Spatial Moment Models for Collective Cell Behaviour

A thesis submitted in partial fulfilment of the requirements for the Degree of
Doctor of Philosophy in Mathematics

by Rachelle N. Binny

under the supervision of
Assoc. Prof. Michael Plank and Assoc. Prof. Alex James
and Prof. Matthew Simpson

School of Mathematics and Statistics
University of Canterbury

2016

Abstract

The ability of cells to undergo collective movement plays a fundamental role in tissue repair, development and cancer. Interactions occurring at the level of individual cells may give rise to spatial structure, such as clustering, in a moving population. *In vitro* cell culture studies have shown that the presence of such spatial structure can play an important role in determining the dynamics of migrating cells at a population level. However, mathematical models that consider population-level behaviour often take a mean-field approach, which assumes that individuals interact with one another in proportion to their average density and neglects the effects of spatial structure.

In this work, we develop a lattice-free individual-based model (IBM) for collective movement in one-dimensional space. The IBM uses random walk theory to model the stochastic interactions occurring at the scale of individual migrating cells. In particular, our model allows an individual's direction of movement to be affected by interactions with other cells in its neighbourhood, providing insights into how directional bias generates spatial structure. As an alternative to the mean-field approach, we employ spatial moment theory to develop a population-level model which accounts for spatial structure and predicts how these individual-level interactions propagate to the scale of the whole population. The IBM is used to derive an equation for dynamics of the second spatial moment (the average density of pairs of cells) which incorporates the neighbour-dependent directional bias and we solve this numerically for a spatially homogeneous case.

Extending our model to consider cell behaviour in two-dimensional space makes it more amenable for use alongside experimental data. Using imaging data from *in vitro* experiments, we estimate parameters for the two-dimensional model and show that it can generate similar spatial structure to that observed in a 3T3 fibroblast cell population. Finally, we incorporate cell birth and death into our two-dimensional model to consider how these processes give rise to spatial structure and how, in turn, this spatial structure affects the collective dynamics.

Acknowledgements

First and foremost, I would like to thank my fantastic supervisors Mike Plank, Alex James and Mat Simpson, for their guidance, time and unwavering support. I am immensely grateful to them for their encouragement and for inspiring me to pursue a career in research. I look forward to our future collaborations!

Thank you to Mat, for the invitation to visit Queensland University of Technology to run cell experiments in the lab. Thanks also to Parvathi Haridas for her support in carrying out the experimental work. I am incredibly grateful to them both for their time, advice and for making the entire Brisbane trip so enjoyable from start to finish.

Further thanks go to Richard Law, for sharing his expertise in spatial moment theory both on my visit to the University of York and on his visits to the University of Canterbury. Richard's enthusiasm for the field made working with him a real privilege and a pleasure. I thoroughly enjoyed all of our discussions and trips into the stunning New Zealand backcountry.

I would also like to thank the staff and postgraduates in the School of Mathematics and Statistics at UC, for creating such a positive and inspiring working environment. In particular, thanks to my officemate and friend Shakira Suwan with whom I have immensely enjoyed sharing the PhD experience.

Finally, a huge thank you to my family and friends, I am incredibly grateful to you all for encouraging and supporting me every step of the way.

This PhD was supported by a scholarship funded by the Royal Society of New Zealand Marsden Fund, Grant Number 11-UOC-005.

"It is not the mountain we conquer, but ourselves."
$\sim$ Sir Edmund Hillary

Contents

Nomenclature

Functions

Individual-based Model

ψ_i	Movement rate for an individual i, T^{-1}
$\hat{B}_i$	Birth rate for an individual i, T^{-1}
$\hat{D}_i$	Death rate for an individual i, T^{-1}
$w(z)$	Interaction kernel for movement rate, T^{-1}
$w_\rho(z)$	Interaction kernel for birth rate, T^{-1}
$\mu(x, y)$	Movement PDF, L^{-d} in $d-$dimensional space
$b(x)$	Neighbour-dependent directional bias, dimensionless
$\hat{h}(x)$	Spatially-dependent weight, dimensionless
$u(x, y)$	Spatially-dependent neighbour contribution to directional bias, dimensionless
$v(z)$	Interaction kernel for directional bias, L
p_r	Probability of moving right, dimensionless
$g(\theta; b)$	Probability of moving in direction θ, dimensionless
$\mu^{(\rho)}(x, y)$	Dispersal PDF, L^{-d}

Spatial Moment Dynamics

Z_n	n^{th} spatial moment, L^{-dn}
$\tilde{Z}_3(x, y, z)$	Moment closure for third spatial moment, L^{-dn}
$M_1(x)$, $M_2(x, y)$	Expected movement rates for first and second moment dynamics, T^{-1}
$B_1(x)$, $B_2(x, y)$	Expected birth rates for first and second moment dynamics, T^{-1}
D	Expected death rate, T^{-1}
$\mu_1(x, y)$, $\mu_2(x, y, z)$	Movement PDFs for first and second moment dynamics, L^{-d}
$b_1(x)$, $b_2(x, y)$	Directional bias for first and second moment dynamics, dimensionless

