008.

[100] H Jochen Schenk and Robert B Jackson. Rooting depths, lateral root spreads and below-ground/above-ground allometries of plants in water-limited ecosystems. *J. Ecol.*, 90(3):480, 2002.

[101] Ilya Prigogine and Réné Lefever. Symmetry breaking instabilities in dissipative systems. ii. *J. Chem. Phys.*, 48(4):1695, 1968.

[102] Olivier Lejeune and Mustapha Tlidi. A model for the explanation of vegetation stripes (tiger bush). *J. Veg. Sci.*, 10(2):201, 1999.

[103] Olivier Lejeune, Mustapha Tlidi, and René Lefever. Vegetation spots and stripes: dissipative structures in arid landscapes. *Int. J. Quantum Chem.*, 98(2):261, 2004.

[104] Max Rietkerk and Johan Van de Koppel. Regular pattern formation in real ecosystems. *Trends Ecol. Evol.*, 23(3):169, 2008.

[105] Jonathan A Sherratt. History-dependent patterns of whole ecosystems. *Ecol. Complex.*, 14:8, 2013.

[106] O Lejeune, M Tlidi, and P Couteron. Localized vegetation patches: a self-organized response to resource scarcity. *Phys. Rev. E*, 66(1):010901, 2002.

[107] Max Rietkerk, Stefan C Dekker, Peter C De Ruiter, and Johan van de Koppel. Self-organized patchiness and catastrophic shifts in ecosystems. *Science*, 305(5692):1926, 2004.

[108] Yves Pomeau. Front motion, metastability and subcritical bifurcations in hydrodynamics. *Physica D*, 23(1-3):3–11, 1986.

[109] Mustapha Tlidi, Paul Mandel, and René Lefever. Localized structures and localized patterns in optical bistability. *Phys. Rev. Lett.*, 73(5):640, 1994.

[110] Alan R Champneys. Homoclinic orbits in reversible systems and their applications in mechanics, fluids and optics. *Physica D*, 112(1-2):158, 1998.

[111] Pierre Coullet, C Riera, and Charles Tresser. Stable static localized structures in one dimension. *Phys. Rev. Lett.*, 84(14):3069, 2000.

[112] Mustapha Tlidi, Kęstutis Staliunas, Krassimir Panajotov, AG Vladimirov, and MG Clerc. Localized structures in dissipative media: from optics to plant ecology. *Philos. Trans. R. Soc. A*, 372:20140101, 2014.

[113] E Knobloch. Spatial localization in dissipative systems. *Annu. Rev. Condens. Matter Phys.*, 6(1):325, 2015.

[114] N Barbier. Interactions spatiales et auto-organisation des végétations semi-arides. *These de doctorat, Université libre de Bruxelles, Belgique*, 2006.

[115] Matt Thomson, Siyuan John Liu, Ling-Nan Zou, Zack Smith, Alexander Meissner, and Sharad Ramanathan. Pluripotency factors in embryonic stem cells regulate differentiation into germ layers. *Cell*, 145(6):875, 2011.

[116] Adam J Engler, Shamik Sen, H Lee Sweeney, and Dennis E Discher. Matrix elasticity directs stem cell lineage specification. *Cell*, 126(4):677, 2006.

[117] Myriam Zecca, Konrad Basler, and Gary Struhl. Direct and long-range action of a wingless morphogen gradient. *Cell*, 87(5):833, 1996.

[118] Shigeru Kondo and Takashi Miura. Reaction-diffusion model as a framework for understanding biological pattern formation. *Science*, 329(5999):1616, 2010.

[119] Patrick Müller, Katherine W Rogers, R Yu Shuizi, Michael Brand, and Alexander F Schier. Morphogen transport. *Development*, 140(8):1621, 2013.

[120] Cyrille Alexandre, Alberto Baena-Lopez, and Jean-Paul Vincent. Patterning and growth control by membrane-tethered wingless. *Nature*, 505(7482):180, 2014.

[121] Henner F Farin, Ingrid Jordens, Mohammed H Mosa, Onur Basak, Jeroen Korving, Daniele VF Tauriello, Karin de Punder, Stephane Angers, Peter J Peters, Madelon M Maurice, et al. Visualization of a short-range wnt gradient in the intestinal stem-cell niche. *Nature*, 530(7590):340, 2016.

