THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Statistical inference on interacting particle systems with applications to cancer biology

Gustav Lindwall

CHALMERS
UNIVERSITY OF TECHNOLOGY

UNIVERSITY OF GOTHENBURG

Division of Mathematical Statistics
Department of Mathematical Sciences
Chalmers University of Technology and University of Gothenburg
Göteborg, Sweden 2023

Statistical inference on interacting particle systems with applications to cancer biology
Gustav Lindwall
Göteborg 2023
ISBN 978-91-7905-899-9

Doktorshavhandlingar vid Chalmers tekniska högskola
Ny serie nr 5365
ISSN 0346-718X

Division of Mathematical Statistics
Department of Mathematical Sciences
Chalmers University of Technology and University of Gothenburg
SE-412 96 Göteborg
Sweden
Telephone +46 (0)31 772 1000

Typeset with LaTeX
Printed by Chalmers Reproservice
Göteborg, Sweden 2023

Statistical inference on interacting particle systems with applications to cancer biology

Gustav Lindwall

Division of Mathematical Statistics
Department of Mathematical Sciences
Chalmers University of Technology and University of Gothenburg

Abstract

Interacting particle systems is a mathematical framework which allows for condensed modelling of complex phenomena undergoing both deterministic and random dynamics. While there are several ways to formulate an interacting particle system, this thesis focuses on modelling such dynamics using stochastic differential equations (SDE:s). The SDE framework was constructed in order to describe the *in vitro* population dynamics of cancer cells.

This thesis introduces the necessary mathematical and biological context, and formulates a model that is subsequently studied in the appended research papers. In the first of three papers, we introduce a novel method of inferring the diffusive properties in such systems based on a higher order numerical approximation of the underlying stochastic differential equations. In the second paper, we model the effect of cell-to-cell interactions, and conduct inference on this model using microscopy data. The third and last paper concerns modelling how the spatial distribution of the cell population affects the cell division rate, and apply our theoretical results to microscopy data.

Put together, the three papers present a cohesive package for modelling and parameter inference that can be applied to population data that is spatial and time-resolved.

Keywords: interacting particle systems, mathematical biology, bayesian inference, stochastic differential equations, reaction-diffusion equations.

List of publications

This thesis is based on the work represented by the following papers:

Paper I. **Lindwall, G.**, Gerlee, P. (2023). Fast and precise inference on diffusivity in interacting particle systems. *Journal of Mathematical Biology, **86**:64*, https://doi.org/10.1007/s00285-023-01902-y.

Paper II. **Lindwall, G.**, Gerlee, P. (2023). Inference on an interacting diffusion system with application to in vitro glioblastoma migration. *Manuscript revision submitted*.

Paper III. **Lindwall, G.**, Gerlee, P. (2023). Bayesian inference on the Allee effect in cancer cell populations using time-lapse microscopy images. *Manuscript submitted, under review*.

Author contributions

Paper I. Formulated the core result, implemented the method in Matlab and conducted the numerical experiments. Wrote the paper.

Paper II. Developed the model in conjunction the co-author. Formulated the inference algorithm, implemented the method in Matlab and conducted the numerical experiments. Wrote the paper.

Paper III. Developed the model in conjunction the co-author. Formulated the inference algorithm, implemented the method in Matlab and conducted the numerical experiments. Wrote the paper.

Acknowledgements

I would first and foremost like to thank my supervisor, Philip Gerlee, for his immense knowledge and helpfulness throughout all aspects of writing this thesis. I would also like to thank my co-supervisor, Umberto Picchini for insightful guidance in the world of statistics. A special thank you goes out to all of my friends who have made both my time at the university and beyond a delightful one. We're talking about quiz nights, concerts, hours at the gym, brisk walks in the rain and all manners of tomfoolery - you know who you are. A special thanks will be directed to my cats, Per and Irma, for all the funny noises they make, and for having fur that I can pet.

Last, I would like to thank my family for continued faith in me, even if I sometimes lack it myself.

Contents

1 Introduction

This is a thesis on the topic of inference methods in mathematical biology. Although mathematical and statistical reasoning has long been the cornerstone in all of the natural sciences, biology and medicine has until the 20th century been thought of and taught without an agreed upon mathematical mode of analysis. However, in the last couple of decades, there has been a burgeoning surge of research into mathematical biology, establishing it as a field of its own. Thus, biology in the 21st century has served as a driver for innovative research in applied mathematics [1].

1.1 A brief history of mathematical biology

The history of mathematical biology is intertwined with the history of science in general and goes back for centuries. However, the recent interest in mathematical biology, and the mutually beneficial relationship it has had with modern mathematics, is beyond parallel when compared to just a century ago. Thus, as a framing device for this brief history of biomathematics, we will begin by making a comparison to the development of physics, pointing out similarities and contrasts.

Up until the Scientific Revolution of the 16th and 17th centuries, physics was a largely qualitative field akin to traditional biology. This period was the advent of modern astronomy, where Johannes Kepler formulated his laws of planetary motion between 1609-1619. Inspired by the heliocentric model of Nicolaus Copernicus and the observations of Tycho Brahe, he detailed the elliptical trajectories of celestial bodies using the notions of geometry that was known to him at this point in history, with remarkable accuracy. Kepler's work in turn inspired and was corroborated by Newtonian mechanics, laying the ground for classical physics.

Crucially, Isaac Newton's systematic study of mechanics made him one of the inventors of calculus, his contributions to the nascent field of mathematics being motivated by a desire to formalize the rules governing the motion of bodies. Ever since the publishing of *Principia* in 1687, physics has arguably been *the* driving force stimulating research in applied mathematics, with new scientific theories demanding new tools. This in turn has stimulated research in pure mathematics, seeking to generalize and rigorously prove the relationships conjectured by physicists. In the end, a feedback loop between physics and mathematics has been the established order in basic science over the last half millennia, with new mathematics facilitating the discovery of new physics - in turn facilitating the need for new mathematics.

One can argue that the reason that biology has been, so to say, late to the mathematical game is the inherent complexity of life. In [2], the authors lay out how even the simplest biological phenomena is by necessity not amenable to the elegant one size fits all formulations underpinning classical physics. As an example, they argue that physics share a common element with mathematics in that the foundational objects are generic - an electron remains interchangeable with any other electron. Meanwhile, every object in the biological world is unique with a unique history, be it on the scale of a single cell to a whole ecosystem. Furthermore, the atomic unit in biology, i.e. a single cell, necessarily adapts and changes in response to its environment, ensuring that biological systems can never be at equilibrium. Thus, a mathematical model in biology demands a lot from the mathematician constructing it, as accounting for the full complexity of even a single cell leads to an intractable mess of interactions. Knowing what you are interested in studying and the limitations of your problem is more important in biomathematics than perhaps any other field.

1.1.1 Population dynamics and epidemiology

Interest in developing mathematical models to describe biological phenomena has existed for a long time. An argument can be made that the Fibonacci sequence, first formulated in the year 1202, is the first model of *population dynamics* [3], as it arises when considering the size of a hare population enjoying unrestricted reproduction. The Fibonacci numbers grow exponentially, and to this day the reasoning behind their construction linger in population dynamics modelling.

Centuries later Leonhard Euler noted that exponential growth could feasibly explain the rate at which populations increase; but he also realized that

exponential growth was not indefinitely sustainable [3]. However, at this point one cannot discuss population dynamics as a true field of science - the observations of Euler and his peers were more curious observations than anything else. The agreed upon founder of population dynamics was the British demographer Thomas Malthus, who in 1798 published the first systematic treaty of exponential population growth. A few decades later, the Belgian mathematician Pierre-François Verhulst introduced the logistic equation; a modification the Mathusian model of growth that takes the phenomena noted by Euler into account.

By the start of the 20th century, the field of population dynamics was in full swing. Seminal models such as the Lotka-Volterra equations modelling predator-prey dynamics was introduced, inspired by the law of mass action from chemistry [4]. With x being the size of the prey population and y the predator population, they are given as

$$\begin{aligned}\dot{x} &= \alpha x - \beta xy,\\ \dot{y} &= \gamma xy - \delta y.\end{aligned}$$

Essentially, the Lotka-Volterra system describes a set of 'chemical reactions' where an unlimited source of prey is available and is replenished at a per-capita rate given by α. When predators comes into contact with the prey, the resulting reaction is that β prey animals are consumed, resulting in γ additional predators. However, the predator population decays at a rate δ. This simple model demonstrates two factors that are ubiquitous throughout mathematical biology - the liberal use of external inspiration, and the need for extreme simplifications to arrive at tractable models.

Parallel to the developments in population dynamics, mathematical modelling of epidemics also began to take form. The early models in epidemiology are similar in spirit to what we see in population dynamics, with the SIR-model for the spread of infectious diseases being the key stone toy model in this field. With S denoting the fraction of the population susceptible to infection, I denoting the infected and R those who have recovered, it is stated as

$$\begin{aligned}\dot{S} &= -rSI,\\ \dot{I} &= rSI - aI,\\ \dot{R} &= aI.\end{aligned}$$

Once again, the analogy to chemical reactions remain, where mixing susceptible and infected populations resolves into additional infected at a rate r, and the infected population decay at a per capita rate a into the recovered state. To

summarize, by the 1920s mathematicians were already employing several techniques to model everything from epidemics to ecosystems, and this is not taking the role of statistics in traditional biology into account. The Lotka-Volterra and SIR models remain important learning tools within their respective fields to this day.

1.1.2 Modern developments - pattern formations, genomics and game theory

In the first half of the 20th century, mathematics underwent an explosive growth spurt, in no small part stimulated by the formulation of modern physics. This period also coincides with the rapid growth of mathematical biology, far too expansive to feasible chronicle in this brief history. Of most importance to this thesis is the study of reaction-diffusion systems, first introduced to study spatial population dynamics in 1937 [5]. Systems of such equations were later studied by Alan Turing, who in 1952 hypothesized that *morphogenesis*, how organisms take their form, was the result of reaction and diffusion of morphogenic chemical substances [6].

Coinciding with the publishing of Turing's paper, DNA was discovered in 1953, and the systematic studying of the genome has since been a driver in mathematical research. Efforts to describe the genome has led to advances in algebra and graph theory [7], and the Human Genome Project started in 1990 demanded the development of new techniques in statistics and computer science [8].

Game theory was established as a field of applied mathematics in 1928 by John von Neumann, with much of the initial applications found in economics. In 1974, the milestone paper *The theory of games and the evolution of animal conflicts* by John Maynard Smith was published [9], establishing evolutionary game theory as a field. Ever since, biology has joined economy in being the spark generating research into game theory.

The emergence of computers has made it possible to study models that are beyond the scope of the theories glanced through in this section. It should come as no surprise then that two of the aforementioned mathematicians, Turing and Neumann, not only took interest in biology, but are also among the founders of computer science. In the 1940s, Neumann was instrumental in the construction of the first *cellular automaton* [10], simple lattice-based models where the state of a lattice point evolves according to its neighbourhood. With the emergence of computer graphics, such systems could be studied through

simulation, including the seminal *Game of Life* constructed by John Conway in 1970. Cellular automata and similar lattice-based models remain ubiquitous in mathematical biology, and computer simulations remain the chief tool employed by practioners in the field.

1.2 Mathematical oncology

Oncology is the branch of medicine that deal with the study, treatment, diagnosis and prevention of cancerous tumours. In the developed world, cancer is among the leading causes of death, and a great impairment to the quality of life of the patient [11]. As we have laid out throughout this chapter, biology and medicine has benefited greatly from mathematical research over the last century. This include oncologists, who have found use of mathematical models in their understanding of cancer [12].

Cancer is an incredibly complex process, and the mathematical modelling of its progression requires one to consider every aspect of mathematical biology considered up to this point.

- At a coarse-grained scale, one frequently models tumours and their interaction the surrounding tissue using reaction-diffusion models. Reaction-diffusion models are also used to model biochemical processes in the body with which the cancer interacts.
- Cancer is a genetic disease, meaning that exploring the human genome is vital in understanding the causes and behaviours of an individual patient's tumour.
- Cancer is an evolutionary process, and in evolution the genotype with the better strategy survives. Such processes are studied using game theory.
- At the fine-grained scale, one can model tumours using *agent based models*. Here, an individual cancer cell constitutes an agent. Cellular automata are common ways to model interacting agents.
- Finally, the use of computer simulation to study hypothetical tumour behaviour is vital to mathematical oncology. Computers are of equal importance when it comes to analyzing experimental data.

In this thesis, we will focus on agent based modelling of tumours, formulate models of *in vitro* cancer cell migration and use statistical tools to infer what

parameters govern the behaviour detected in experimental data. Thus, we will survey the field of cell migration, and give some suggestions on how to formulate this process in mathematical terms.

1.3 Mechanics behind cancer cell migration

Cell migration is vital in the formation and perpetuation of life, whether we are discussing single cell organisms such as bacteria seeking sustenance or human skin cells migrating to close a wound. Understanding cell migration is essential to understanding life. However, not all aspects of cell migration is benign – it is also responsible for the occurrence of tumours, which is what we are to focus on in this thesis.

From a mathematical modelling perspective, a tumour is at the macroscopic level characterised by two main features; the proliferation rate and the cell migration speed. Both of these features are emergent phenomena stemming from complex dynamics at the cell level [13].

On a microscopic level, individual cells migrate throughout its local environment, whose non-cellular components is called the extra-cellular matrix (ECM) and consists of water, proteins and polysaccharides. It acts like a scaffolding for cells migration [14]. The mode of migration of a cell has bio-mechanical explanations on the individual cell level that is a field of research in its own [15, 16]. The process of cell migration starts with cell polarisation; a protrusion is created in the direction that the migration will take place. This protrusion then adheres to the ECM, acting like a cellular 'foot'. The cell then contracts at its new site, resulting in a crawling-like movement. The direction of cell migration in a homogeneous chemical environment is thought of as random, and in mathematics we commonly model it using stochastic processes. The ECM however affects this stochastic process in question. As an example, it has been noted that glioblastoma multiforme cancer cells migrate more than twice as fast in white brain matter as compared to gray [17].

In chemically heterogeneous environments, we observe phenomena such as *chemotaxis*, where perceived changes in the concentration of chemicals around the cell lead to a directed movement in the cell migration process [18]. This can be both a movement towards an attractive chemical such as a source of sustenance, or away from chemicals toxic to the cell in question. Cells also have the ability to communicate with one another by the means of signal substances [18]. Other taxi that influence cell migration is *haptotaxis* and *durotaxis*, governed by

structure of the ECM. These taxis, along with the phenomena of cell adhesion, makes it possible for cells to migrate en masse, as is seen in tumour growth.