Parameters

m	Intrinsic movement rate, T^{-1}
ρ	Intrinsic birth rate, T^{-1}
d	Intrinsic death rate, T^{-1}
α	Strength of interaction for movement rate, T^{-1}
β	Strength of interaction for directional bias, L
γ	Strength of interaction for birth rate, T^{-1}
σ_w	Spatial range of interactions for movement rate, L
σ_v	Spatial range of interactions for directional bias, L
σ_ρ	Spatial range of interactions for birth rate, L
σ_{μ_ρ}	Spatial range of dispersal PDF, L
λ_μ	Rate parameter of PDF for movement distance, L^{-1}
σ_μ	Spatial range of PDF for movement distance, L
δr	Bin width for PCF, L
Δ	Grid spacing for discretisation of spatial displacement $\boldsymbol{\xi}$, L
ξ_{max}	Maximum distance of ξ_1, ξ_2 for computing $Z_2(\boldsymbol{\xi})$, L

Chapter 1

Introduction

1.1 Motivation and Objectives

The ability of cells to migrate as a collective, either to local sites or distant parts of the body, is fundamental for a number of physiological processes. For example in wound healing there is an orchestrated response involving the migration of various cell types to the wound site in order to combat infection and instigate tissue regeneration [1, 2]. The behaviour of cells undergoing collective movement is governed by a highly complex network of biochemical and biophysical factors acting across a range of scales. When the mechanisms regulating cell movement break down or are disrupted the outcome is either a lack of migration or migration to improper sites, consequently leading to defects which in the worst case may be lethal. In cancer the transition of a tumour from a benign state to a malignant state is marked by tumour cells acquiring the ability to invade surrounding tissues and metastasise to distant parts of the body [3].

Achieving a comprehensive understanding of collective cell movement is a hurdle that must be overcome if new and effective therapeutic strategies targeting migration are to be developed [4, 5]. Therapies that promote movement would enhance the synthesis of artificial tissues [6] while movement-inhibiting drugs could slow or halt the spread of a malignant tumour [7]. Mathematical modelling has proved a useful tool for the substantial task of dissecting the intricacies of collective movement and can often provide explanations to problems for which an experimental approach alone is insufficient [1, 8, 9].

Gaining insight into collective cell movement at different scales is important. A highly dynamic actin cytoskeleton comprises the core component of the intracellular machinery that facilitates a cell's migration [10]. At a higher organisational level, extracellular chemical signals, cell-cell adhesions and interactions with the extracellular matrix all contribute to the movement of cells [11]. The combined effect of these factors has implications for the moving cell population at a tissue level. For example, interactions occurring between neighbouring

cells can generate spatial structure which will, in turn, have significant consequences for the population-level dynamics. Thus a mathematical model of collective movement should be designed to predict behaviour across multiple biological scales [9].

In addition, a theoretical model for collective cell movement should be compatible with experimental data at different scales. Data generated from cell migration experiments may consist of measurements made at the scale of individual cells, for instance cell tracking techniques can be used to monitor the trajectories of single cells migrating across a surface [12]. Other measurements may be taken on a larger scale such as that of a whole population. Experimental data involving average cell densities or the speed of invading fronts of cells fall into this category [1]. Together, these requirements have led to the development of multi-scale models for collective movement [9, 13].

In this thesis we develop a mathematical model that accurately predicts the behaviour of cells undergoing collective movement. Our model acts on two scales: a microscopic scale which describes the behaviour of migrating cells at the level of an individual, and a macroscopic scale that describes the emergent dynamics of the population as a whole. At the microscopic level where cell behaviour is governed by stochastic forces, we employ an individual-based model (IBM) to predict how short-range interactions, such as crowding effects, influence collective movement. In particular, we incorporate a mechanism which allows a cell's direction of movement to be determined by the degree of crowding in its neighbourhood. This local directional bias is representative of attractive or repulsive forces occurring between cells and generates spatial structure in the population. The IBM is constructed using a lattice-free framework where cells are represented as individual agents that can wander freely across a continuous domain, rather than being constrained to discrete lattice sites. As we will discuss, this lattice-free approach leads to more realistic cell behaviour compared to a lattice-based approach. In order to investigate how the underlying individual-level information propagates to a larger scale we use our IBM to derive a corresponding description in terms of the dynamics of spatial moments. Unlike the majority of cell movement models, our spatial moment model accounts for the presence of spatial structure and allows us to explore population-level properties such as the dynamics of the average density of single cells and the average density of cell pairs. The model will be designed such that it is compatible with a variety of cell types and experimental situations. We use *in vitro* data generated from experiments with 3T3 fibroblast cells to parameterise our model and show that it is capable

of generating similar spatial structure to that observed in fibroblast populations. Finally, we extend the model to incorporate cell birth and death to consider the role these processes play in determining collective behaviour.

1.2 Thesis Structure

Chapter 2 (Background) Some key concepts of collective cell movement are introduced and we carry out a review of the relevant biological and mathematical modelling literature. In particular, the importance of spatial structure for collective cell behaviour is discussed.