[122] Roel Nusse and Hans Clevers. Wnt/β-catenin signaling, disease, and emerging therapeutic modalities. *Cell*, 169(6):985, 2017.

[123] M Cristina Marchetti, Jean-François Joanny, Sriram Ramaswamy, Tanniemola B Liverpool, Jacques Prost, Madan Rao, and R Aditi Simha. Hydrodynamics of soft active matter. *Rev. Mod. Phys.*, 85(3):1143, 2013.

[124] Umberto Marini Bettolo Marconi, Matteo Paoluzzi, and Claudio Maggi. Effective potential method for active particles. *Mol. Phys.*, 114(16–17):2400, 2016.

[125] Thomas E Angelini, Edouard Hannezo, Xavier Trepat, Manuel Marquez, Jeffrey J Fredberg, and David A Weitz. Glass-like dynamics of collective cell migration. *Proc. Natl. Acad. Sci. USA*, 108(12):4714, 2011.

[126] Kyogo Kawaguchi, Ryoichiro Kageyama, and Masaki Sano. Topological defects control collective dynamics in neural progenitor cell cultures. *Nature*, 545(7654):327, 2017.

[127] Simon Garcia, Edouard Hannezo, Jens Elgeti, Jean-François Joanny, Pascal Silberzan, and Nir S Gov. Physics of active jamming during collective cellular motion in a monolayer. *Proc. Natl. Acad. Sci. USA*, 112(50):15314, 2015.

[128] Duyu Chen, Wen Yih Aw, Danelle Devenport, and Salvatore Torquato. Structural characterization and statistical-mechanical model of epidermal patterns. *Biophys. J.*, 111(11):2534, 2016.

[129] Shukry J Habib, Bi-Chang Chen, Feng-Chiao Tsai, Konstantinos Anastassiadis, Tobias Meyer, Eric Betzig, and Roel Nusse. A localized wnt signal orients asymmetric stem cell division in vitro. *Science*, 339(6126):1445, 2013.

[130] Molly Lowndes, Sergi Junyent, and Shukry J Habib. Constructing cellular niche properties by localized presentation of wnt proteins on synthetic surfaces. *Nat. Protoc.*, 12(7):1498, 2017.

[131] Derk Ten Berge, Dorota Kurek, Tim Blauwkamp, Wouter Koole, Alex Maas, Elif Eroglu, Ronald K Siu, and Roel Nusse. Embryonic stem cells require wnt proteins to prevent differentiation to epiblast stem cells. *Nat. Cell Biol.*, 13(9):1070, 2011.

[132] Bryan T MacDonald, Annie Hien, Xinjun Zhang, Oladoyin Iranloye, David M Virshup, Marian L Waterman, and Xi He. Disulfide bond requirements for active wnt ligands. *Journal of Biological Chemistry*, 289(26):18122, 2014.

[133] Gavin E Crooks. Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Phys. Rev. E*, 60(3):2721, 1999.

[134] Udo Seifert. Entropy production along a stochastic trajectory and an integral fluctuation theorem. *Phys. Rev. Lett.*, 95(4):040602, 2005.

[135] Thomas Speck, Valentin Blickle, Clemens Bechinger, and Udo Seifert. Distribution of entropy production for a colloidal particle in a nonequilibrium steady state. *Europhys. Lett.*, 79(3):30002, 2007.

[136] Klebert Feitosa and Narayanan Menon. Fluidized granular medium as an instance of the fluctuation theorem. *Phys. Rev. Lett.*, 92(16):164301, 2004.

[137] Bart Cleuren, Christian Van den Broeck, and R Kawai. Fluctuation theorem for the effusion of an ideal gas. *Phys. Rev. E*, 74(2):021117, 2006.

[138] Daniel S Seara, Vikrant Yadav, Ian Linsmeier, A Pasha Tabatabai, Patrick W Oakes, SM Ali Tabei, Shiladitya Banerjee, and Michael P Murrell. Entropy production rate is maximized in non-contractile actomyosin. *Nat. Commun.*, 9(1):4948, 2018.

[139] Cesare Nardini, Étienne Fodor, Elsen Tjhung, Frédéric Van Wijland, Julien Tailleur, and Michael E Cates. Entropy production in field theories without time-reversal symmetry: quantifying the non-equilibrium character of active matter. *Phys. Rev. X*, 7(2):021007, 2017.