1.4 My contribution to the field

The main concern of the journal articles on which this thesis is based is mathematical modelling of *in vitro* cancer cell migration, and statistical inference on key parameters in these models using microscopy imaging data. The guiding principle behind the research is that first principles models derived from physical interactions can aid in the understanding of how cancer cells interact with one another. Subsequent clinical applications of both the modelling and inference presented here can for example be profiling of cells sampled from a specific patient, aiding the physician in choice of clinical intervention.

Paper I: The first paper concerns estimating the diffusivity in a population of identical cells migrating *in vitro*. It expands upon the standard methodology, and derives a robust estimator applicable to any type of spatio-temporal data where random walkers interact with one another.

Paper II: The second paper concerns methods for inferring the nature of local interactions between cancer cells.

Paper III: The third paper introduces a model for how cancer cells up-regulate the division rate of neighbouring cells. Statistical inference on said model is also considered.

The main tools used throughout all three of these papers are *stochastic differential equations (SDE:s)*, used to formulate an agent based model for a cell population. These in turn have an intimate connection to diffusion equations. The outline is as follows; in Chapter 2-4 we introduce the basic theory behind diffusion and SDE:s at a level suitable for the modestly mathematically mature. In Chapter 5-6, we focus on the modelling of cell populations in particular. Chapter 7 is a brief survey of statistical tools used in the papers.

2 Continuum description of diffusion

Diffusion describes the process by which physical media tends to spread from areas of high density to areas of lower density over time. In modern science, the first systematic study of diffusion was made by British chemist Thomas Graham in the 1830's. He noted the following;

> "...gases of different nature, when brought into contact, do not arrange themselves according to their density, the heaviest undermost, and the lighter uppermost, but they spontaneously diffuse, mutually and equally, through each other, and so remain in the intimate state of mixture for any length of time." ([19])

Two decades later, physician Adolf Fick set out formulate a universal law of diffusion, based on Grahams research. He drew inspiration from Fourier's law of heat conduction, formulated in 1822.

2.1 Origins: Fourier's law and Fickian diffusion

The original, phenomenological basis for Fickian diffusion is based in the theory of conservation laws, an already well studied concept in physics at the time, and an assertion of how material *flux* relates to its local density. We denote by J the flux, and by $c(x,t)$ the concentration of a medium at location x at time t. Fick then concluded that the flux is proportional to the gradient of the concentration. Joseph Fourier drew the same conclusion regarding the transfer of heat, and Fick conjectured that the same formalism is applicable to

the diffusion of gases. In one dimension, we state this as

$$J = -D\frac{\partial}{\partial x}c(x,t)$$

where the proportionality constant D is called the *diffusion coefficient*, and the negative sign indicates a flux from higher to lower concentrations. Fick then formulated the following conservation law;

$$\frac{\partial}{\partial t}\int_{x_0}^{x_1} c(x,t)\mathrm{d}x = J(x_1,t) - J(x_0,t) \tag{2.1.1}$$

which intuitively can be understood in the following way: the time evolution of the medium concentration in segment of space $[x_0, x_1]$ equals the difference of the flux at the segments boundaries. In higher dimensions, this result is usually referred to as Gauss' law. By setting $x_1 = x_0 + \Delta$, taking the limit of $\Delta \to 0$ and applying the fundamental theorem of calculus reduces (2.1.1) to the following partial differential equation;

$$\frac{\partial}{\partial t}c(x,t) = D\frac{\partial^2}{\partial x^2}c(x,t) \tag{2.1.2}$$

which is commonly referred to as the *diffusion equation* or *heat equation*. Within this chapter the independent variables will be suppressed in all future mentions of diffusion-type equations for readability.

2.2 Fisher's equation, reaction, diffusion and convection

The diffusion equation is one of the fundamental building blocks in the field of mathematical physics, and its application has indeed diffused into almost every field of science [20]. In the 1930's, the British statistician and biologist Ronald Fisher applied diffusion to a new subject; biology. More precisely, in his paper *The Wave of Advance of Advantageous Genes* [21], Fisher studied how a certain variant of a gene, a so called allele, would spread throughout a uniform population on a line, given that natural selection favored this new mutation. The application in mind were simple lifeforms such as slugs living along a shoreline. If we by $c(x,t)$ denote the concentration of individuals that express

the advantageous gene, *Fisher's equation* in one dimension is given by

$$\frac{\partial c}{\partial t} = D\frac{\partial^2 c}{\partial x^2} + rc\big(1 - \frac{c}{K}\big) \tag{2.2.1}$$

$$c(x, 0) = c_0(x) \tag{2.2.2}$$

where the new parameters r and K are known as the *growth rate* and the *carrying capacity*. The lack of boundary conditions indicate that the diffusion takes place on the entire real line $\mathbf{R}$. Fisher's equation has since its introduction been the fundamental object in spatial ecology and related fields, but it is an idealised equation to be used as a starting point, not applied directly to novel problems and domains. In fact, Fisher himself was adamant about this upon the equation's introduction 1937.

Nevertheless, Fisher's equation stands today as a powerful tool to express spatial evolution of populations, and the way it succinctly summarizes complex emergent behaviours using three macroscopic and measurable parameters gives it an unparalleled place in the field of mathematical oncology, especially so in the development of brain tumours [22]. As a differential equation, it belongs to the class of equations known as semi-linear reaction-diffusion equations; generally such equations are expressed as

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\big(D\frac{\partial u}{\partial x}\big) + f \tag{2.2.3}$$

where $f(u, x, t)$ is the reaction term, and D can now depend on u, x and t. With the reaction term, we aim to encode how a solution u interacts with both the environment and itself. In the original paper by Fisher, the shape of the reaction term is inspired by the law of mass action, commonly employed in chemistry to model chemical reactions in well-stirred mixtures. Today, the most common interpretation of the reaction term is that it is analogous to logistic growth. The logistic differential equation

$$u' = ru\big(1 - \frac{u}{K}\big) \tag{2.2.4}$$

is commonly used to model population growth in the presence of some limiting factor encoded by K, such as competition for resources. In the modelling of tumours, this is the interpretation most commonly taken.

One of the most interesting aspects of solutions to (2.2.1) is their *travelling wave* property; given an initial condition of compact support on $\mathbf{R}$, the Fisher equation is satisfied by a solution sporting a clear, sharp wave front that traverses outward from the initial distribution, exemplified in Figure 2.2.1. This

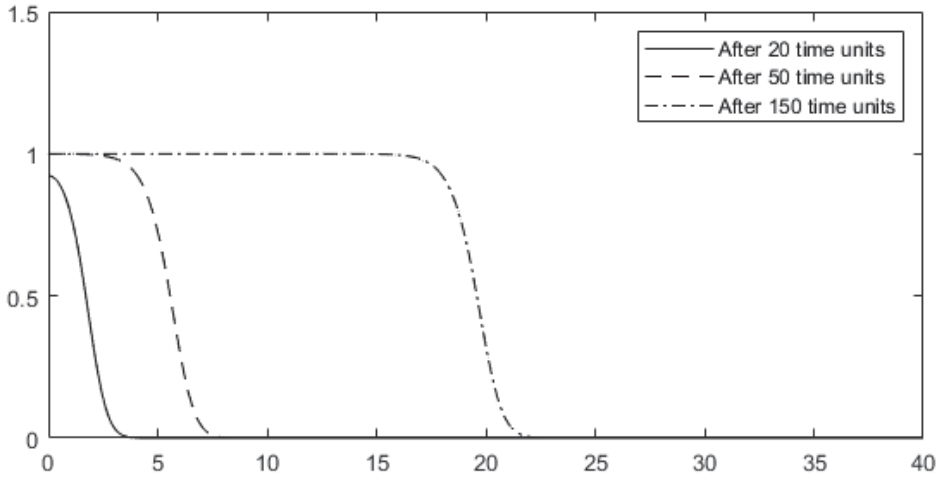

Figure 2.2.1: Solutions to (2.2.1)-(2.2.2) with $D = 0.1$, $r = 1.25$, $K = 1$ and $u_0(x) = \frac{4}{5}(1 + \exp(30(x^2 - 0.05)))^{-1}$ at three different times. Note the consistent wave front. In the modelling of cancer, this wave front is interpreted as the edge of the tumour, moving with some degree of infiltration (given by the slope of the wave front) towards the surrounding tissue.

is in sharp contrast to classical diffusion; solutions to (2.1.2) tend to showcase much wider 'tails'. However, the 'bulk' of the solution to (2.1.2) explores space at a very slow pace. Another important distinction is that solutions to the diffusion equation conserve mass; solutions to Fisher's equation do not.

An intuitive argument in favor of reaction-diffusion as the driver of biological phenomena is provided in Chapter 11 of Murray's excellent text book *Mathematical Biology I* [23]. Here, Murray argues that pure diffusion is simply too slow to be an adequate model of biological phenomena, no matter what D is. He finds that under similar circumstances, the reaction term in even a simple model such as (2.2.1) works as a driving factor, increasing the transportation of biological media by several orders of magnitude. Thus, a common approach to this day in mathematical biology is to tweak the reaction term to suit the circumstances of the phenomena that is being modeled. In addition to the reaction term added to the basic diffusion in (2.2.3), one may also add a convection term, resulting in a reaction-convection-diffusion equation

$$\frac{\partial u}{\partial t} + \frac{\partial h}{\partial x} = \frac{\partial}{\partial x}\big(D\frac{\partial u}{\partial x}\big) + f \tag{2.2.5}$$

where $h(u, x, t)$ is the convection term, and models a directed transport phenomena with velocity h'. The non-linear toy example equation of this kind is

the Burger's equation, given by setting $h = \nu u^2/2$;

$$\frac{\partial u}{\partial t} - \nu u \frac{\partial u}{\partial x} = D \frac{\partial^2 u}{\partial x^2}. \tag{2.2.6}$$

This equation describes a self-propelling behaviour, where the convection speed is proportional to the local concentration. This equation was originally formulated to study shock waves in liquids, but found some use in biology when studying cell cultures where volume exclusion is taken into account [24]. Most importantly, these types of equations arise when considering the *diffusion scaling* of transport equations of the Boltzmann kind [25], but that type of equations lie beyond the scope of this thesis.

3 Stochastic processes and diffusion

We will now shift our focus away from the macroscopic perspective of tumours that makes use of partial differential equations in favour of formulating models based on single-cell modelling. The modelling of cancer cells using stochastic processes is inspired by the field of statistical mechanics, and as such we find it fitting to begin this treatise by its most fundamental construct, Brownian motion.

3.1 The simple random walk and Brownian motion

Definition 3.1.1 (Simple random walk)**.** *Consider a uniform lattice on* $\mathbf{R}$ *with spacing* Δx*, and further consider a particle being located at* $x = 0$ *at time* $t = 0$*. In every time step* Δt*, the particle jumps to either* $-\Delta x$ *or* Δx *with equal probability, and this process is repeated every time step.*

The probability that a simple random walker reaches the lattice point m after n time steps, where $m \in \mathbb{Z}$, $n \in \mathbb{N}$, is given by

$$p(m,n) = \frac{1}{2^n}\frac{n!}{a!(n-a)!}, \qquad a = \frac{n+m}{2}.$$

Now assume that $n \gg 1$, i.e the random walker has been jumping around for a very long time. We find by using Stirling's formula $n! \sim (2\pi n)^{1/2} n^n e^{-n}$ that asymptotically,

$$p(m,n) \sim \Big(\frac{2}{\pi n}\Big)^{1/2} e^{-m^2/(2n)}.$$

Now say that we are interested in the limit of an infinitely fine grid, i.e $\Delta x \to 0$,

$\Delta t \to 0$, and declare the continuous variables $m\Delta x := x$ and $n\Delta t := t$. The probability of finding the particle in the small interval $(x, x + 2\Delta x)$ is then given by

$$u(x,t) := \frac{p(\frac{x}{\Delta x}, \frac{t}{\Delta t})}{2\Delta x} \sim \Big(\frac{\Delta t}{2\pi t(\Delta x)^2}\Big)^{\frac{1}{2}} \exp\Big(-\frac{x^2}{2t}\frac{\Delta t}{(\Delta x^2)}\Big).$$

Finally, by considering the limit where Δx and Δt approach zero so that

$$\lim_{\substack{\Delta x \to 0 \\ \Delta t \to 0}} \frac{(\Delta x)^2}{2\Delta t} = D > 0 \tag{3.1.1}$$

we get the classical result

$$u(x,t) = \frac{1}{\sqrt{4\pi D t}} e^{-x^2/(4Dt)} \tag{3.1.2}$$

where D is a diffusion coefficient and (3.1.1) is known as the *diffusion limit*. Note that (3.1.2) is the distribution of a normal random variable with mean 0 and variance $2Dt$. Denote by $X(t)$ the random walker's location at time t. In the diffusion limit, we have that a random walker starting at $X(0) = x_0$ satisfies the following

$$\mathbb{P}(X(t) \in \Omega \mid X(0) = x_0) = \int_\Omega \frac{1}{\sqrt{4\pi D t}} e^{-(y-x_0)^2/(4Dt)} \mathrm{d}y, \quad \Omega \subset \mathbf{R}$$

and most importantly, we have that

$$X(t) - x_0 \sim \mathcal{N}(0, 2Dt), \tag{3.1.3}$$

i.e the continuous random walk has *Gaussian increments*. This property, along with independence of increments and continuity of paths (not shown here) are the defining properties of *Brownian motion*, also referred to as the *Wiener process*, which is the essential building block in continuous time stochastic processes. Note that this derivation of Brownian motion is quite informal, and there exists a rich literature on the subject for readers interested in a more rigorous treatment of its fundamentals, see for example [26]. The treatise given here was mainly inspired by [23] and [27].

3.2 Itô calculus and the Fokker-Planck equation

We will now head straight into the main application of Brownian motion, namely stochastic calculus.

Definition 3.2.1 (Itô process on integral form)**.** *Let $W(t)$ be a standard Brownian motion, i.e a continuous stochastic process with independent Gaussian increments with variance t. Furthermore, assume that $\int_0^t |a(x,s)|\mathrm{d}s < \infty$ and $\int_0^t b^2(x,s)\mathrm{d}s < \infty$ for all $x \in \mathbf{R}$. An Itô-process $X(t)$ is then given as*

$$X(t) = X_0 + \int_0^t a(X(s),s)\mathrm{d}s + \int_0^t b(X(s),s)\mathrm{d}W(s) \tag{3.2.1}$$

where we can interpret both of these integrals in a Riemann-Stieltjes sense.

The first integral is referred to as the *drift term*, and models deterministic dynamics driving the stochastic process $X(t)$. The second integral is an Itô-integral, where $\mathrm{d}W(s)$ is to be interpreted as a Gaussian increment of infinitesimal size. If $b = 0$, we see that by differentiating (3.2.1) we get a general ordinary differential equation in $X(t)$. With a slight abuse of notation, we can thus discuss *stochastic differential equations* (SDE:s).