Chapter 3 (Individual-Based Model) Presents the derivation of an individual-based model (IBM) for collective cell movement through one-dimensional space. The model allows interactions with neighbouring cells to determine an individual's rate and direction of movement. We also discuss how a pair-correlation function (PCF) can be computed from a particular configuration of cells, to quantify the extent of spatial structure present in the population.

Chapter 4 (Spatial Moments Model) The one-dimensional IBM is used to derive corresponding population-level approximations by employing theory from the dynamics of spatial moments. Equations for the rates of change of spatial moments up to second order are derived. This allows the effects of spatial structure to be accounted for at a population level. The model can be applied in a spatially non-homogeneous setting, however we only solve for a homogeneous space.

Chapter 5 (Numerical Results for One-dimensional Model) Presents numerical results for the one-dimensional model for a simple case where cells are distributed homogeneously throughout space. The different spatial structures generated by neighbour-dependent movement are discussed and we analyse the performance of the spatial moment model as an approximation to the underlying IBM.

Chapter 6 (Two-dimensional Model and *In Vitro* Data) Extends the model for collective movement in two spatial dimensions. We present and discuss numerical results for the extended model, again for a spatially homogeneous case. The analysis of experimental data is also dealt with in Chapter 6 and the data is used to parameterise and validate

the two-dimensional model.

Chapter 7 (Model of a Birth-death-movement Process) The IBM and spatial moment models are extended to incorporate cell birth and death as well as movement. Numerical results from the extended model are presented and discussed. In particular, we explore how these different cell processes generate spatial structure and what effect this structure has on the dynamics of the population.

Chapter 8 (Concluding Remarks) Summarises the outcomes and importance of the work. We also discuss some possible directions for future research.

1.3 Thesis Publications

A paper presenting the derivation and results for the one-dimensional model, from Chapters 3, 4 and 5, has been published by Journal of the Royal Society Interface [14]. The two-dimensional model, numerical results and analysis of experimental data from Chapter 6, are presented in a paper which has been accepted for publication by PeerJ [15].

Chapter 2

Background

2.1 Biological Background

Cell movement plays a fundamental role in a number of physiological processes, where it is highly regulated. For example, controlled cell migration facilitates the growth of new blood vessels from existing vasculature (angiogenesis) [3], embryonic development [16], wound healing [17], and is also a component of the immune response [11]. Abnormal migration is known to contribute to pathological processes such as vascular disease, chronic inflammatory syndromes and cancer [6, 11].

While the role of migratory cells varies between different cell types, there is evidence to suggest the underlying mechanisms controlling migration are similar. The key components of these mechanisms are also evolutionarily conserved between different species [16]. In general, migratory cells sense and respond to external migration-inducing signals by undergoing internal changes involving components of their cytoskeleton, allowing them to move through their external environment [6]. Cells may move up a chemical signal concentration gradient by chemotaxis or up a gradient of cellular adhesion sites by haptotaxis, in addition to undergoing random migration [3].

The intracellular machinery that facilitates the migration of a single cell through its environment is built from a highly dynamic actin cytoskeleton and involves a cyclical progression through four steps, namely polarisation, protrusion, adhesion and retraction [6, 11, 16]. In response to an external migratory signal, a cell will first undergo polarisation, involving specification of the front of the cell (the leading edge) and the rear (the trailing edge) so that it now faces a certain direction [6, 16]. The next stage in the cycle is the protrusion of the cell's leading edge in the direction of migration. These finger-like projections are called pseudopodia and arise through rearrangements in cytoskeletal components [10]. In the third stage, the protrusions at the leading edge undergo adhesion by fusing to the extracellular substrate. This provides the cell with strong and stable external contacts from which it can push off,

and enables the cell to propel itself forwards through the surrounding medium [11]. The final stage in the migration cycle is the retraction of the rear of the cell. The trailing edge retracts into the cell body while adhesion contacts at the rear are released, allowing the cell to pull itself forwards, completing the cycle [6, 11, 16].

While many cell types migrate as individuals, for instance lymphocytes involved in the immune response, others undergo collective migration [7, 11]. Migrating epithelial cells can form a tightly associated continuous sheet, within which the cells align and polarise themselves such that they share a unified direction of movement [18, 19]. Such sheets often have a clear invasion front and are observed in physiological processes such as wound healing [8, 11]. In addition to the protrusions extended by individual cells, multi-cellular projections guided by leader cells occur [18,20]. Other forms of collective migration include sprouting and branching, for example the growth of new blood vessels from pre-existing vasculature in angiogenesis [3], and streams, such as neural crest cells involved in vertebrate embryogenesis [16]. In some cases, cells also migrate as free groups, as is observed in ovarian border cell clusters of *Drosophila melanogaster* [11]. A desire to understand how these moving collectives arise and how it is possible for such large numbers of individuals to coordinate their movement has led to extensive studies into the interactions occurring between migrating cells [12, 20, 21].