[140] Regina Raz, Chien-Kuo Lee, Linda A Cannizzaro, Peter d'Eustachio, and David E Levy. Essential role of stat3 for embryonic stem cell pluripotency. *Proc. Natl. Acad. Sci. USA*, 96(6):2846, 1999.

[141] Jean-Pierre Hansen and Ian Ranald McDonald. *Theory of simple liquids: with applications to soft matter*. Academic Press, 2013.

[142] Malte Henkel and Michel Pleimling. *Non-Equilibrium Phase Transitions: Volume 2: Ageing and Dynamical Scaling Far from Equilibrium*. Springer, 2011.

[143] Arieh Ben-Naim. Statistical potentials extracted from protein structures: are these meaningful potentials? *The Journal of Chemical Physics*, 107(9):3698, 1997.

[144] Thomas Hamelryck, Mikael Borg, Martin Paluszewski, Jonas Paulsen, Jes Frellsen, Christian Andreetta, Wouter Boomsma, Sandro Bottaro, and Jesper Ferkinghoff-Borg. Potentials of mean force for protein structure prediction vindicated, formalized and generalized. *PloS one*, 5(11):e13714, 2010.

[145] Ulrich Hersel, Claudia Dahmen, and Horst Kessler. Rgd modified polymers: biomaterials for stimulated cell adhesion and beyond. *Biomaterials*, 24(24):4385, 2003.

[146] Andre E Nel, Lutz Mädler, Darrell Velegol, Tian Xia, Eric MV Hoek, Ponisseril Somasundaran, Fred Klaessig, Vince Castranova, and Mike Thompson. Understanding biophysicochemical interactions at the nano–bio interface. *Nat. Mater.*, 8(7):543, 2009.

[147] Tania Patiño, Jorge Soriano, Lleonard Barrios, Elena Ibáñez, and Carme Nogués. Surface modification of microparticles causes differential uptake responses in normal and tumoral human breast epithelial cells. *Sci. Rep.*, 5:11371, 2015.

[148] David Chandler. *Introduction to Modern Statistical Mechanics*. Oxford University Press, 1987.

[149] AM Berezhkovskii, Yu A Makhnovskii, and RA Suris. Wiener sausage volume moments. *J. Stat. Phys.*, 57(1-2):333, 1989.

[150] Stefan Nekovar and Gunnar Pruessner. A field-theoretic approach to the wiener sausage. *J. Stat. Phys.*, 163(3):604, 2016.

[151] Marie Chupeau, Olivier Bénichou, and Raphaël Voituriez. Cover times of random searches. *Nat. Phys.*, 11(10):844, 2015.

[152] Harold J Raveché. Entropy and molecular correlation functions in open systems. i. derivation. *J. Chem. Phys.*, 55(5):2242, 1971.

[153] T Kunath, M.K Saba-El-Leil, M Almousailleakh, J Wray, S Meloche, , and A. Smith. Fgf stimulation of the erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. *Development*, 134(16):2895, 2007.

[154] A Mishra, B Oules, A.O Pisco, T Ly, K Liakath-Ali, G Walko, P Viswanathan, M Tihy, J Nijjher, S.J Dunn, AI Lamond, and FM Watt. A protein phosphatase network controls the temporal and spatial dynamics of differentiation commitment in human epidermis. *Elife*, 6:e27356, 2017.

[155] M.L Dequeant, E Glynn, K Gaudenz, M Wahl, J Chen, A Mushegian, , and O. Pourquie. A complex oscillating network of signaling genes underlies the mouse segmentation clock. *Science*, 314(58-5):1595, 2006.

[156] John J Tyson, Katherine C Chen, and Bela Novak. Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. *Curr. Opin. Cell Biol.*, 15(2):221, 2003.

[157] Boris N Kholodenko. Cell-signalling dynamics in time and space. *Nat. Rev. Mol. Cell Biol.*, 7(3):165, 2006.

[158] J.H Levine, Y Lin, , and M.B. Elowitz. Functional roles of pulsing in genetic circuits. *Science*, 342(6163):1193, 2013.

[159] J.E Purvis, , and G. Lahav. Encoding and decoding cellular information through signaling dynamics. *Cell*, 152(5):945, 2013.