Definition 3.2.2 (Itô process on SDE form)**.** *Let $W(t)$, $a(x,t)$ and $b(x,t)$ be as in Definition 3.2.1. An Itô-process $X(t)$ is expressed on SDE form as*

$$\mathrm{d}X(t) = a(X(t),t)\mathrm{d}t + b(X(t),t)\mathrm{d}W(t), \quad X(0) = X_0. \tag{3.2.2}$$

Stochastic calculus give rise to a multitude of interesting phenomena not observed in deterministic calculus. Perhaps the most important result in all of stochastic calculus is the stochastic analog to the chain rule, known as Itô's lemma.

Theorem 3.2.1 (Itô's lemma)**.** *Let $\varphi(x) \in C_0^2(\mathbf{R})$, and consider the stochastic process $\varphi(X(t))$. The stochastic differential $\mathrm{d}\varphi(X(t))$ is given by*

$$\mathrm{d}\varphi(X(t)) = \big(a\frac{\partial\varphi}{\partial x} + \frac{b^2}{2}\frac{\partial^2\varphi}{\partial x^2}\big)\mathrm{d}t + b\frac{\partial\varphi}{\partial x}\mathrm{d}W(t). \tag{3.2.3}$$

Note that under the assumptions made on $\varphi(x)$, we have that $\varphi(X(t))$ also is an Itô process. The linear differential operators constituting the drift- and

diffusion-term in (3.2.3) are commonly given the shorthand

$$\mathcal{L}_0 = a\frac{\partial}{\partial x} + \frac{b^2}{2}\frac{\partial^2}{\partial x^2}, \tag{3.2.4}$$

$$\mathcal{L}_1 = b\frac{\partial}{\partial x}. \tag{3.2.5}$$

An intrinsic property of Itô integrals is their *martingale property*,

Definition 3.2.3 (Martingale)**.** *Let $X(t)$ be a stochastic process, and $\mathcal{F}_s$ be the filtration of $X(t)$ up to time s. We call $X(t)$ a martingale if*

$$\mathbf{E}[X(t) \mid \mathcal{F}_s] = X(s) \tag{3.2.6}$$

for $t > s$. The Wiener process $W(t)$ is an example of a martingale, and crucially we have that for any reasonable function $f(x)$,

$$\mathbf{E}[\int_0^t f(X(s))\mathrm{d}W(s)] = 0.$$

With this in mind, let us now take the expectation of the stochastic process $\varphi(X(t))$ with respect to a probability measure generated by the stochastic process (3.2.1). This measure is given by the probability density at a point x at time t given an initial distribution $p_0(x)$, and we will call this measure $p(x,t)$. By considering (3.2.3) and using the martingale property, we get

$$\begin{aligned}
\frac{\mathbf{E}_p[\varphi(X(t))]}{\mathrm{d}t} &= \mathbf{E}_p\big[a\frac{\partial\varphi}{\partial x} + \frac{b^2}{2}\frac{\partial^2\varphi}{\partial x^2}\big] \Longrightarrow \\
\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbf{R}} \varphi p \mathrm{d}x &= \int_{\mathbf{R}} \big[a\frac{\partial\varphi}{\partial x} + \frac{b^2}{2}\frac{\partial^2\varphi}{\partial x^2}\big] p \mathrm{d}x \Longrightarrow \\
\frac{\mathrm{d}}{\mathrm{d}t}\langle\varphi, p\rangle &= \langle a\frac{\partial\varphi}{\partial x}, p\rangle + \langle\frac{b^2}{2}\frac{\partial^2\varphi}{\partial x^2}, p\rangle
\end{aligned} \tag{3.2.7}$$

where we have used that expectation with respect to a probability measure p defines a linear operator $\mathbf{E}_p[\cdot] = \langle\cdot, p\rangle$. We note that (3.2.7) is a differential equation in p written in weak form. By integration by parts and using that φ is of compact support, we can rewrite (3.2.7) as

$$\langle\varphi, \frac{\partial p}{\partial t}\rangle = -\langle\varphi, \frac{\partial}{\partial x}[ap]\rangle + \frac{1}{2}\langle\varphi, \frac{\partial^2}{\partial x^2}[b^2 p]\rangle$$

which weakly defines a partial differential equation known as the *Fokker-Planck equation*.

Definition 3.2.4 (The Fokker-Planck equation)**.** *Consider an SDE of the form given by Definition 3.2.2. The Fokker-Planck equation for this SDE is given by*

$$\begin{aligned} \frac{\partial}{\partial t}p(x,t) &= -\frac{\partial}{\partial x}[a(x,t)p(x,t)] + \frac{1}{2}\frac{\partial^2}{\partial x^2}[(b(x,t))^2 p(x,t)] \\ p(x,0) &= p_0(x). \end{aligned} \tag{3.2.8}$$

This equation has the remarkable property that given some known initial distribution $p_0(x)$, the solution to (3.2.8) gives the probability distribution for where an Itô process (3.2.1) will be at time t. We also note that (3.2.8) is a diffusion-style differential equation, indeed being a linear convection-diffusion equation. The Fokker-Planck equation serves as a bridge from the microscopic description of diffusion phenomena described by random walks and the macroscopic description given by Fickian diffusion. As a final exercise, we shall consider the Fokker-Planck equation for standard Brownian motion with diffusion coefficient $b = \sqrt{2D}$ and drift coefficient $a = 0$. The Fokker-Planck equation in this case becomes

$$\frac{\partial p}{\partial t} = D\frac{\partial^2 p}{\partial x^2}$$

which is the exact same equation as (2.1.2). Thus we have established that Brownian motion is the microscopic equivalent of standard diffusion. We note once again that the formality has been kept to a minimum in this chapter; we have foregone to mention precise conditions on a and b for existence of solutions and have been playing fast and loose with subtle measure theoretic considerations when manipulating expectations. For a rigorous treatment of Itô's lemma, we once again refer to [26].

3.3 Itô-Taylor expansion and numerical schemes

Cases when an explicit solutions to a stochastic differential equation exists are rare, and solutions to the Fokker-Planck equation are equally challenging to find. Thus, numerical solutions to SDE:s are often what one has to resort to. The theory for finding such solutions is similar to that of numerically solving ordinary differential equations, and as such we will now provide a warm-up example of how the *Taylor*-schemes for solving ODE:s can be found, a family of explicit formulas for numerically solving ODE:s. Consider the first order ODE given by

$$\dot{x}(t) = a(x(t), t). \tag{3.3.1}$$

A Taylor expansion of $x(t)$ around $t = t_0$ is given by

$$x(t) = x(t_0) + (t - t_0)\dot{x}(t_0) + \frac{(t - t_0)^2}{2}\ddot{x}(t_0) + \dots$$

Note that this Taylor expansion can be rewritten using (3.3.1);

$$x(t) = x(t_0) + (t - t_0)a(x(t_0), t_0) + \frac{(t - t_0)^2}{2}\frac{\mathrm{d}}{\mathrm{d}t}\Big(a(x(t), t)\Big) + \dots$$

Truncating after the first order term and setting $t = t_0 + \Delta := t_1$ gives us that

$$x(t_1) \approx x(t_0) + \Delta a(x(t_0), t_0) \tag{3.3.2}$$

which is the first iteration of the familiar *Euler forward* scheme, the simplest numerical scheme for approximating solutions to ODE:s. Truncation at the second order term gives us a more refined approximation of $x(t_1)$,

$$\begin{aligned} x(t_1) \approx x(t_0) &+ \Delta a(x(t_0), t_0) \\ &+ \frac{\Delta^2}{2}\Big(\partial_t a(x(t_0), t_0) + [\partial_x a(x(t_0), t_0)]a(x(t_0), t_0)\Big) \end{aligned}$$

where the chain rule has been used to calculate the second-order term. Repeated use of the chain-rule on higher order terms will give us more refined explicit numerical schemes. This demonstrates that one can use the right-hand side of an ODE such as (3.3.1) to express the Taylor expansion of its solution $x(t)$, which is the key idea in the construction of approximate solutions to SDE:s. The solution to an SDE is not a function as in (3.3.1), but rather an Itô process. Luckily, a stochastic analogue to the Taylor expansion from traditional calculus exists, known as the *Itô-Taylor* expansion.

Now let $X(t)$ be an Itô process given by the (3.2.2). For the sake of simplicity we restrict ourselves to the case of autonomous SDE:s where the coefficient functions a and b have no time-dependence. Remembering Itô's lemma (3.2.3) and the proper formulation of $X(t)$ in (3.2.1), we have

$$X(t) = X(t_0) + \int_{t_0}^{t} a(X(s))\mathrm{d}s + \int_{t_0}^{t} b(X(s))\mathrm{d}W(s), \tag{3.3.3}$$

$$\varphi(X(t)) = \varphi(X(t_0)) + \int_{t_0}^{t} \mathcal{L}_0\varphi(X(s))\mathrm{d}s + \int_{t_0}^{t} \mathcal{L}_1\varphi(X(s))\mathrm{d}W(s).$$

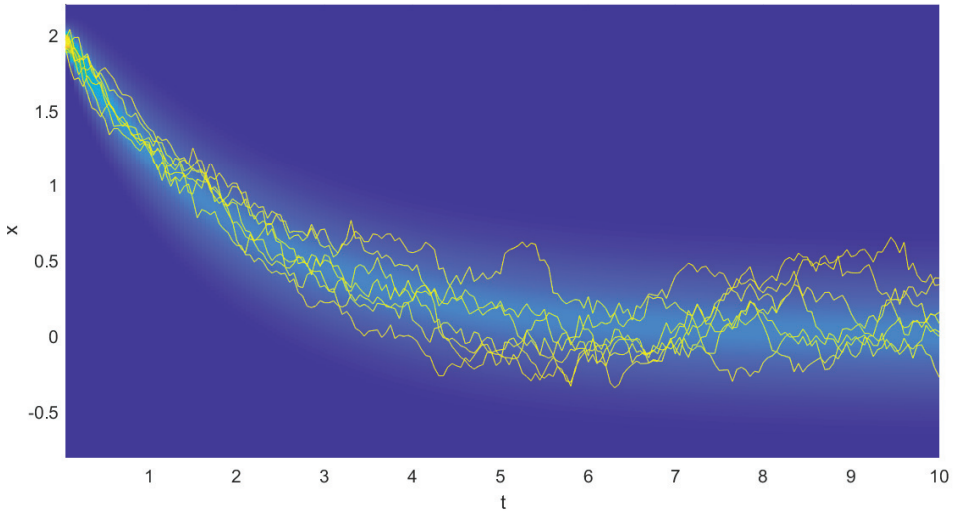

Figure 3.3.1: Demonstration of the duality between an Itô-process and the Fokker-Planck equation. Seven realisations of the SDE $\mathrm{d}X(t) = -0.5X(t)\mathrm{d}t + 0.225\mathrm{d}W(t)$ has been simulated using the Euler-Maruyama scheme (3.3.6), along with the solution to the corresponding Fokker-Planck equation $p(x,t)$. For any time $t > 0$, the state of the SDE solution is a sample from a probability distribution given by $p(x,t)$.

Now, by choosing $\varphi(x) = a(x)$, and $\varphi(x) = b(x)$ respectively, we get that

$$a(X(t)) = a(X(t_0)) + \int_{t_0}^{t} \mathcal{L}_0 a(X(s))\mathrm{d}s + \int_{t_0}^{t} \mathcal{L}_1 a(X(s))\mathrm{d}W(s), \tag{3.3.4}$$

$$b(X(t)) = b(X(t_0)) + \int_{t_0}^{t} \mathcal{L}_0 b(X(s))\mathrm{d}s + \int_{t_0}^{t} \mathcal{L}_1 b(X(s))\mathrm{d}W(s). \tag{3.3.5}$$

By substituting (3.3.4) and (3.3.5) into (3.3.3) and integrating up to $t_1 = t_0 + \Delta$, we get

$$\begin{aligned} X(t_1) =& x(t_0) + \int_{t_0}^{t_1} \Big[a(X(t_0)) + \int_{t_0}^{s} \mathcal{L}_0 a(X(u))\mathrm{d}u + \int_{t_0}^{s} \mathcal{L}_1 a(X(u))\mathrm{d}W(u)\Big]\mathrm{d}s \\ &+ \int_{t_0}^{t_1} \Big[b(X(t_0)) + \int_{t_0}^{s} \mathcal{L}_0 b(X(u))\mathrm{d}u + \int_{t_0}^{s} \mathcal{L}_1 b(X(u))\mathrm{d}W(u)\Big]\mathrm{d}W(s) \end{aligned}$$

and from this exercise, we are ready to piece together the *Euler-Maruyama* scheme for approximating solutions to SDE:s.

Definition 3.3.1 (Euler-Maruyama scheme with remainder term)**.** *Let $X(t)$ be an Itô process, and let $t_k = k\Delta$ for some grid size $\Delta > 0$ and $k = 0, 1, \ldots$. Furthermore let $Z_k \sim \mathcal{N}(0,1)$. Given an initial observation $X(t_0) = X_0$ that might be random or*

deterministic, an approximate sample $\hat{X}_k$ of $X(t_k)$ is given by the numerical scheme

$$\hat{X}_{k+1} = \hat{X}_k + \Delta a(\hat{X}_k) + \sqrt{\Delta} b(\hat{X}_k) Z_k, \tag{3.3.6}$$

$$\begin{aligned} X(t_k) - \hat{X}_k \sim \int_{t_0}^{t_k} \Big[\int_{t_0}^{s} \mathcal{L}_0 \big(a(X(u)) + b(X(u))\big) \mathrm{d}u \\ + \int_{t_0}^{s} \mathcal{L}_1 \Big(a(X(u)) + b(X(u))\Big) \mathrm{d}W(u) \Big] \mathrm{d}W(s). \end{aligned} \tag{3.3.7}$$

We refer to (3.3.6) as the *Euler-Maruyama* scheme for the SDE (3.2.2), with remainder term given by (3.3.7). More exact numerical schemes can be derived for stochastic differential equations by further application of Itô's lemma to aspects of the remainder; such techniques are studied in great detail in [28] and will prove pivotal to the first paper in this thesis.

4 Counting processes and population dynamics

We have until this point chiefly discussed stochastic processes describing the spatial-temporal evolution for a random walker, but this is not the only type of stochastic process that will be useful when modelling cell populations that vary in size over time. A stochastic process describing changes in the number of individuals in a population is called a *birth-death*-process, which we will now give a basic characterization of.