Interactions occurring between collectively moving cells can be mechanical. For instance contacts via cell-cell junctions provide a physical coupling between neighbouring cells. Tensile stresses are transmitted through these intercellular junctions as a result of traction forces exerted between migrating cells and their surrounding substrate [18–20, 22]. Non-mechanical interactions, such as the ability of certain cell types to detect diffusible chemical factors, also impact collective behaviour [12]. Chemical signals can enhance [23] or inhibit a cell's rate of movement, or affect its direction of movement [24]. Quorum-sensing in bacterial populations enables cells to sense their population density by detecting the concentration of a cell-secreted chemical signal in the surrounding medium. At high cell densities the increased concentration of signal induces a physiological response, such as the initiation of swarming movement in *Burkholderia cepacia* [25–27]. Similar behaviour has been observed in eukaryotic cells for example the induction of aggregation in the slime mould *Dictyostelium discoideum* [25, 28].

Some interactions within cell collectives act as attractive forces to drive cells together. These include cell-cell adhesions [29] and cell-released chemoattractant signals which give rise to chemotaxis towards the signal source [25, 30]. On the other hand, movement in response

to a cell-secreted chemorepellant can have a repulsive effect where cells are biased to move away from their nearest neighbours [24,31]. Other interactions affecting cell motility include crowding effects which can occur at high cell densities. One such effect is contact inhibition of locomotion whereby, after colliding with another individual, a moving cell will slow down then alter its direction of movement in an attempt to avoid future collisions [12,32]. This behaviour has been observed experimentally in invading malignant fibroblast cells [32] and in wound healing [1,33].

The short-range interactions experienced by cells often lead to self-generated spatial structure which can in turn have a significant impact on the dynamics of the cell population [34–36]. For instance, many cell types are known to form clusters or aggregates as a result of attractive interactions [37,38]. Examples include breast cancer cells [29] and hepatocyte-stellate aggregates [37]. Others, such as retinal neurons [31,39], arrange themselves into patterns that minimise their proximity to neighbouring cells. This behaviour can be observed in cell populations cultured *in vitro*, however it is not always obvious which underlying mechanisms are responsible for pattern formation, particularly when multiple types of interaction are involved [37]. Therefore there is good motivation for the development of techniques that give more insight into the effects of these mechanisms.

2.1.1 Experimental Techniques

The migration of cells in two dimensions has been studied extensively *in vitro* and a range of experimental techniques have been developed for this purpose [13,40,41]. Time-lapse photography and image analysis methods are commonly used in conjunction with cell migration assays to capture the distribution of cells across a substrate at discrete time intervals and allow measurements to be taken at these fixed time points. Microscopy techniques generate high quality images at the level of individual cells in addition to the population scale [1,12,42,43].

One such assay is a cell scrape (or wound healing) assay which initially involves the growth of cells to form a confluent monolayer. A small section of the monolayer is then scratched away to leave a cell-free gap or wound. The cells are monitored as they invade into the cell-free gap to close the wound and an invasion front is often observed. Population-scale measurements such as the invasion rate of the leading edge can be taken [8,13,42]. A similar but more reproducible experiment is a two-dimensional barrier assay, which typically involves attaching a hollow circular barrier to the surface of the culture dish and introducing cells

into the internal barrier chamber. The barrier is then removed and the cells are monitored as they spread into the cell-free regions of the dish [40, 41]. Monolayer formation assays are a third type of experiment used in studies of cell population growth. A sparse population of cells is initially uniformally distributed across a culture dish, then allowed to migrate and proliferate until the population reaches confluence. This type of experiment allows large-scale measurements of population growth rate and maximum cell density to be taken. At a smaller scale, individual cell trajectories can also be tracked and their rate and direction of movement monitored over time [1]. In addition, cell-cell contact effects such as contact inhibition of migration and contact inhibition of proliferation can be observed experimentally [32, 44]. Techniques such as the Transwell Invasion assay are used for *in vitro* studies of cell migration through three-dimensional space. In a Transwell assay, cells are placed on the top surface of a cell-permeable membrane and allowed to invade through the membrane, usually towards a chemoattractant [45].

A number of different cell types can be used in migration assays depending on their focus. For instance, epithelial, endothelial or fibroblast cells are often used in studies of wound healing [1, 19, 33] while invasive sarcoma, glioma or melanoma cells have been used in cancer studies [32, 45]. Invasion assays can also be used to test the effect of migration-inducing or -inhibiting agents [45]. For example, the cell-secreted protein Slit has been shown to have a repulsive effect on the migration of neurons [24] while the chemo-attractant acidic fibroblast growth factor (aFGF) has been used to study the role of endothelial cell migration in angiogenesis [46]. The data generated from experimental studies such as these can be used for the parameterisation of mathematical or computational models for cell movement [13, 33, 43, 46].

2.2 Mathematical Modelling Background

Collective movement arises in a variety of contexts from insect swarming [47], to pedestrian flow in traffic [48], to cell migration in tissue [11]. The desire to better understand this type of behaviour has motivated the development of mathematical models for collective movement. A variety of approaches have been applied to the modelling of cell migration across a range of biological scales and these can be divided into two broad categories [5, 9, 13]. Firstly, microscopic models which describe collective cell behaviour at the level of individuals and

where the movement of cells is governed by a stochastic process [46]. Secondly, macroscopic models which provide insight into collective behaviour at the level of a cell population and are deterministic [5, 42, 49].