[160] G Lahav, N Rosenfeld, A Sigal, N Geva-Zatorsky, A.J Levine, M.B Elowitz, , and U. Alon. Dynamics of the p53-mdm2 feedback loop in individual cells. *Nat. Genet.*, 36(2):147, 2004.

[161] JG Albeck, GB Mills, and JS Brugge. Frequency-modulated pulses of erk activity transmit quantitative proliferation signals. *Mol. Cell*, 49(2):249, 2013.

[162] M Cargnello, , and P.P. Roux. Activation and function of the mapks and their substrates, the mapk-activated protein kinases. *Microbiol. Mol. Biol. Rev.*, 75(1):50, 2011.

[163] J.W. Ramos. The regulation of extracellular signal-regulated kinase (erk) in mammalian cells. *Int. J. Biochem. Cell Biol.*, 40(12):2707, 2008.

[164] K Aoki, Y Kumagai, A Sakurai, N Komatsu, Y Fujita, C Shionyu, and M Matsuda. Stochastic erk activation induced by noise and cell-to-cell propagation regulates cell density-dependent proliferation. *Mol. Cell*, 52(4):529, 2013a.

[165] T Hiratsuka, Y Fujita, H Naoki, K Aoki, Y Kamioka, , and M. Matsuda. Intercellular propagation of extracellular signal-regulated kinase activation revealed by in vivo imaging of mouse skin. *Elife*, 4:e05178, 2015.

[166] H.E Johnson, Y Goyal, N.L Pannucci, T Schupbach, S.Y Shvartsman, , and J.E. Toettcher. The spatiotemporal limits of developmental erk signaling. *Dev. Cell*, 40(2):185, 2017.

[167] Y Ogura, F.L Wen, M.M Sami, T Shibata, , and S. Hayashi. A switch-like activation relay of egfr-erk signaling regulates a wave of cellular contractility for epithelial invagination. *Dev. Cell*, 46(2):162, 2018.

[168] Ingo Haase, Robin M Hobbs, M Rosario Romero, Simon Broad, and Fiona M Watt. A role for mitogen-activated protein kinase activation by integrins in the pathogenesis of psoriasis. *J. Clin. Invest.*, 108(4):527, 2001.

[169] B Bailleul, M.A Surani, S White, S.C Barton, K Brown, M Blessing, J Jorcano, , and A. Balmain. Skin hyperkeratosis and papilloma formation in transgenic mice expressing a ras oncogene from a suprabasal keratin promoter. *Cell*, 62(4):697, 1990.

[170] R.M Hobbs, V Silva-Vargas, R Groves, , and F.M. Watt. Expression of activated mek1 in differentiating epidermal cells is sufficient to generate hyperproliferative and inflammatory skin lesions. *J. Invest. Dermatol.*, 123(3):503, 2004.

[171] E Hoste, E.N Arwert, R Lal, A.P South, J.C Salas-Alanis, D.F Murrell, G Donati, , and F.M. Watt. Innate sensing of microbial products promotes wound-induced skin cancer. *Nat. Commun.*, 6:5932, 2015.

[172] F.A Scholl, P.A Dumesic, , and P.A. Khavari. Mek1 alters epidermal growth and differentiation. *Cancer Res.*, 64(17):6035, 2004.

[173] M Sibilia, , and E.F. Wagner. Strain-dependent epithelial defects in mice lacking the egf receptor. *Science*, 269(5221):234, 1995.

[174] Boyan Yordanov, Sara-Jane Dunn, Hillel Kugler, Austin Smith, Graziano Martello, and Stephen Emmott. A method to identify and analyze biological programs through automated reasoning. *NPJ Syst. Biol. Appl.*, 2:16010, 2016.

[175] N Komatsu, K Aoki, M Yamada, H Yukinaga, Y Fujita, Y Kamioka, , and M. Matsuda. Development of an optimized backbone of fret biosensors for kinases and gtpases. *Mol. Biol. Cell*, 22(23):4647, 2011.

[176] H. Green. The birth of therapy with cultured cells. *Bioessays*, 30(9):897, 2008.

[177] J.M Carroll, , and L.B. Taichman. Characterization of the human involucrin promoter using a transient beta-galactosidase assay. *J. Cell. Sci.*, 103(4):925, 1992.