4.1 Birth-death process and survival analysis

Definition 4.1.1 (Birth-death process)**.** *Let $Q(t) = 0, 1, 2, \dots$ be the number of individuals in a population at time t. For times $0 \leq t < s$, we have a transition density of the form*

$$p_{m,n}(s,t) = P(Q(s) = m \mid Q(t) = n), \tag{4.1.1}$$

with support on the non-negative integers $m = 0, 1, 2, \dots$. Birth-death processes are subclass of Markov processes, meaning that for $u < t$ we have

$$P(Q(s) = m \mid Q(t) = n_t, Q(u) = n_u) = P(Q(s) = m \mid Q(t) = n_t].$$

When modelling biological systems, the rate at which a population grows is usually made to be a function of the population size at that particular time. Thus, for long time horizons, i.e when $s \gg t$, an expression such as (4.1.1) becomes intractable due to the multiple ways the population structure can change over such a long duration. Thus some approximation is necessary in order to formulate a workable equation that describes the dynamics of the stochastic process $Q(t)$. Our main focus will be on the expected population

size at time t, i.e $\mathbf{E}[Q(t)] := N(t)$.

We constructed the Fokker-Planck equation, which forecasts the spatial distribution of a random walker long into the future, by considering local dynamics over an infinitesimal time scale. We can thus attempt a similar approach when finding an evolution equation for $Q(t)$. More precisely, we will see that the logistic equation can be derived as the mean-value process of an individual-based birth-death process.

Let's start by assuming that some individual cell was born at time $t = 0$, and the time it divides is given by a random variable β with support on $[0, \infty)$. We refer to the probability density function for β using $b(t)$ and the cumulative probability function using $B(t)$. We assume that β is Markovian, i.e that for $s > t > u > 0$ we have

$$P(\beta > s \mid \beta > t, \beta > u) = P(\beta > s \mid \beta > t)$$

and by Bayes theorem we furthermore get that

$$P(\beta > s \mid \beta > t) = \frac{P(\beta > t \mid \beta > s)P(\beta > s)}{P(\beta > t)} = \frac{P(\beta > s)}{P(\beta > t)}. \tag{4.1.2}$$

where we have used that $P(\beta > t \mid \beta > s) = 1$, since $s > t$. We refer to a realization of the random variable β as the *holding time* until the cell divides. To further characterize our holding time, we let $s = t + \Delta$ for some $\Delta > 0$. By utilizing the integration trick

$$\begin{aligned}\int_{t+\Delta}^{\infty} b(\tau)\mathrm{d}\tau &= \int_{t}^{\infty} b(\tau)\mathrm{d}\tau - \int_{t}^{t+\Delta} b(\tau)\mathrm{d}\tau \\ &= \big[1 - B(t)\big] - \big[B(t+\Delta) - B(t)\big],\end{aligned}$$

(4.1.2) then becomes

$$\begin{aligned}\frac{P(\beta > t+\Delta)}{P(\beta > t)} &= \frac{1 - B(t+\Delta)}{1 - B(t)} \\ &= \frac{\big[1 - B(t)\big] - \big[B(t+\Delta) - B(t)\big]}{1 - B(t)} \\ &= 1 - \frac{B(t+\Delta) - B(t)}{1 - B(t)} := H(t, t+\Delta)\end{aligned} \tag{4.1.3}$$

where $H(t, t+\Delta) \in (0, 1)$ is a shorthand notation for this conditional probability. We now rearrange the left and right hand side of (4.1.3) by moving all B-

dependence to the right hand side, and divide both sides by Δ. We get

$$\frac{1 - H(t, t+\Delta)}{\Delta} = \frac{B(t+\Delta) - B(t)}{\Delta} \cdot \frac{1}{1 - B(t)}. \tag{4.1.4}$$

Assuming that $H(t, t+\Delta)$ is continuous in Δ so that the limit

$$\lim_{\Delta \to 0} \frac{1 - H(t, t+\Delta)}{\Delta} := h(t) \tag{4.1.5}$$

exists,(4.1.4) defines an ordinary differential equation as $\Delta \to 0$ that is commonly referred to as the *survival equation*. Note that this condition is not hard to meet for the vast majority of probability distributions defined on the real half-line; one can note that (4.1.5) can be stated as

$$\lim_{s \to t^+} \frac{H(t,t) - H(t,s)}{s - t}$$

so the only condition is that (4.1.3) is differentiable in s.

Definition 4.1.2 (The survival equation, general definition)**.** *Denote by $B(t)$, $t \geq 0$ the probability that some event of interest has occurred at time t. Furthermore denote by $h(t)$ the hazard rate function, that describes the time-dependent accumulation of probability for said event. The survival equation is given by the ODE*

$$h(t) = \frac{B'(t)}{1 - B(t)}, \qquad B(0) = 0. \tag{4.1.6}$$

The initial condition means that a cell cannot divide at the same time it was born. We can note that by setting $h(t) = \lambda > 0$, the solution to (4.1.6) will be

$$P(\beta < t) = B(t) = 1 - e^{-\lambda t},$$

meaning this choice of $h(t)$ characterizes an exponentially distributed holding time. Another example would be $h(t) = b\lambda t^{b-1}$, that gives us the solution

$$B(t) = 1 - e^{-\lambda t^b}.$$

This is the CDF for a Weibull(λ, b)-distributed random variable, where the hazard rate for the event in question increases as time goes by for $b > 1$, and decreases for $0 < b < 1$. As stated before, the holding time must be dependent on the current population structure in a biologically feasible model, and we model this by making the hazard function depend on $Q(t)$. Our proposed

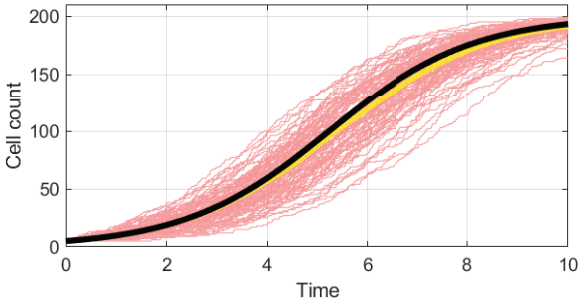

Figure 4.1.1: A demonstration of how the logistic equation arises as the mean of a counting process. In light red, we have 100 realizations of the stochastic process (4.1.8) with $R(N)$ given by (4.1.10), using the parameters $r = 0.7$, $K = 200$, simulated using Algorithm 1. In yellow, the average of these 100 processes. In black, the solution to (4.1.9).

equation that gives us the CDF of β is thus

$$R(Q(t)) = \frac{B'(t)}{1 - B(t)}, \qquad B(0) = 0. \tag{4.1.7}$$

where R is a positive bounded function. Let us now write an approximation of the transition density (4.1.1) in terms of the survival equation formulation of holding times, valid over short time spans. We restrict ourselves to the case of a *pure birth process* for the time being.

Assume that at time t, there are n cells. Let $\beta_1, \beta_2, \ldots, \beta_n$ be n IID copies of β. For some future time $t + \Delta$, the number of cells will be approximately distributed as

$$\big(Q(t+\Delta) \mid Q(t) = n\big) \sim n + \sum_{i=1}^{n} \Big(1 - \mathbf{I}[\beta_i > t + \Delta \mid \beta_i > t]\Big) \tag{4.1.8}$$

given that Δ is small enough to ensure that cells born in the interval $[t, t + \Delta]$ will not divide before the time $t+\Delta$. Now, we take the expected value of (4.1.8), subtract $Q(t) = n$ from both sides, divide both sides by Δ and consider the limit $\Delta \to 0$;

$$\begin{aligned}
\lim_{\Delta\to 0} \Big(\frac{1}{\Delta}\mathbf{E}[Q(t+\Delta) - Q(t) \mid Q(t) = n]\Big) &= \lim_{\Delta\to 0} \frac{1}{\Delta} \sum_{i=1}^{n} \Big(1 - \frac{P(\beta > t + \Delta)}{P(\beta > t)}\Big) \\
&= \underbrace{\lim_{\Delta\to 0} n \frac{1 - H(t, t+\Delta)}{\Delta}}_{\text{see (4.1.5)}} \\
&= nR(n).
\end{aligned}$$

Thus, we have now arrived at an ordinary differential equation that describes

the evolution of $\mathbf{E}[Q(t)] = N(t)$, namely that

$$N'(t) = N(t)R(N(t)). \tag{4.1.9}$$

This equation is used widely within mathematical biology, where $R(\cdot)$ is said to model the per capita reproduction rate of the population. The quintessential form is given by the logistic equation, where

$$R(N) = r(1 - \frac{N}{K}). \tag{4.1.10}$$

Here $r > 0$ is a baseline reproduction rate and $K > 0$ is the carrying capacity of our population, governing the maximum number of individuals that our ecosystem can sustain. Equation (4.1.10) is the simplest type of per-capita reproduction that captures the biological reality of resource scarcity, and we have already by addressing the Fisher equation (2.2.1) demonstrated how logistic growth permeates population dynamics modelling. A typical solution to the logistic equation is visualized in Figure 4.1.1, along with simulations of the process (4.1.8).

Remember that what we have treated in this section is a *pure birth process*, but death can be added to the process quite easily. In the same way that each cell is given a holding time distributed according to the random variable β, one can add an independent random variable ω (and its IID copies ω_i) that governs the holding time until death. Furthermore, let the cumulative distribution $O(t)$ of ω satisfy the survival equation

$$M(Q(t)) = \frac{O'(t)}{1 - O(t)}, \qquad O(0) = 0. \tag{4.1.11}$$

If β and ω are taken to be independent, the transition density (4.1.8) can be modified as

$$\begin{aligned}\big(Q(t+\Delta) \mid Q(t) = n\big) \sim n &+ \sum_{i=1}^{n} \Big(1 - \mathbf{I}[\beta_i > t + \Delta \mid \beta_i > t]\Big) \\ &- \sum_{i=1}^{n} \Big(1 - \mathbf{I}[\omega_i > t + \Delta \mid \omega_i > t]\Big)\end{aligned}$$

and we get a resulting equation for $N(t)$ in

$$N'(t) = N(t)\big(R(N(t)) - M(N(t))\big).$$

4.2 Alternative frameworks for the birth-death process

What we have presented here, formulating and simplifying the birth-death process (4.1.1) in the context of survival analysis, is just one of many different perspectives one can take on this problem. This particular framework lends itself well to individual-based modelling, something that will be justified in the following chapter. A different formulation would perhaps be to consider the *Kolmogorov equations* for continuous-time Markov chains.

Definition 4.2.1 (Kolmogorov equations for birth-death processes)**.** *Let $Q(t)$ be a birth-death process, and set $p_n(t) = P(Q(t) = n)$. Furthermore, let β_n be the rate at which one individual is added to the population when $Q(t) = n$, and ω_n be the rate at which one individual is removed. $p_n(t)$ satisfies the differential equation*

$$p_n'(t) = -(\beta_n + \omega_n)p_n(t) + \beta_{n-1}p_{n-1}(t) + \omega_{n+1}p_{n+1}(t).$$

One can arrive at (4.1.9) by considering the expected value of $Q(t)$ given that $\beta_n = R(n)$ and $\omega_n = 0$ [29]. Yet another, and far more technical approach, would be to use the framework of Itô calculus for jump processes, a field that lie beyond the scope of this thesis but that has interesting implications in providing a rigorous derivation of the Fisher equation [30].

4.3 An algorithm for simulating birth-death processes

We ended the chapter on Brownian motion and Itô processes with a brief summary of numerical methods used for simulating such processes. One can make a symmetry argument that we have arrived at the time to do the same for birth-death processes. The simulation is pleasingly simple - as the time-varying CDF $B(t)$ is defined by an ordinary differential equation, one can quite easily implement a birth-death process along with a simulation of an Itô process by simply using some explicit ODE-solver to step up $B(t)$ along with the Euler-Maruyama scheme used to generate the Itô process for the cell positions. Such an algorithm is exemplified in Algorithm 1. A simulation using this numerical scheme is illustrated in Figure 4.1.1, where we also empirically demonstrate that the logistic equation arises as the average of multiple runs of this stochastic simulation.

Algorithm 1 An algorithm for simulating a pure birth process

Require: $N_0 > 0$ ▷ Initial population size
Require: $r > 0$ ▷ Base division rate
Require: $K > 0$ ▷ Carrying capacity
Require: $\Delta > 0$ ▷ Time step in Euler scheme
Require: $T > 1$ ▷ Number of simulation steps
  $N \leftarrow N_0$
  $B \leftarrow \texttt{zeros}[N, 1]$ ▷ Initiate $B_i(0)$ for cell $i = 1, 2, \ldots$
  $\beta \leftarrow \texttt{rand}[N, 1]$ ▷ Uniform random numbers for simulating holding times
  $Nt \leftarrow N \times \texttt{ones}[1, T]$
  for $t = 2 : T$ **do**
    $n_t = \texttt{zeros}[N, 1]$ ▷ Track if a cell divides this time step
    for $i = 1 : N$ **do**
      $B[i] \leftarrow B[i] + \Delta R(N)(1 - B[i])$
      if $B[i, t] > \beta[i]$ **then**
        $n_t[i] \leftarrow 1$ ▷ Flag that this cell has divided
        $B[i, t] \leftarrow 0$ ▷ Reset the division probability for this cell
        $\beta[i] \leftarrow \texttt{rand}$ ▷ Give it a new holding time
      end if
    end for
    $\eta_t \leftarrow \sum_{i=1}^{N} n_t[i]$ ▷ Calculate the number of new cells
    $B_t \leftarrow \texttt{zeros}[\eta_t, 1]$ ▷ Initiate division probability for new cells
    $\beta_t \leftarrow \texttt{rand}[\eta_t, 1]$ ▷ Give them holding times
    $B \leftarrow [B; B_t]$ ▷ Add newcomers to population
    $\beta \leftarrow [\beta; \beta_t]$
    $N \leftarrow N + \eta_t$
    $Nt[t] \leftarrow N$
  end for

5 Microscopic model of cell migration and proliferation

> It can scarcely be denied that the supreme goal of all theory is to make the irreducible basic elements as simple and as few as possible without having to surrender the adequate representation of a single datum of experience.
>
> *Albert Einstein, 1933*

Now that we have the concept of stochastic differential equations, the Fokker-Planck equation and birth-death processes freshly in our minds, we will move on to the primary subject for this thesis, namely interacting particle systems as models for cancer cell populations.

5.1 Kernel-based modelling of cell interactions

In a mathematical oncology setting, agent-based models are a fairly recent development, but similar methods have a rich history in chemistry and physics. There, agent-based models are the fundamental construct underpinning statistical mechanics on the sub-microscopic level, and thus we will apply similar SDE models to two-dimensional cell migration. In a system of N cells, we will label each individual cell as $i = 1, 2, \ldots, N$ and let $\mathbf{X}_i(t) \in \mathbf{R}^2$ be the location of cell i at time t.