Ideally, a model for collective movement should incorporate both microscopic and macroscopic behaviour and multi-scale models for collective movement are now well discussed in the literature [3, 9, 13]. In many cases a PDE is derived initially for the population-level behaviour and then undergoes a decomposition to a continuous-time master equation to give insight into the individual-level properties [3, 13]. However this trend has now shifted to a "bottom-up" approach where a microscopic model is developed first and then, by first writing conservation statements and taking the continuum limit, is used to derive a population-level description [4, 5, 49].

2.2.1 Microscopic Models

Probabilistic models are used to model migration at an individual-level, from isolated single cells to large numbers of individuals [50]. The earliest stochastic models of cell migration tended to focus on isolated or small numbers of individuals, such as a model of bacterial chemotaxis proposed by Stroock (1974) [50] or a correlated random walk model for slime-mold amoebae movement developed by Hall (1977) [51]. The strategy of using random walks [52] to describe the movement of eukaryotic cells, such as leukocytes and fibroblasts, found in more complex multicellular organisms was popularised by models developed by Peterson & Noble (1972) [53], Alt (1980) [54] and by Dunn & Brown (1987) [55]. In 1991, Stokes & Lauffenberger proposed a model for the migration of endothelial cells involved in angiogenesis that incorporated random movement as well as chemotaxis [46].

More recently, stochastic models for simulating the movement of large numbers of individuals have been developed [33, 42]. These include individual-based models (IBMs) or agent-based models where each cell is represented as an individual agent and the movements of all agents are tracked over time [56]. IBMs can be divided into two classes. In the first lattice-based class (which includes cellular automata), movement is constrained to a regular or irregular lattice with pre-defined grid points. The lattice framework restricts the location of agents to discrete grid sites and agents can only move in a discrete set of directions to neighbouring sites [5, 42]. Cell interactions and crowding effects can be incorporated into lattice-based models through exclusion processes which ensure that only a single agent can

occupy any one lattice site [43, 44]. Other individual-level factors, such as cell-cell adhesion which has a role in malignant invasion, have been explored using lattice-based models [29]. Lattice-based approaches have been shown to provide a good match to behaviour observed experimentally, particularly for low cell densities [13, 42]. However, one drawback to this approach is that the carrying capacity (i.e. the maximum number of cells that can occupy the domain) is a property that is artificially pre-determined by the number of grid points available in the chosen lattice. Also, at high cell densities where cells are tightly packed into the available space, an unrealistic spatially-regular configuration of cells arises due to their alignment along the lattice [44].

The second class are lattice-free models where cells are not restricted to discrete lattice sites but can wander freely across a continuous domain. These have been shown to result in a more realistic spatially-irregular configuration of cells [57]. In a lattice-free framework the maximum cell density is an emergent outcome of the model as opposed to being pre-determined by the size of the lattice [44]. A drawback of employing a lattice-free strategy is that running simulations can become more computationally intensive at high cell densities than using an equivalent lattice-based approach [57].

Recent research has highlighted the importance of volume exclusion, the concept that the cells themselves occupy space in the domain and may obstruct other individuals from occupying the same space [58, 59]. In lattice-based approaches, volume-exclusion is incorporated by assuming that only one individual can be present at each lattice site and the lattice-spacing corresponds to a cell diameter. Methods for deriving population-level cell models from lattice-based IBMs can be found in the literature [42, 49]. Different approaches to deriving macroscopic descriptions from lattice-free IBMs with volume exclusion have also been discussed. One method involves defining individuals as hard spheres with fixed diameter, around which may lie an exclusion area that other individuals cannot occupy. In this case, the domain can be divided into strips and an approximate population-level description found by writing a conservation statement for the density of agents in each strip [29, 44, 57]. Other models describe a hard spherical core but apply an external force which can be representative of either inter-particle or external interactions [60]. Frameworks for deriving population-level descriptions from lattice-free IBMs in one-, two- and three-dimensions, and incorporating a mechanism for global directional bias, have also been described [58, 59].

Simulations of IBMs produce synthetic data that can be compared to experimental im-

ages [43] and may shed some light on the underlying mechanisms responsible for emerging spatial structure [37], however they are quite intractable mathematically. Deriving a formal mathematical representation gives more insight into the population-level dynamics of such systems and provides scope for a more rigorous analysis [61].

2.2.2 Macroscopic Models

For simplicity, macroscopic models often neglect the effects of spatial structure on a cell population. They typically deal with a density of individuals that has been averaged over space and explore the evolution of this average density over time. Such models are termed 'mean-field' and assume that individuals are well-mixed or undergo long-range interactions. 'Local mean-field' models, such as reaction-diffusion equations, allow the average density of individuals to be expressed as a function of the location in space however they still tend to ignore the effects of small-scale spatial structure on the population [61]. For example, the Fisher-Kolmogorov equation [62, 63] has been used to describe both cell migration, incorporated in a diffusion term, and proliferation in the form of a logistic growth function [8, 43]. The use of reaction-diffusion equation models was pioneered largely by Keller & Segel (1970) when they proposed a partial-differential equation (PDE) for modelling chemotaxis in bacteria [64]. More recently, models for cell movement have been described in a range of contexts including angiogenesis [3], wound healing [1, 8] and malignant spreading [4, 5].