[178] P.H Jones, , and F.M. Watt. Separation of human epidermal stem cells from transit amplifying cells on the basis of differences in integrin function and expression. *Cell*, 73(4):713, 1993.

[179] K Aoki, K Takahashi, K Kaizu, , and M. Matsuda. A quantitative model of erk map kinase phosphorylation in crowded media. *Sci. Rep.*, 3:1541, 2013b.

[180] T Matsuyama, Y Izumi, , and T. Sueda. Culture and characterization of human junctional epithelial cells. *J. Periodontol.*, 68(3):229, 1997.

[181] A.J Zhu, I Haase, , and F.M. Watt. Signaling via beta1 integrins and mitogen-activated protein kinase determines human epidermal stem cell fate in vitro. *Proc. Natl. Acad. Sci. USA*, 96(12):6728, 1999.

[182] F.M Watt, , and H. Green. Stratification and terminal differentiation of cultured epidermal cells. *Nature*, 295(5848):434, 1982.

[183] J Younus, , and B.A. Gilchrest. Modulation of mrna levels during human keratinocyte differentiation. *J. Cell Physiol.*, 152(2):232, 1992.

[184] A Sharma, C.T Luke, N.A Dower, J.C Stone, , and P.S. Lorenzo. Rasgrp1 is essential for ras activation by the tumor promoter 12-o-tetradecanoylphorbol-13-acetate in epidermal keratinocytes. *J. Biol. Chem.*, 285(21):15724, 2010.

[185] M Blumenberg. Profiling and metaanalysis of epidermal keratinocytes responses to epidermal growth factor. *BMC Genomics*, 14(1):85, 2013.

[186] Q.L Ying, J Wray, J Nichols, L Batlle-Morera, B Doble, J Woodgett, P Cohen, , and A. Smith. The ground state of embryonic stem cell self-renewal. *Nature*, 453(7194):519, 2008.

[187] John Cardy. Reaction-diffusion processes. In Sergey Nazarenko and Oleg V. Zaboronski, editors, *Non-equilibrium Statistical Mechanics and Turbulence*, pages 108–161. Cambridge University Press, Cambridge, UK, 2008. London Mathematical Society Lecture Note Series: 355, preprint available from `http://www-thphys.physics.ox.ac.uk/people/JohnCardy/warwick.pdf`.

[188] Gunnar Pruessner. Lecture notes on non-equilibrium statistical me- chanics. *(not published)*, 2011.

[189] H. W. Diehl and Felix M. Schmidt. The critical casimir effect in films for generic non-symmetry-breaking boundary conditions. *New J. Phys.*, 13(12):123025, 2011.

[190] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical Recipes in C*. Cambridge University Press, New York, NY, USA, 2nd edition, 1992.

[191] Ruma Dasgupta, T. K. Ballabh, and S Tarafdar. Scaling exponents for random walks on sierpinski carpets and number of distinct sites visited: a new algorithm for infinite fractal lattices. *J. Phys. A: Math. Gen.*, 32(37):6503–6516, 1999.

[192] Claudio Destri and Luca Donetti. On the growth of bounded trees. *J. Phys. A: Math. Gen.*, 35(25):5147, 2002.

[193] Pierre Couteron and Olivier Lejeune. Periodic spotted patterns in semi-arid vegetation explained by a propagation-inhibition model. *J. Ecol.*, 89(4):616, 2001.

[194] Vincent Deblauwe, Pierre Couteron, Olivier Lejeune, Jan Bogaert, and Nicolas Barbier. Environmental modulation of self-organized periodic vegetation patterns in sudan. *Ecography*, 34(6):990, 2011.

[195] Eldad Kepten, Aleksander Weron, Grzegorz Sikora, Krzysztof Burnecki, and Yuval Garini. Guidelines for the fitting of anomalous diffusion mean square displacement graphs from single particle tracking experiments. *PLoS One*, 10(2):e0117722, 2015.

[196] OR Vincent and Olusegun Folorunso. A descriptive algorithm for sobel image edge detection. In *Proceedings of Informing Science & IT Education Conference (InSITE)*, volume 40, pages 97–107, 2009.

[197] Richard Courant and David Hilbert. *Methods of Mathematical Physics: Partial Differential Equations*, volume II. John Wiley & Sons, New York, 1962.