Before moving on to an explicit formulation of an SDE system for $\mathbf{X}_i(t)$, we have to acknowledge a limit in what SDE:s are capable of modelling. A random walker occupies a *point* in space at a specific time, but since the scale of interest here is microscopical, cells have a 'size' that has to be accounted for. We will model the physical presence of the cells using a radially symmetric interaction kernel $\varphi_d(r)$, dependent on a scaling parameter d. We assume that φ is smooth and of compact support, i.e $\varphi(r) \in C_0^\infty(\mathbf{R}^+)$, and propose

$$\varphi_d(r) = \frac{\varphi(r)}{\varphi(d)}. \tag{5.1.1}$$

In Figure 5.1.1, we see an example of how we have adjusted the kernel scaling to be that of an average cell diameter. If we furthermore set the length scale so that $d = 1$, we achieve a microscopic unit suitable for modelling individual cells in a petri dish, while also simplifying our simulations. Note that for longer-range interactions between cells, we might introduce additional kernels with a scaling that differs from the d featured in (5.1.1).

The main purpose of introducing the kernels is to let the cells interact with one another at a short length scale, by forming the basis in an *interaction potential* $U(r)$. In mathematical biology, interaction potentials are used to aggregate mechanical effects that can be difficult to disentangle. Potentials can be either *repulsive*, *attractive* or both. Every cancer cells i is assigned a potential, and it is through this potential that the cell interact with its environment. If another cell is at a distance r from cell i where $\nabla U(r) < 0$, the cells repel one another. If $\nabla U(r) > 0$, they attract. See Figure 5.2.1 for an illustration. The mechanism behind repulsion and attraction can be for any number of reasons, as the potential in itself encode some type of average behaviour, and is not a model of a specific biophysical phenomena. Example of phenomena which results in attraction cell is adhesion, while volume exclusion is a source of repulsion.

A construct that will come of use later on in this chapter is the *empirical measure* $\mu_t(\mathbf{x})$ generated by our population.

Definition 5.1.1 (The empirical measure)**.** *Assume that at time t we have observed N cells centered at $\mathbf{X}_i(t) \in \Omega \subseteq \mathbf{R}^2$, $i = 1, 2, \ldots, N$. The empirical measure generated by this observation is given by*

$$\mu_t(\mathbf{x}) = \frac{1}{N} \sum_{i=1}^{N} \delta(\mathbf{x} - \mathbf{X}_i(t)) \tag{5.1.2}$$

where $\delta(\mathbf{x})$ is the Dirac delta distribution, with the property $\langle f(\mathbf{x}), \delta(\mathbf{x} - \mathbf{a}) \rangle = f(\mathbf{a})$.

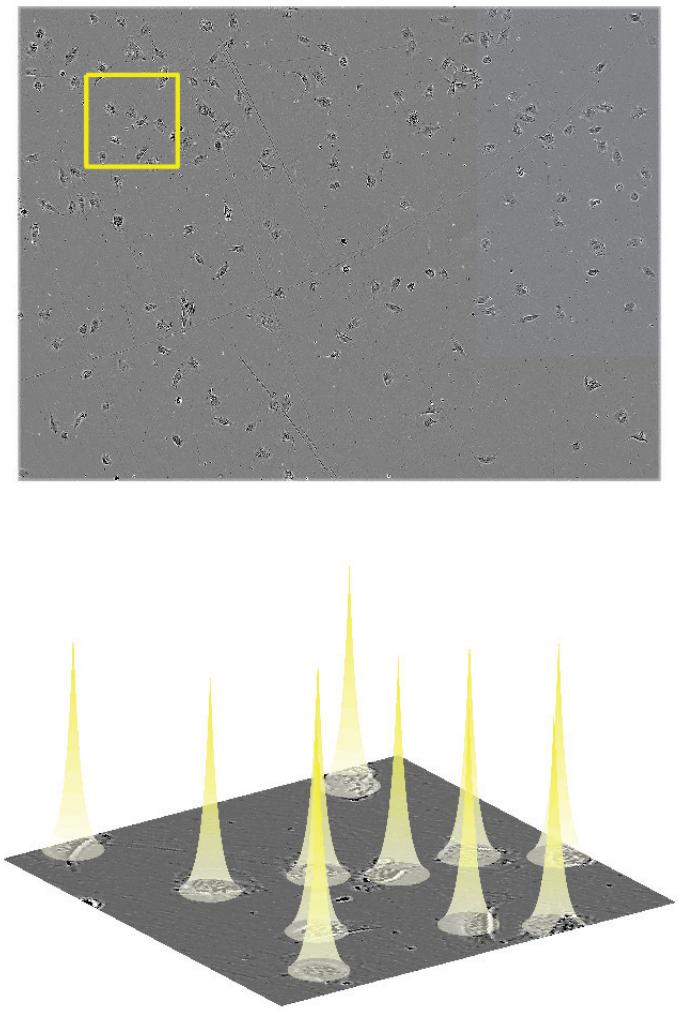

Figure 5.1.1: A demonstration of how kernels are necessary to model our cell population using SDE:s. In the left panel, we see a representative microscopy image of *glioblastoma multiforme* cancer cells. In the right panel, we focus on the marked area in the left panel, adding kernels to the cells. The average diameter $\bar{d}$ of the cells is calculated using the radial distribution function for the cell population, and an exponential kernel with scaling parameter $\bar{d}$ (see (5.1.1)) has been placed at the center of a cell of choice. The SDE system only models the dynamics of the center of the cells, but their interactions and physical presence is accounted for by the kernel.

The empirical measure is to be interpreted as the distribution that indicates whether or not a cell is located at some particular $\mathbf{x} \in \Omega$. If we integrate $\mu_t(\mathbf{x})$ over some domain $A \subseteq \Omega$, we get the *fraction* of the total population that is found within A. Thus, we can define a new distribution $n(A,t) : \Omega \mapsto [0,1]$

$$n(A,t) = \int_\Omega \mathbf{1}_A(\mathbf{x})\mu_t(\mathbf{x})\mathrm{d}\mathbf{x} = \langle \mathbf{1}_A, \mu_t \rangle \tag{5.1.3}$$

which is visualized in the left panel of Figure 5.3.1. One can further define the *number density* at a point $\mathbf{x} \in \Omega$ by setting $A = B(\mathbf{x}, \mathbf{r})$, where $B(\mathbf{x}, \mathbf{r})$ is a ball centered around $\mathbf{x}$ of radius $\mathbf{r}$, and considering the limit

$$\lim_{\mathbf{r}\to\mathbf{0}} N(A,t) = n(\mathbf{x},t) \tag{5.1.4}$$

but this expression only makes sense as a distribution.

5.2 SDE model for the microscopic dynamics

We are now ready to introduce the system of stochastic differential equations that form the cornerstone of the entire thesis. The model choice is meant to be as simple as possible, but not simpler, and thus we adapt ourselves to the setting of our main data source, *in vitro* migration of single cell populations. Denote by $\Omega \subseteq \mathbf{R}^2$ the domain on which our cell migration takes place. We assume the time evolution of $\mathbf{X}_i(t) \in \Omega$ can be described by an SDE

$$\mathrm{d}\mathbf{X}_i(t) = V_i(\mathbf{X}_i(t))\mathrm{d}t + \sqrt{2D}\mathrm{d}W(t) \tag{5.2.1}$$

$$V_i(\mathbf{y}) = -\nabla_{\mathbf{y}} \sum_{j\neq i} U(\|\mathbf{y} - \mathbf{X}_j(t)\|), \tag{5.2.2}$$

$$U(r) = D_e\Big[1 - (\varphi_d(r))^a\Big]^2 - D_e, \tag{5.2.3}$$

where $\sigma := \sqrt{2D}\mathbf{I}$ is a diagonal matrix corresponding to an isotropic Brownian motion. The environment in which the cell migration takes place is in all our applications homogeneous and rich in nutrients and oxygen, thus making the isotropic diffusion a plausible assumption. We have chosen this formulation of the interaction potential due to the wide array of attractive-repulsive behaviours one can extract from it depending on ones choice of kernel $\varphi(r)$. In Figure 5.2.1, we have formulated the *Morse* potential by setting $\varphi(r) = e^{-r}$, and one can formulate the *Lennard-Jones* potential by setting $\varphi(r) = r^{-\alpha}$, i.e letting the attraction-repulsion follow an inverse power law.

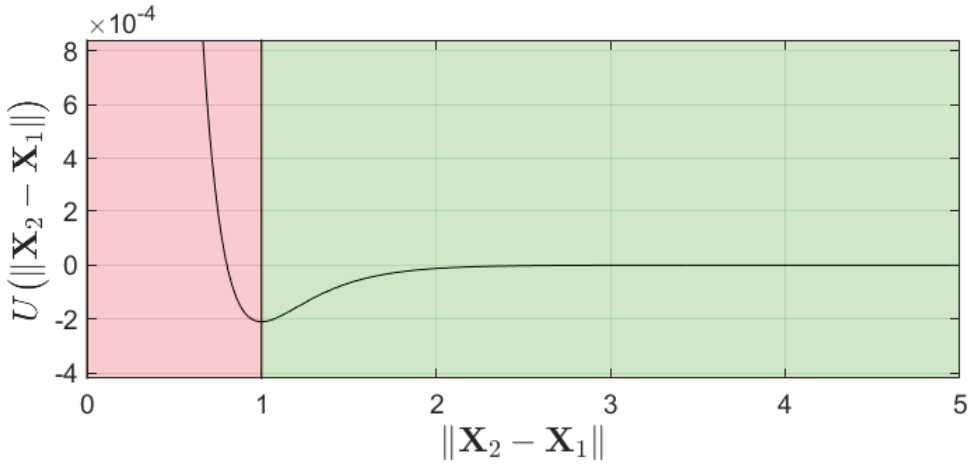

Figure 5.2.1: An example of an interaction potential frequently employed in cancer cell population modelling, the *Morse potential*. It is acquired by using the kernel $\varphi(r) = e^{-r}$ in (5.2.3), and sports an attractive-repulsive behaviour. Let $\mathbf{X}_1$ be the position of cell 1, and set that as the origin. The "force" with which it acts upon a neighbouring cell located at $\mathbf{X}_2$ depends on their distance, being *repulsive* at close ranges and *attractive* outside of this range. The amount of attraction or repulsion is proportional to the gradient of the potential, meaning that very little interaction takes place at distances longer than a few cell diameters. Parameters used: $D_e = 2.1 \cdot 10^{-4}$, $a = 3.5$ and $d = 1$.

Given this SDE system, we are now interested in what the solution to the corresponding Fokker-Planck equation might look like. The transition density from a known population distribution to a possible future distribution is of crucial importance in the inference problem this thesis is centered around.

Denote by $P_N(\vec{\mathbf{x}}, t)$ the joint probability distribution for our population, where $\vec{\mathbf{x}} = [\mathbf{x}_1\ \mathbf{x}_2\ \dots\ \mathbf{x}_N]^T$. Here, we interpret $\mathbf{x}_i \in \Omega$ as the variable for the space where the probability density for cell i is considered. We get an equation for $P_N(\vec{\mathbf{x}}, t)$ by deriving the Fokker-Planck equation for the system (5.2.1), given as

$$\partial_t P_N(\vec{\mathbf{x}}, t) = \nabla_{\vec{\mathbf{x}}} \cdot \Big[D \nabla_{\vec{\mathbf{x}}} P_N(\vec{\mathbf{x}}, t) + \sum_{i=1}^{N} \mathcal{V}_i(\mathbf{x}_i) P_N(\vec{\mathbf{x}}, t) \Big]. \tag{5.2.4}$$

$$\mathcal{V}_i(\mathbf{x}_i) = -\nabla_{\mathbf{x}_i} \sum_{j \neq i} U(\|\mathbf{x}_i - \mathbf{x}_j\|) \tag{5.2.5}$$

However, this equation is unwieldy, difficult to solve and of spatial dimension $2N$. One needs to approximate $P(\vec{\mathbf{x}}, t)$ using a simpler set of objects, something we will now explore.

5.3 Deriving the marginal distributions

The traditional way to deal with such massive joint probability distributions is to derive the marginal distribution for each individual cell, and then approximate the joint distribution using the marginals. The cells are assumed to be identical in this model, making their labels $i = 1, 2, \ldots, N$ interchangeable. Thus, we will consider the marginal distribution for cell 1, knowing that the same procedure can be used to find the marginal distribution for any cell i. The marginal for cell 1 is given by

$$P_1(\mathbf{x}_1, t) = \int_{\Omega^{N-1}} P_N(\vec{\mathbf{x}}, t) \mathrm{d}\mathbf{x}_2 \cdots \mathrm{d}\mathbf{x}_N. \tag{5.3.1}$$

As we will come to see in a little bit, finding the marginal distribution for just a single cell will not be enough. We will also need the *pairwise* marginal densities, given by

$$P_2(\mathbf{x}_1, \mathbf{x}_2, t) = \int_{\Omega^{N-2}} P_N(\vec{\mathbf{x}}, t) \mathrm{d}\mathbf{x}_3 \cdots \mathrm{d}\mathbf{x}_N \tag{5.3.2}$$

for the cell-pair 1 and 2; the marginalization was performed with respect to these two labels to simplify the notation. Note that due to the invariance under labelling, one can interpret (5.3.1) as the probability of finding *any cell* at location $\mathbf{x}_1 \in \Omega$ at time t. Likewise, a valid interpretation of (5.3.2) is the probability that at time t, one cell is centered at $\mathbf{x}_1 \in \Omega$ and another at $\mathbf{x}_2 \in \Omega$.