In reality the short-range interactions experienced by cells can often lead to self-generated spatial structure which can in turn have a significant impact on the dynamics of the cell population [34–37]. As mean-field models do not account for these local interactions they do not always provide a good representation of real behaviour [65]. Spatial moment theory, originally developed in statistical physics [66–69], can be used to investigate the effect of spatial structure on population-level dynamics [61, 65, 70, 71]. The average density of individuals dealt with in mean-field models is the first spatial moment which holds no information on small-scale spatial structure. One way to access such information is to consider the second spatial moment, the average density of pairs of cells, expressed as a function of the distance r between them. The second moment is often dealt with as a pair correlation function (PCF) $C(r)$ in which it is normalised by the square of the first moment such that in the absence of spatial structure $C(r) \approx 1$. Figure 2.1 shows the PCF for three spatial point patterns. Figures 2.1(a)-(c) can each be considered as a snapshot in time from a realisation of an IBM.

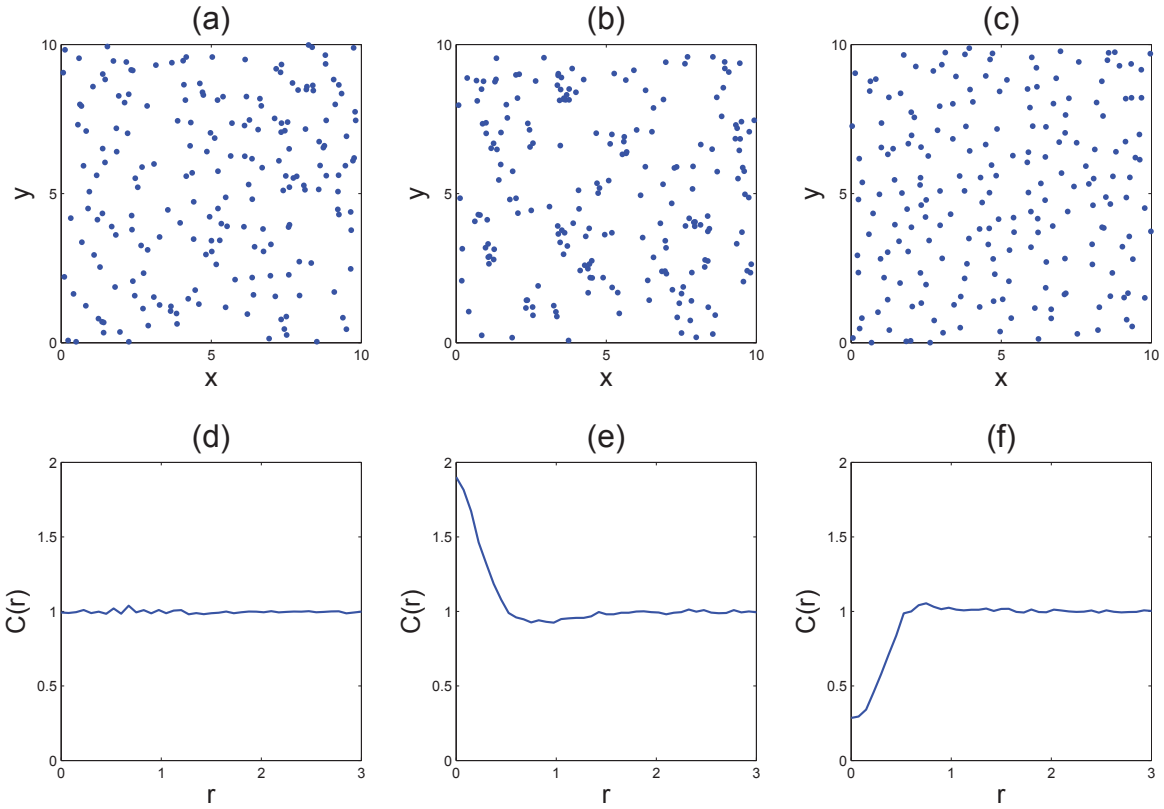


Figure 2.1: Three different spatial patterns in two-dimensional homogeneous space (a)-(c) and the pair correlation functions $C(r)$ for each pattern (d)-(f). (a) Poisson spatial pattern (or complete spatial randomness) in which there is no spatial structure present; (b) Cluster spatial pattern; (c) Regular spatial pattern.

Figure 2.1(a) describes a spatial Poisson point process (sometimes referred to as complete spatial randomness) in which all locations of individuals are independent of one another. For this case $C(r) \approx 1$ and no spatial structure is present (Fig. 2.1(d)). Figure 2.1(b) shows a cluster pattern, in which pairs of cells are more likely to be found in close proximity. This corresponds to $C(r) > 1$ for short distances r as shown in Figure 2.1(e). The opposite effect can also arise, whereby cells are less likely to be found close together resulting in a regular pattern (Fig. 2.1(c)). Figure 2.1(f) shows that $C(r) < 1$ at short distances r for this type of spatial structure [72].