We will now attempt to apply the marginalization procedure (5.3.1) to (5.2.4), hopefully giving us an equation from which we can find $P_1(\mathbf{x}_1, t)$ without too much trouble. For the sake of simplicity, we will assume that $\Omega = \mathbf{R}^2$ for the remainder of this section; i.e the domain on which the cell migration takes place is very large compared to the area covered by the cell population. Integrating with respect to $\mathbf{x}_2, \ldots, \mathbf{x}_n$ on both sides of (5.2.4), we have

$$\int_{\Omega^{N-1}} \partial_t P_N \mathrm{d}\mathbf{x}_2 \ldots \mathrm{d}\mathbf{x}_N = \int_{\Omega^{N-1}} \nabla_{\vec{\mathbf{x}}} \cdot \Big[D \nabla_{\vec{\mathbf{x}}} P_N + \sum_{i=1}^{N} \mathcal{V}_i P_N \Big] \mathrm{d}\mathbf{x}_2 \ldots \mathrm{d}\mathbf{x}_N \Rightarrow$$

$$\partial_t P_1 = D \nabla^2_{\mathbf{x}_1} P_1 + \int_{\Omega^{N-1}} \nabla_{\mathbf{x}_1} \cdot \Big[\mathcal{V}_1 P_N \Big] \mathrm{d}\mathbf{x}_2 \ldots \mathrm{d}\mathbf{x}_N \tag{5.3.3}$$

$$+ \sum_{i=2}^{N} \int_{\Omega^{N-1}} D \nabla^2_{\mathbf{x}_i} P_N + \nabla_{\mathbf{x}_i} \Big[\mathcal{V}_i P_N \Big] \mathrm{d}\mathbf{x}_2 \ldots \mathrm{d}\mathbf{x}_N \tag{5.3.4}$$

where all independent variables have been suppressed for readability. Let us begin by dealing with (5.3.4). Our domain of integration is the product space of $N-1$ copies of $\mathbf{R}^2$. Let us say that $\mathbf{x}_i$ is our variable of interest; we can then approximate our domain as

$$\Pi_i = \Omega \times \Omega \times \ldots \times \underbrace{\big([-\xi,\xi]\times[-\xi,\xi]\big)}_{\text{the } i\text{:th term}} \times \Omega \ldots$$

and consider the limit $\xi \to \infty$. Let $\hat{\mathbf{n}}_i$ be the normal vector pointing outwards from the ξ-dependent square, which we will call Ξ. Setting $i = 2$ and using Green's theorem, the first half of the integral in (5.3.4) thus becomes

$$\lim_{\xi\to\infty}\int_{\Pi_2} D\nabla^2_{\mathbf{x}_2} P_N \mathrm{d}\mathbf{x}_2 \ldots \mathrm{d}\mathbf{x}_N = \lim_{\xi\to\infty} D\int_{\Omega^{N-2}}\Big[\int_{\partial\Xi}\nabla_{\mathbf{x}_2}P_N\cdot\hat{\mathbf{n}}_2\mathrm{d}s_2\Big]\mathrm{d}\mathbf{x}_3\ldots\mathrm{d}\mathbf{x}_N = 0$$

as the probability flux $\nabla_{\mathbf{x}_i}P$ must vanish as $\xi\to\infty$. Likewise, we perform the same integration trick for the second half of the integrand, and arrive at

$$\lim_{\xi\to\infty}\int_{\Pi_2}\nabla_{\mathbf{x}_2}\Big[\mathcal{V}_2 P_N\Big]\mathrm{d}\mathbf{x}_2\ldots\mathrm{d}\mathbf{x}_N = 0.$$

The same results hold for any other choice of i, meaning that (5.3.4) equals zero. Thus, what remains a potential issue is the integral term in (5.3.3), so let us handle it with some extra care. We have that

$$\int_{\Omega^{N-1}}\nabla_{\mathbf{x}_1}\cdot\Big[\mathcal{V}_1 P_N\Big]\mathrm{d}\mathbf{x}_2\ldots\mathrm{d}\mathbf{x}_N =$$
$$-\int_{\Omega^{N-1}}\nabla_{\mathbf{x}_1}\cdot\Big[\nabla_{\mathbf{x}_1}\sum_{j=2}^{N}U(\|\mathbf{x}_1-\mathbf{x}_j\|)P_N(\vec{\mathbf{x}},t)\Big]\mathrm{d}\mathbf{x}_2\ldots\mathrm{d}\mathbf{x}_N$$

can be split up into the sum of $N-1$ identical integrals owing the independence of cell labels;

$$\int_{\Omega^{N-1}}\nabla_{\mathbf{x}_1}\cdot\Big[\nabla_{\mathbf{x}_1}U(\|\mathbf{x}_1-\mathbf{x}_2\|)P_N(\vec{\mathbf{x}},t)\Big]\mathrm{d}\mathbf{x}_2\ldots\mathrm{d}\mathbf{x}_N =$$
$$\int_{\Omega^{N-1}}\nabla_{\mathbf{x}_1}\cdot\Big[\nabla_{\mathbf{x}_1}U(\|\mathbf{x}_1-\mathbf{x}_j\|)P_N(\vec{\mathbf{x}},t)\Big]\mathrm{d}\mathbf{x}_2\ldots\mathrm{d}\mathbf{x}_N,\quad j\neq 2.$$

For the cell pair 1 and 2, we have that

$$\int_{\Omega^{N-1}} \nabla_{\mathbf{x}_1} \cdot \Big[\nabla_{\mathbf{x}_1} U(\|\mathbf{x}_1 - \mathbf{x}_2\|) P_N(\vec{\mathbf{x}}, t)\Big] \mathrm{d}\mathbf{x}_2 \dots \mathrm{d}\mathbf{x}_N = \tag{5.3.5}$$

$$\nabla_{\mathbf{x}_1} \cdot \int_{\Omega} \Big[\nabla_{\mathbf{x}_1} U(\|\mathbf{x}_1 - \mathbf{x}_2\|) P_2(\mathbf{x}_1, \mathbf{x}_2, t)\Big] \mathrm{d}\mathbf{x}_2. \tag{5.3.6}$$

Renaming the spatial variables as the general $\mathbf{x}_1 := \mathbf{x}$, $\mathbf{x}_2 := \mathbf{x}'$, this procedure gives us the closest possible explicit equation for the one particle density that is available,

$$\partial_t P_1(\mathbf{x}, t) = D\Delta_{\mathbf{x}} P_1(\mathbf{x}, t) - (N-1)\nabla_{\mathbf{x}} \cdot \int_{\Omega} \Big[\nabla_{\mathbf{x}} U(\|\mathbf{x} - \mathbf{x}'\|) P_2(\mathbf{x}, \mathbf{x}', t)\Big] \mathrm{d}\mathbf{x}'. \tag{5.3.7}$$

As is usually the case in the mathematics of bridging stochastic models with their deterministic counterparts, there are multiple ways to formulate the process covered in this section. One alternative way is to apply Itô's lemma to the expected value of (5.1.2); in [31], the authors demonstrate that (5.3.7) can be derived using that method. Thus, one can have the perspective on (5.3.7) that rather than describing the marginal distribution for one particular cell, it rather describes the time evolution of the empirical measure. This is visualized in the right panel of Figure 5.3.1, where $P_1(\mathbf{x}, T)$ for some $T > 0$ has been approximated using an empirical measure as initial condition, i.e $P_1(\mathbf{x}, 0) = \mu_0(\mathbf{x})$. Furthermore, this interpretation of $P_1(\mathbf{x}, t)$ will be important when cell division is introduced into the model.

The equation for the 2-particle density can likewise be recovered by applying Itô's lemma to the expected value of

$$\nu_t(\mathbf{x}, \mathbf{x}') = \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{j \neq i} \delta(\mathbf{x} - \mathbf{X}_i(t)) \delta(\mathbf{x}' - \mathbf{X}_j(t)),$$

which is to be interpreted as an indicator of whether or not the statement "given a cell is located at $\mathbf{x}'$, there is a cell at $\mathbf{x}$" is true. Note that the calculations performed in this section can be adapted to a compact subset Ω of $\mathbf{R}^2$ by imposing a no-flux boundary condition on Ω, but performing these calculations would yield little further insight.

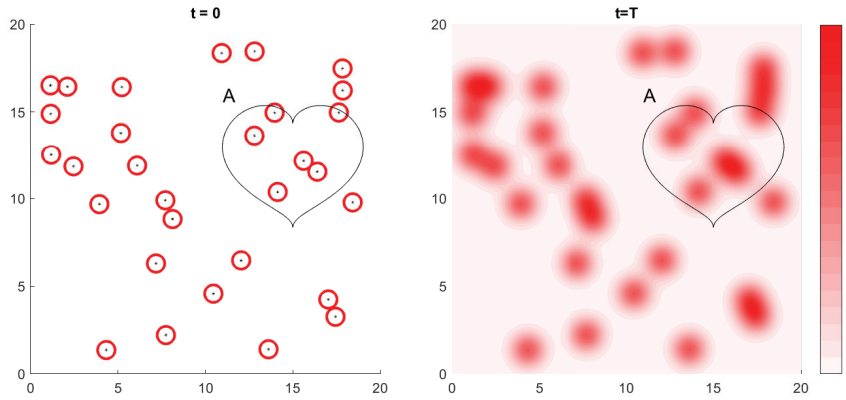

Figure 5.3.1: An example of how the number density for a small population of 30 "cells" is calculated on the domain $\Omega = [0, 20] \times [0, 20]$, at two different times. At $t = 0$, we know where all cells are and the domain A contains six cell centers, marked by black dots, and thus (5.1.3) gives us $n(A, 0) = 1/5$. The red circles indicate the cell diameters. At some time $t = T$ in the future, the cells have diffused throughout Ω, and the probability of finding a cell centered at a certain location $\mathbf{x}$ is given by the the probability distribution $P_1(\mathbf{x}, t)$, given by (5.3.7). We now have $n(A, T) \approx 0.1417$, as $n(x, T)$ is the solution to (5.3.7) at time T using (5.1.2) as initial condition.

5.4 Closure methods for N-particle systems

We note that (5.3.7), the single cell marginal density, turned out to be dependent on the pairwise density. In turn, the equation for the pairwise density depends on the *three-particle* density, demonstrating that there is a hierarchical structure to the equations giving us the marginals, called the BBGKY hierarchy [32]. In fact, if one were to perform the marginalisation done in (5.3.1) and (5.3.2) N times, one would end up with a system of N PDE:s, with (5.2.4) being the equation at the top of the hierarchy! Thus, one has to *close* the hierarchy at some low level, by approximating the k-particle density P_k using P_1, P_2, $\dots P_{k-1}$. This is usually done at the two or three particle level, and we will now briefly cover two such methods for closure at the pairwise density, and how one goes about reconstructing $P(\vec{\mathbf{x}}, t)$ using these closure methods.

The simplest closure available is called the *mean field closure*, where we assume that the marginal distributions are independent of one another, and thus

$P_N(\vec{\mathbf{x}}, t)$ can be factorized as

$$P_N(\vec{\mathbf{x}}, t) = \prod_{i=1}^{N} P_1(\mathbf{x}_i, t) \Rightarrow$$
$$P_2(\mathbf{x}, \mathbf{x}', t) = P_1(\mathbf{x}, t) P_1(\mathbf{x}', t) \tag{5.4.1}$$

Under this closure, (5.3.7) becomes

$$\partial_t P_1(\mathbf{x}, t) = D\Delta_{\mathbf{x}} P_1(\mathbf{x}, t) - (N-1)\nabla_{\mathbf{x}} \cdot \Big[P_1(\mathbf{x}, t)(\nabla_{\mathbf{x}} U * P_1)(\mathbf{x}, t) \Big] \tag{5.4.2}$$

where $\nabla_{\mathbf{x}} U * P_1(\mathbf{x}, t) = \int_\Omega \nabla_{\mathbf{x}} U(\|\mathbf{x} - \mathbf{x}'\|) P_1(\mathbf{x}', t) \mathrm{d}\mathbf{x}'$. However, there is a glaring issue with this closure method, demonstrated with Figure 5.4.1. In the exact pairwise density, one should have that $P_2(\mathbf{y}, \mathbf{y}, t) = 0$ for $y \in \Omega$, as two particles of non-zero size should not be able to occupy the same place at the same time. However, the single particle marginal densities does not take such spatial structures into account, resulting in a joint distribution that *peaks* for the region where $\mathbf{x} \approx \mathbf{x}'$. One can remedy this by the *method of matched asymptotic expansion* [33], where one can find an approximate formula for $P_2(\mathbf{x}, \mathbf{x}', t)$ in

$$P_2(\mathbf{x}, \mathbf{x}', t) \approx P_1(\mathbf{x}, t) P_1(\mathbf{x}', t) e^{-U(\|\mathbf{x} - \mathbf{x}'\|)}. \tag{5.4.3}$$

This expression takes the repulsion into account (see panel (c) in Figure 5.4.1) and thus provides an adequate approximation of the pair density at the small price of some added complexity. We arrive at a PDE for the single particle density in

$$\partial_t P_1(\mathbf{x}, t) = D\Delta_{\mathbf{x}} P_1(\mathbf{x}, t) - (N-1)\nabla_{\mathbf{x}} \cdot \Big[P_1(\mathbf{x}, t)(\nabla_{\mathbf{x}} U * (P_1 e^{-U}))(\mathbf{x}, t) \Big]. \tag{5.4.4}$$

5.5 Introducing cell division to the particle system

The model would not be complete without a brief discussion on how to introduce new cells to the population. We will do this in the context of a survival analysis formulation of the cell division probability, but unlike the model outlined in Chapter 3, we will make the hazard function for cell i depend on the *local cell density* around i, and not the total population size as in (4.1.7).

Let $\varphi_b(r)$ be a kernel as in (5.1.1), but now with a different scaling b than the one used for cell-to-cell interactions. The different scaling stems from that

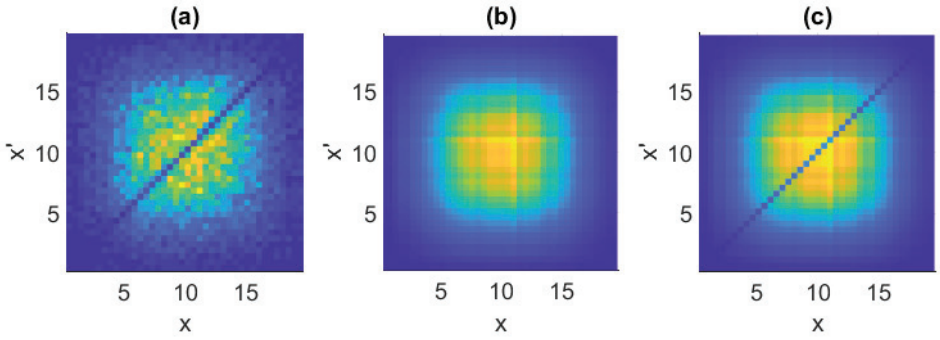

Figure 5.4.1: A demonstration of how the mean field and matched asymptotic expansion closure operates. In panel (a), we have the pairwise density for a system of five interacting cells diffusing along the line segment $\Omega = [0, 20]$, recovered by Monte Carlo simulations. In panel (b), we demonstrate how mean field approximation (5.4.1) fails to take the spatial anti-correlation along the diagonal into account. In panel (c), we observe how the correction term included in the matched asymptotic expansion-derived closure remedies this issue.

cells might influence the division rates of one another at longer ranges than the mechanical interactions modelled by U, something called *autocrine signalling* [34]. We can then compute the local density $\rho_i(t)$ around cell i at time t as

$$\rho_i(t) = \frac{1}{C}\sum_{j\neq i}\varphi_b(\|\mathbf{X}_i(t) - \mathbf{X}_j(t)\|) = \frac{1}{C}\varphi_b * \left(\mu_t(\mathbf{x}) - \delta(\mathbf{x} - \mathbf{X}_i(t))\right) \quad (5.5.1)$$

where C is a normalizing constant, and $\delta(\mathbf{x} - \mathbf{X}_i(t))$ is subtracted as cell i does not contribute to its own local density. The idea is now to assign a birth process $B_i(t)$ to cell i. Assuming that cell i was born at time $t = 0$, we formulate it as

$$R(\rho_i(t)) = \frac{B_i'(t)}{1 - B_i(t)}, \qquad B_i(0) = 0. \quad (5.5.2)$$

The algorithm for adding new agents to an interacting particle system is essentially the same as Algorithm 1, and allows us to dynamically add new cells to our population with ease. The exact method is covered in greater detail in the third paper included in this thesis.