In previous studies, PCFs have been calculated from experimental images to quantify the extent of spatial structure in live cell populations which adopt Poisson [73], cluster or regular patterns [38, 74] to varying degrees. For instance, time-lapse imaging of *in vitro* cell migration assays, such as circular barrier assays [73] and scratch assays [75], generates data in two spatial dimensions. Image analysis techniques can then be employed to measure the distances between cell pairs and this data used to calculate a PCF. PCFs have also been

used alongside experimental data to give insight into the mechanisms responsible for pattern formation [37].

Exploring the dynamics of the second moment can provide insight into how the spatial structure is changing over time and whether the state of the system converges. The dynamics of the third moment, the average density of triplets, can be derived to provide further information still, and so on up to the n-th moment however the descriptions of the dynamics become increasingly complex for higher moments [65, 70]. In order to solve the dynamical system a suitable closure is also required because the dynamics of each moment depend on the next moment in the hierarchy. Mean-field models employ a first order closure (the mean-field assumption) in which the second moment is assumed to equal the first moment squared. In other words, it is assumed that individuals encounter one another in proportion to their average density. Thus, in the mean-field assumption any spatial information that was held in the second moment is lost. However, models which close the dynamics at higher orders retain the spatial information held by the second moment. At second order a number of different closures are possible, for example the Kirkwood superposition approximation [76], which is commonly employed in cell movement models [71,77,78]. Murrell, Dieckmann and Law (2004) discussed the suitability of various closures and found that the asymmetric power-2 closure also performs well in general [79].

The types of local interactions inherent to migrating cell populations are also of importance in other contexts, such as in animal or plant communities. Many of the modelling tools that employ spatial moment theory were developed for ecological problems [80–83]. Models for dynamics of the second moment which incorporate mechanisms for birth, growth, death and movement (either in isolation or combination) have been derived. One of the earliest ecological studies employing spatial moment theory was that of Bolker and Pacala (1997), who explored growth and competition in plants using the dynamics of spatial moments with density-dependent mortality [80]. This study was extended by Law, Murrell and Dieckmann (2003), who showed that self-generated spatial structure can cause the growth of a population to differ significantly from the non-spatial dynamics [65]. Their work revealed that both the rate of population growth and the steady state average density can be greatly affected by spatial structure, generated from the dispersal of offspring or local competition, as shown in Figure 2.2 [65].

Lewis and Pacala (2000) considered the invasion of populations by incorporating density-

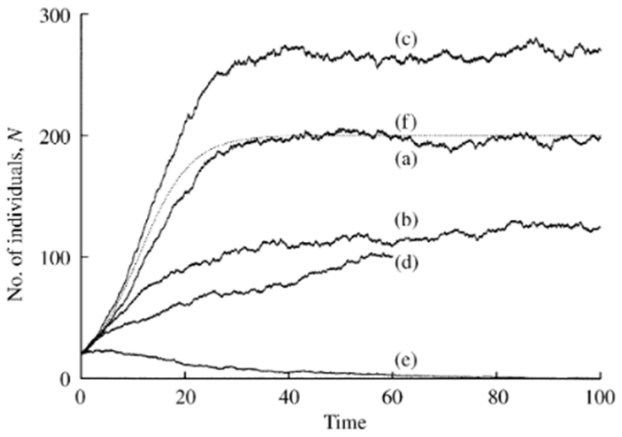


FIG. 1. Number of individuals, N, in an arena of unit area over time from IBMs (individual-based models) with contrasting kernel parameters for dispersal S_m and competition S_w. All simulations began with 20 individuals randomly dispersed across the unit arena; the time series shown are averages of 20 realizations of the IBMs. Kernel parameters are: (a) $S_m = 0.12$, $S_w = 0.12$; (b) $S_m = 0.02$, $S_w = 0.12$; (c) $S_m = 0.12$, $S_w = 0.02$; (d) $S_m = 0.04$, $S_w = 0.04$; the simulation is stopped at $t = 60$, after which it overlaps with (b); (e) $S_m = 0.02$, $S_w = 0.02$. (f) Growth of an equivalent nonspatial logistic population, with density (number per unit area) shown as the dotted line. Parameters held constant throughout are as follows: b (per capita birth rate) = 0.4, d (intrinsic per capita death rate) = 0.2, and $d' = 0.001$.

Figure 2.2: The effects of spatial structure on the dynamics of a growing population, studied by Law, Murrell and Dieckmann (2003) [65]. Spatial structure was generated both by the short-range dispersal of offspring and local competition. Figure used by permission of Ecological Society of America.

dependent birth into the moment dynamics [83]. The relationship between spatial arrangement and plant size distribution have also been explored [84]. Models for animal populations undergoing movement have been derived both for the case where movement is dependent on local neighbourhood and the independent case [70, 85, 86]. These have also been extended to describe the role of spatial structure in predator-prey relationships [86].