With cell division accounted for and an interpretation of $P_1(\mathbf{x}, t)$ as the evolution of the empirical measure for our population, one can follow the procedure laid out in [30] and derive a PDE that describes the mean field evolution of the

number density $n(\mathbf{x}, t)$, as defined by (5.1.4). Without delving into too much detail, we find that

$$\partial_t n = D\Delta_{\mathbf{x}} n - (N-1)\nabla_{\mathbf{x}} \cdot \Big[n(\nabla_{\mathbf{x}} U * n) \Big] + n[R(n) - M(n)] \tag{5.5.3}$$

where $R(n)$ and $M(n)$ are the reproduction and mortality rates as laid out in (4.1.7) and (4.1.11), respectively. We have then come full circle; by assuming no cell interactions and setting $R(n) = r(1 - n/K)$ and $M(n) = 0$, (5.5.3) becomes the Fisher equation, giving us the microscopic derivation of it hinted at in the end of Chapter 3.

6 Model expansions and future considerations

While the model presented in the previous chapter is capable of modelling a wide array of different cell migration behaviours, we would be remiss not to mention some extensions that could be considered in future work on this problem. This includes further modification of the system of diffusion equations (5.2.1)-(5.2.3), as well as more advanced variants that take cell velocity into account.

6.1 Models for heterogeneous media

The cell cultures that have laid the foundation of the data sets in this thesis have been kept under quite optimal conditions; their access to nutrients and oxygen has been unhindered, and the medium through which the cell migrates is homogeneous. This is not the case when one models tumour growth in real tissue, however. Thus, we will now introduce a few constructs that must be accounted for when modelling cell migration *in vivo*, first one out being what is summarily called the *extracellular matrix*, or ECM.

The ECM is a complex and dynamic network of molecules that surrounds the cells in a multicellular organisms, composed of proteins, glycoproteins, polysaccharides, and more. The ECM acts as a scaffold for cells, offering mechanical stability and transmitting mechanical cues that influence cell behaviour [35]. When modelling cell migration, the ECM can for example regulate the diffusion coefficient in (5.2.1), representing that denser tissue is more difficult for cells to migrate through. It can also provide an additional term to the drift part of (5.2.1), through the processes of *durotaxis* and *heptotaxis* [36]. However, ECM modelling is most commonly found in the literature as part in

reaction-diffusion models [37, 38], although agent-based approaches have been studied as well [39, 40].

Next up, one can introduce the influence of chemical compounds and oxygen levels to the model, referred to as *chemotaxis*. Here, one typically model the molecules of the chemicals in mind as diffusive, a consequence of that a single agent (i.e a molecule) in the chemical solution is much smaller than an agent (i.e a cell) in the tumour [41]. Common approaches is that chemorepellants, representing toxins, steer the cells away and chemoattractants, for example nutrients, have an attractive effect, feeding into the drift term of (5.2.1) as a flow along the gradient of the chemical concentration. Oxygen levels in the tissue is known to have a crucial effect on cell division rate [42], and the common approach is to model the oxygen as diffusing from blood vessels around the tumour. Detailed *in vivo* models of tumour evolution shall thus take *angiogenesis* into account, the process in which the tumour facilitates the generation of new blood vessels around itself to keep itself alive [43].

6.2 Velocity-driven equations

Though diffusion models are ubiquitous in mathematical biology, the fundamental assumption they rest on is physically impossible. Remember that diffusion is derived as a limit for the simple random walk, which is based on discrete, instantaneous jumps in the position of our random walker. As such, these models are referred to as *position jump models* or *kangaroo processes*. A more physically grounded view point is to instead formulate a Newtonian mechanics-based model for the position of our cells.

It has been long noted that certain microbiological organisms, famously the bacteria *E. coli*, displays a *persistence* in its migratory behaviour [44, 23]. These bacteria tend to express a 'run-and-tumble' pattern in its migration whereas it swims in a straight line for some time, stops, and then picks a new direction, perhaps stimulated by some chemotactic factor. This behaviour has also been observed in several types of cancer [45]. The 'run-and-tumble' migration behaviour is not possible to model using the SDE model (5.2.1)-(5.2.3), but several models have been suggested by the literature. We will now briefly discuss those, for cells migrating in $\mathbf{R}^2$.

Assume that the cell or bacteria maintains a constant speed during the runs, and a run lasts for a random holding time $\tau > 0$. After this, a new direction for the run is picked according to a probability distribution corresponding to the

boundary of the unit circle (i.e the interval $[0, 2\pi]$), and the run begins anew in this new direction. In [46], one can find a pleasant review article laying out the basics of this paradigm.

A second variant of a similar idea can be proposed, this time rooted in SDE modelling, making it more readily applicable to the work laid out in this thesis. Note that the drift term in (5.2.1) corresponds to a *velocity* derived from the potential field U acting on the cell position $\mathbf{X}_i(t)$. Thus, we add a term $\mathbf{V}_i(t)$ to the drift part, corresponding to cell i:s velocity, something not taken into account in a kangaroo process. We will now let the randomness in the system act upon the velocity of the particle instead. This type of equation is known in statistical mechanics as a *Langevin* equation and is formulated as

$$\mathrm{d}\mathbf{X}_i(t) = \mathbf{V}_i(t)\mathrm{d}t, \tag{6.2.1}$$

$$\mathrm{d}\mathbf{V}_i(t) = -\big[\mu\mathbf{V}_i(t) + \frac{1}{m}\mathbf{F}_i(\mathbf{X}_i(t))\big]\mathrm{d}t + \sigma\mathrm{d}W(t). \tag{6.2.2}$$

In this equation, μ is analogous to mechanical friction, m is the mass of a cell and $\mathbf{F}_i(\mathbf{X}_i)$ is a force acting on cell i, that may include both external sources and cell-to-cell interactions. Thus, (6.2.1)-(6.2.2) can be viewed as Newton's equations of motion as seen in classical mechanics, but with a noise term added to the acceleration. This model is studied in detail in [47].

The drawback with this much more physically realistic model is the added complexity. Not only does it contain more parameters, but now each cell i correspond to a point in the *phase space* $\mathbf{R}^2 \times \mathbf{R}^2$ associated with it; the velocity is considered to be a point in $\mathbf{R}^2$. Alas, this has doubled the dimension of our previous problem, but non-the-less a Fokker-Planck equation can still be derived. Under the assumption that $\mathbf{F}_i$ involves no cell-to-cell interactions, the marginal distribution for a single particle is given by Itô's lemma, and is given by [47]

$$\frac{\partial}{\partial t}p + \mathbf{v}\cdot\nabla_{\mathbf{x}}p + \nabla_{\mathbf{v}}\cdot\big((-\mu\mathbf{v} + \frac{\mathbf{F}}{m})p\big) = \frac{\sigma^2}{2}\nabla_{\mathbf{v}}\cdot(\nabla_{\mathbf{v}}p) \tag{6.2.3}$$

where $p := p(\mathbf{x}, \mathbf{v}, t)$ is the probability of finding a cell located at $\mathbf{x}$, with velocity $\mathbf{v}$, at time t. The index has been dropped for $\mathbf{F}$ as all cells follow the same dynamics in this case. If one were to include cell-to-cell interactions into this equation, this would serve as a derivation of a Boltzmann-like equation [47, 14], who are famously difficult to handle in even the simplest cases [48].

7 Elements of computational statistics

The goal of this thesis is two-fold. On one hand, we aim to evaluate methods of modelling cancer migration by using stochastic differential equations. On the other, we wish to conduct inference on these interacting particle systems based on real data. For this purpose, we will present a short discussion of relevant topics in statistical inference. In this chapter, we will stick to nomenclature common within Bayesian inference; most importantly we will refer to systems of SDE:s such as (5.2.1) as *stochastic dynamical systems*.

7.1 Transition probabilities in dynamical systems and construction of likelihood functions

With our recent discussion of the Fokker-Planck equation, we have illustrated that the state of stochastic dynamical system described by an SDE can be sampled directly from the solution to its corresponding PDE (5.2.4), illustrated by Figure 3.3.1 for a simple one-dimensional case. Given this, assume that we have observed a particle system undergoing stochastic dynamics on Ω at times t_0, t_1, $\dots$ t_K, and refer to these observations as $\mathbf{X}_k$, $k = 0, \dots, K$. Let $P_1^k(\mathbf{x}, t)$ be the solution to the Fokker-Planck equation (5.3.7) on the time interval $[t_k, t_{k+1})$ using the initial condition

$$P_1^k(\mathbf{x}, t_k) = \frac{1}{N} \sum_{i=1}^{N} \delta(\mathbf{x} - \mathbf{X}_{ik}). \tag{7.1.1}$$

Here $\mathbf{X}_{ik}$ is the k:th observation of the i:th particle. The interpretation of this is that given an observation, we are certain where cells are, thus the Dirac

δ-spikes at their centers. As the cells are allowed to diffuse, we become more uncertain of where the cells are, and the probability of finding cells in a certain region $A \in \Omega$ at time t is given by $\int_A P_1^k(\mathbf{x}, t)\mathrm{d}\mathbf{x}$. This is neatly illustrated by Figure 5.3.1, where the observed configuration at time $t = 0$ in the left panel has been smeared out at time $t = T$ in the right panel due to diffusion.

To ease up the notation, the remainder of this chapter will illustrate parameter inference for the case of a one-dimensional SDE such as the one given by (3.2.2). The principles remain the same when applying these methods to more complex models.

Assume that the drift coefficient a or the diffusion coefficient b in (3.2.2) have some parameters θ for which we wish to conduct statistical inference given the observations $X_{0:K}$, where $0 : K$ is used to refer to a collection of observations. Since the transition density will depend on these parameters, we will use the notation $p_k(x, t; \theta)$, where p_k is the solution to (3.2.8) with initial condition $p(x, t_k) = X_k$. We are now ready to construct a transition probability from X_k to X_{k+1} in the following manner;

$$\pi(X_{k+1}|X_k, \theta) := p_k(X_{k+1}, t_{k+1}; \theta). \tag{7.1.2}$$

We get the likelihood for our entire sequence of observations in

$$\pi(X_{0:K}|\theta) = \prod_{k=0}^{K-1} \pi(X_{k+1}|X_k, \theta). \tag{7.1.3}$$

One can then use the likelihood (7.1.3) to evaluate how *likely* a sequence of observations $X_{0:K}$ are given a parameter set θ. The theory presented in this segment is nothing that cannot be found in an ordinary text book on Bayesian inference or machine learning, see Bishops textbook [49] for an excellent overview of many related topics.

7.2 Simulation of SDE:s and Monte Carlo methods

Before diving into the problem of maximizing the likelihood (7.1.3), we should discuss how to approximate the solution to a Fokker-Planck equation using *Monte Carlo* methods. Given an observed state X_0 and setting it as our initial condition in (3.2.8), we wish to find an approximation of $p(x, T)$ for some $T > 0$. We find this by iterating the Euler-Maruyama scheme (3.3.6) on a partitioning

of K grid points on the interval $[0, T]$;

$$\hat{X}_{k+1} = \hat{X}_k + \frac{T}{K} a(\hat{X}_k) + \sqrt{\frac{T}{K}} b(\hat{X}_k) Z_k \qquad \text{for } k = 0, \ldots, K-1$$
$$\hat{X}_0 = x_0.$$

Running this numerical scheme once gives us $\hat{X}_K$ as an approximate sample from $p_k(x, T)$, as the Euler-Maruyama scheme deviates from the underlying distribution by the error term (3.3.7). One sample is not enough to say something about the distribution of $p_k(x, T)$, however. We thus repeat this procedure many times; by performing S independent iterations of (3.3.6) we get the weak convergence

$$\lim_{S \to \infty} \frac{1}{S} \sum_{s=1}^{S} \varphi(\hat{X}_K^s) \xrightarrow{w} \int_{\mathbf{R}} \varphi(x) p_k(x, T) \mathrm{d}x \tag{7.2.1}$$

where φ is a test function and $\hat{X}_K^s$ corresponds to the s:th sample run of the Euler-Maruyama scheme. This is known as the Monte Carlo approach to finding a transition density, and a lot more on this subject can be found in the monolithic text book on the subject by Kloeden and Platen [28]. Through Monte Carlo simulation, we can now look back at Figure 3.3.1 and note a duality. When introduced, we viewed Figure 3.3.1 as an example of how one can obtain the probability distribution for a stochastic dynamical system at time T by solving the PDE (3.2.8) up until that time. Now however, we can see it the other way; how one can approximate a solution to (3.2.8) using simulation by (3.3.6).

7.3 Bootstrap particle filter for likelihood approximation

In theory, the likelihood expression (7.1.3) is readily available to us when performing parameter inference on (5.2.1) by repeatedly solving the PDE (5.2.4) K times, using the observations as initial conditions. But as stated at the beginning of Chapter 5, solutions to (5.2.4) are notoriously difficult to find in most cases. When dealing with interacting particle systems, one usually resolves to solving such equations by simulating the underlying system [50], and then reconstructing $P(\vec{\mathbf{x}}, t)$ using the mean field closure (5.4.1). Performing fast and accurate Monte Carlo simulations of a complex model can be tricky, and special methods are needed to make it computationally feasible. One method is to use *particle filters*, which we will demonstrate the usage of in the case of a simple

one dimensional SDE. This description is more or less based on [51].

Particle filtering is a *Sequential Monte Carlo method* used to sample from *hidden states* of our dynamical system. In our application, a hidden state would be any configuration the particle system takes at times $t \neq t_k$. Intuitively, one can understand that a hidden state "close to t_{k+1}" contains more information about the likelihood structure of at time t_{k+1} than the observed state at t_k. The question is then how to access this hidden state, and the answer to that question is to use the Euler-Maruyama scheme as an *importance sampler*. By letting Y_k be a hidden state on the interval (t_k, t_{k+1}), we can rewrite the left-hand side of (7.1.2) as

$$\pi(X_{k+1}|X_k, \theta) = \int_\Omega \pi(X_{k+1}|Y_k, \theta)\pi(Y_k|X_k, \theta)\mathrm{d}Y_k \tag{7.3.1}$$

using Bayes theorem. We can then make an analogy to (7.2.1), with the transition probability $\pi(X_{k+1}|Y_k, \theta)$ in (7.3.1) takes the role of the test function. We then compute the integral in (7.2.1) using Monte Carlo simulation of the hidden state. With S samples from the hidden state, this gives us

$$\pi(X_{k+1}|X_k, \theta) \approx \frac{1}{S}\sum_{s=1}^{S} \pi(X_{k+1}|Y_k^s, \theta). \tag{7.3.2}$$

For improved accuracy, one can inject multiple hidden states between each observation, and apply variance reduction techniques; see for example [52] for a review article on such techniques.

7.3.1 Numerical example

We finish with a simple numerical example, visualized using Figure 7.3.1. The two first in a sequence of K observations at times t_1 and t_2 of some SDE

$$\mathrm{d}X(t) = -\mu X(t)\mathrm{d}t + \sigma \mathrm{d}W(t) \tag{7.3.3}$$

are marked with red dots, with $t_2 - t_1 = \Delta$. We simulate the process starting from $X_1 = X(t_1)$ a total of $S = 3$ times for some parameters $\hat{\theta} = (\hat{\mu}, \hat{\sigma})$, over a coarse grid of $K = 10$ steps between the observations. If we believe that $\hat{\mu} = \mu$ and $\hat{\sigma} = \sigma$, this should give us three samples from a hidden state,

$$Y_s \sim X\Big(t_1 + \frac{K-1}{K}\Delta\Big)$$

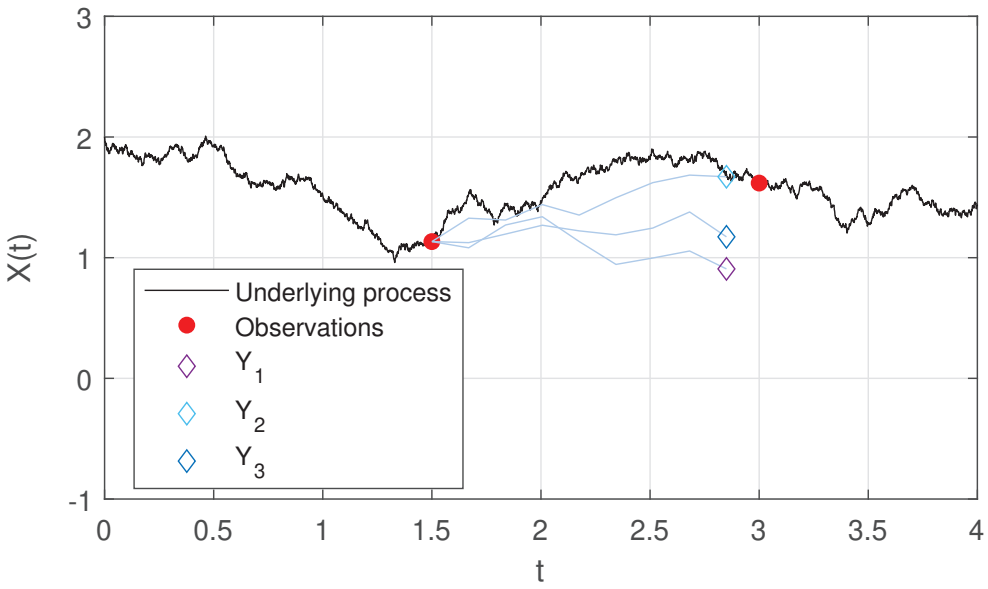

Figure 7.3.1: An example of how to use particle filters and the Euler-Maruyama scheme to get an approximate transition distribution for an Itô process. In black, we have a realization of the SDE (7.3.3). In red, we have the observations at $t_1 = 1.5$ and $t_2 = 3$. Three samples Y_s from a hidden state are shown, achieved through simulation using a proposed set of parameters $\hat{\mu}, \hat{\sigma}$.

where $s = 1, 2, 3$. Our goal now is to evaluate how well $\hat{\theta}$ agrees with the unknown parameters in (7.3.3). The Euler-Maruyama scheme gives us an approximation of $X_2 = X(t_2)$ given Y_s,

$$X_2 \mid Y_s \sim \mathcal{N}(Y_s(1 - \Delta\hat{\mu}), \Delta\hat{\sigma}^2).$$

This in turn lets us approximate of the transition density (7.1.2) using (7.3.2),

$$\pi(X_{k+1} \mid X_k, \hat{\theta}) \approx \frac{1}{S} \sum_{s=1}^{S} \frac{1}{\sqrt{2\pi\Delta\sigma^2}} e^{-\frac{(X_{k+1} - Y_s(1-\Delta\mu))^2}{2\Delta\sigma^2}} \tag{7.3.4}$$

for $k = 1$. This expression is tractable, so with a sequence $0 : K$ of observations we can inject (7.3.4) into (7.1.3) and employ tools from mathematical optimization to maximize (7.1.3) with respect to $\hat{\theta}$.

8 Summary of papers

Paper I. The first paper tackles the problem of estimating the diffusivity of Brownian particles undergoing frequent, short-range interactions. The magnitude of random motion is often quantified using mean squared displacement, which provides a simple estimate of the diffusion coefficient. However, this method often fails when data is sparse or interactions between agents frequent. In order to address this, we derive a conjugate relationship in the diffusion term for large interacting particle systems undergoing isotropic diffusion, giving us an efficient inference method. The method accurately accounts for emerging effects such as anomalous diffusion stemming from mechanical interactions. We apply our method to an agent-based model with a large number of interacting particles, and the results are contrasted with a naive mean square displacement-based approach. We find a significant improvement in performance when using the higher-order method over the naive approach. This method can be applied to any system where agents undergo Brownian motion and will lead to improved estimates of diffusion coefficients compared to existing methods.

Paper II. In this paper, we introduce a stochastic interacting particle system as a model of *in vitro* glioblastoma migration, along with a maximum likelihood-algorithm designed for inference using microscopy imaging data. The inference method is evaluated on *in silico* simulation of cancer cell migration, and then applied to a real data set. We find that the inference method performs with a high degree of accuracy on the *in silico* data, and achieve promising results given the *in vitro* data set.

Paper III. The Allee effect in biology describes the phenomenon that the per capita reproduction rate increases along with the population density at low densities. Allee effects have been observed at all scales, including in microscopic environments where individual cells are taken into account. This is great interest to cancer research, as understanding critical tumour density thresholds

can inform treatment plans for patients. In this paper, we introduce a simple model for cell division in the case where the cancer cell population is modelled as an interacting particle system. The rate of the cell division is dependent on the local cell density, introducing an Allee effect. We perform parameter inference of the key model parameters through Markov Chain Monte Carlo, and apply our procedure to two image sequences from a patient-derived cervical cancer cell line. The inference method is verified on *in silico* data to accurately identify the key parameters, and results on the *in vitro* data strongly suggest an Allee effect.

Bibliography

[1] Reed MC. Mathematical biology is good for mathematics. Notices of the AMS. 2015;62(10):1172-6.

[2] Longo G, Soto AM. Why do we need theories? Progress in Biophysics and Molecular Biology. 2016;122(1):4-10.

[3] Bacaër N. Axplock i den matematiska populationsdynamikens historia. Nicolas Bacaër; 2022.

[4] Goel NS, Maitra SC, Montroll EW. On the Volterra and other nonlinear models of interacting populations. Reviews of modern physics. 1971;43(2):231.

[5] Provine WB. The origins of theoretical population genetics: with a new afterword. University of Chicago Press; 2001.

[6] Turing AM. The chemical basis of morphogenesis. Bulletin of mathematical biology. 1990;52:153-97.

[7] Cohen JE. Mathematics is biology's next microscope, only better; biology is mathematics' next physics, only better. PLoS biology. 2004;2(12):e439.

[8] Percus JK. Mathematics of genome analysis. vol. 17. Cambridge University Press; 2002.

[9] Smith JM. The theory of games and the evolution of animal conflicts. Journal of theoretical biology. 1974;47(1):209-21.

[10] Kari J. Theory of cellular automata: A survey. Theoretical computer science. 2005;334(1-3):3-33.

[11] Nayak MG, George A, Vidyasagar M, Mathew S, Nayak S, Nayak BS, et al. Quality of life among cancer patients. Indian journal of palliative care. 2017;23(4):445.

[12] Altrock PM, Liu LL, Michor F. The mathematics of cancer: integrating quantitative models. Nature Reviews Cancer. 2015;15(12):730-45.

[13] Skog J, Würdinger T, Van Rijn S, Meijer DH, Gainche L, Curry WT, et al. Glioblastoma microvesicles transport RNA and proteins that promote tumour growth and provide diagnostic biomarkers. Nature cell biology. 2008;10(12):1470-6.

[14] Chauviere A, Hillen T, Preziosi L. Modeling the motion of a cell population in the extracellular matrix. Discrete Contin Dyn Syst. 2007:250-9.

[15] Bodor DL, Pönisch W, Endres RG, Paluch EK. Of cell shapes and motion: the physical basis of animal cell migration. Developmental cell. 2020;52(5):550-62.

[16] Malik AA, Gerlee P. Mathematical modelling of cell migration: stiffness dependent jump rates result in durotaxis. Journal of mathematical biology. 2019;78(7):2289-315.

[17] Swanson KR, Alvord Jr EC, Murray J. A quantitative model for differential motility of gliomas in grey and white matter. Cell proliferation. 2000;33(5):317-29.

[18] Alberts B. Molecular biology of the cell. Garland science; 2008.

[19] Philibert J. One and a half century of diffusion: Fick, Einstein before and beyond. Diffusion fundamentals. 2006.

[20] Cussler EL. Diffusion: mass transfer in fluid systems. Cambridge university press; 2009.

[21] Fisher RA. The wave of advance of advantageous genes. Annals of eugenics. 1937;7(4):355-69.

[22] Gatenby RA, Gawlinski ET. A reaction-diffusion model of cancer invasion. Cancer research. 1996;56(24):5745-53.

[23] Murray JD. Mathematical biology: I. An introduction. vol. 17. Springer Science & Business Media; 2007.

[24] Gamba A, Ambrosi D, Coniglio A, De Candia A, Di Talia S, Giraudo E, et al. Percolation, morphogenesis, and Burgers dynamics in blood vessels formation. Physical review letters. 2003;90(11):118101.

[25] Hadeler KP. Reaction transport equations in biological modeling. Mathematical and computer modelling. 2000;31(4-5):75-81.

[26] Mörters P, Peres Y. Brownian motion. vol. 30. Cambridge University Press; 2010.

[27] Klebaner FC. Introduction to stochastic calculus with applications. World Scientific Publishing Company; 2012.

[28] Kloeden PE, Platen E. Numerical Solution of Stochastic Differential Equations. Springer; 1992.

[29] Kendall DG. On the generalized" birth-and-death" process. The annals of mathematical statistics. 1948;19(1):1-15.

[30] Oelschläger K. On the derivation of reaction-diffusion equations as limit dynamics of systems of moderately interacting stochastic processes. Probability Theory and Related Fields. 1989;82(4):565-86.

[31] Middleton AM, Fleck C, Grima R. A continuum approximation to an off-lattice individual-cell based model of cell migration and adhesion. Journal of theoretical biology. 2014;359:220-32.

[32] Born M, Green HS. A general kinetic theory of liquids I. The molecular distribution functions. Proceedings of the Royal Society of London Series A Mathematical and Physical Sciences. 1946;188(1012):10-8.

[33] Bruna M, Chapman SJ, Robinson M. Diffusion of particles with short-range interactions. SIAM Journal on Applied Mathematics. 2017;77(6):2294-316.

[34] Gerlee P, Altrock PM, Malik A, Krona C, Nelander S. Autocrine signaling can explain the emergence of Allee effects in cancer cell populations. PLoS computational biology. 2022;18(3):e1009844.

[35] Hay ED. Cell biology of extracellular matrix. Springer Science & Business Media; 1991.

[36] Reinhardt JW, Krakauer DA, Gooch KJ. Complex matrix remodeling and durotaxis can emerge from simple rules for cell-matrix interaction in agent-based models. Journal of biomechanical engineering. 2013;135(7):071003.

[37] Eikenberry SE, Sankar T, Preul MC, Kostelich EJ, Thalhauser C, Kuang Y. Virtual glioblastoma: growth, migration and treatment in a three-dimensional mathematical model. Cell proliferation. 2009;42(4):511-28.

[38] Bymed HM. Biological inferences from a mathematical model for malignant invasion. Invasion Metastasis. 1996;16:209-21.

[39] Wang Z, Butner JD, Kerketta R, Cristini V, Deisboeck TS. Simulating cancer growth with multiscale agent-based modeling. In: Seminars in cancer biology. vol. 30. Elsevier; 2015. p. 70-8.

[40] Letort G, Montagud A, Stoll G, Heiland R, Barillot E, Macklin P, et al. PhysiBoSS: a multi-scale agent-based modelling framework integrating physical dimension and cell signalling. Bioinformatics. 2019;35(7):1188-96.

[41] Painter KJ. Mathematical models for chemotaxis and their applications in self-organisation phenomena. Journal of theoretical biology. 2019;481:162-82.

[42] Alarcón T, Byrne HM, Maini PK. A mathematical model of the effects of hypoxia on the cell-cycle of normal and cancer cells. Journal of theoretical biology. 2004;229(3):395-411.

[43] Schor AM, Schor SL. Tumour angiogenesis. The Journal of pathology. 1983;141(3):385-413.

[44] Bearon R, Pedley T. Modelling run-and-tumble chemotaxis in a shear flow. Bulletin of mathematical biology. 2000;62(4):775-91.

[45] Theveneau E, Marchant L, Kuriyama S, Gull M, Moepps B, Parsons M, et al. Collective chemotaxis requires contact-dependent cell polarity. Developmental cell. 2010;19(1):39-53.

[46] Othmer HG, Dunbar SR, Alt W. Models of dispersal in biological systems. Journal of mathematical biology. 1988;26(3):263-98.

[47] Othmer HG, Xue C. The mathematical analysis of biological aggregation and dispersal: progress, problems and perspectives. In: Dispersal, individual movement and spatial ecology: a mathematical perspective. Springer; 2013. p. 79-127.

[48] Harris S. An introduction to the theory of the Boltzmann equation. Courier Corporation; 2004.

[49] Bishop CM, Nasrabadi NM. Pattern recognition and machine learning. vol. 4. Springer; 2006.

[50] Bossy M. Some stochastic particle methods for nonlinear parabolic PDEs. In: ESAIM: proceedings. vol. 15. EDP Sciences; 2005. p. 18-57.

[51] Schön T, Lindsten F. Learning of dynamical systems–Particle filters and Markov chain methods. Draft available. 2015.

[52] Durham GB, Gallant AR. Numerical techniques for maximum likelihood estimation of continuous-time diffusion processes. Journal of Business & Economic Statistics. 2002;20(3):297-338.