Spatial moment models often assume a homogeneous spatial distribution (the pattern is stationary over space). Here we use the term spatially homogeneous to refer to a situation where the probability of finding an individual in a given small region does not depend on the location in space. This is the same as assuming that the spatial structure observed in a small window within a larger space is independent of the position of the window, i.e. it has translational invariance. In terms of spatial moments this corresponds to a first moment that is constant over space, while the second and third moments can be expressed in terms of displacements between pairs of agents as opposed to agent locations [61, 72, 87].

In some cases of collective cell movement it is necessary to consider a non-homogeneous setting, where the average density of cells is higher or lower in certain regions. For example, a non-homogeneous initial condition would be required for the modelling of cell invasion assays in which moving fronts of cells are formed [43, 88]. However, solving the spatial moment dynamics up to at least second order is more complicated for the non-homogeneous case and consequently has received significantly less attention than simpler homogeneous systems [61]. Middleton et al. [71] studied a lattice-free cell movement model incorporating a local direc-

tional bias and employing the second spatial moment for a non-homogeneous distribution of cells. This work extended that of Newman and Grima (2004) [89]. Their IBM is based on a Langevin equation defining each cell's velocity as the sum of the interaction forces from neighbours plus a noise term. Thus, neighbour interactions always affect both speed and direction of movement. Comparable population-level models for cell birth-death-movement processes in lattice-based frameworks have been developed, in which pair-approximation methods are used to account for spatial correlations in discrete space. These have been described for a homogeneous setting [78] and have also been extended to account for a global directional bias in non-homogeneous space [88].

In an ecological context, spatial moments models in non-homogeneous space have been developed for the invasion of a population through a birth/dispersal mechanism, incorporating local interactions, but in the absence of movement [82, 83]. Murrell and Law (2000) [70] described a model for local density-dependent animal movement, in which the self-generated spatial structure of the population assumed homogeneity but movement occurred through a fixed heterogeneous landscape. While these models were relatively case-specific, a more general model incorporating terms for density-dependent birth, death and movement was recently derived by Plank and Law (2014) [61] for a spatially non-homogeneous case.

Other methods for describing the dynamics of spatial point processes, which do not require a closure assumption, are also discussed in the literature. Stochastic differential equations, such as Langevin-type equations, capture fluctuations arising due to short-range interactions via a noise term and can be used to investigate spatio-temporal patterns at different scales [90]. Blath et al. [91] analysed a stochastic, lattice-based model using stochastic differential equations to explore whether spatial structure could give rise to coexistence between two competing species. A closed system of equations for the whole hierarchy of moments was derived by Ovaskainen et al. [92] using techniques from Markov evolutions and a perturbation expansion around the spatial mean-field model. Bruna and Chapman [60] employed a perturbation method to describe the dynamics of moving particles by using matched asymptotic expansions in a small parameter $\epsilon \ll 1$ to derive a nonlinear diffusion equation.

Chapter 3

Individual-Based Model

In this chapter we describe an IBM for collective cell movement in one-dimensional space. At this scale we are interested in the stochastic interactions occurring between individual cells. Our model is constructed using a lattice-free framework where cells are represented as individual agents that can wander freely across a continuous domain, rather than being constrained to discrete lattice sites. We incorporate short-range interactions by allowing an individual's rate and direction of movement to depend on the degree of crowding in its neighbourhood. This local directional bias is representative of attractive or repulsive forces occurring between cells, such as in response to a chemoattractant or repellant, and generates spatial structure in the population. In the next chapter, we use our IBM to derive a corresponding description in terms of the dynamics of spatial moments.

Motile cells possess dynamic cytoskeletons which enable them to change their shape and flex around neighbouring cells [10, 32]. For this reason, cells rarely form perfect spheres and it can be difficult to accurately estimate their average diameter. Therefore, we choose not to use a hard-core approach to account for volume exclusion but instead represent the location of a cell by its coordinates in space. Rather than explicitly excluding neighbours from the space surrounding an individual, we consider a kernel (a Gaussian function) which weights the strength of an individual's interaction with its neighbours. The kernel width corresponds to the range over which an individual will affect other cells in its neighbourhood and can be considered a proxy for average cell diameter.

The majority of cell biology experiments are carried out in two or three spatial dimensions. However, numerically solving the moment dynamics up to second order can become quite complicated in higher dimensions and so here we consider a simpler case of movement through one-dimensional space. We will show that the one-dimensional model can still capture the qualitative traits of spatial structure inherent to populations in which short-range interactions are important, i.e. clustering and regular patterns observable in cell populations cultured *in vitro*. In Chapter 6 we extend the IBM for movement in two spatial dimensions.

3.1 Setting

The stochastic theory employed in many microscopic models of collective movement revolves around the concept of Markov processes which we will now briefly describe in a general context.

3.1.1 Poisson and Markov Processes

The *Poisson process* is a stochastic process that counts the number of events that occur in a given time interval. The waiting time between a pair of consecutive events is exponentially distributed with rate λ and is independent of all other waiting times. In order to give a formal definition of this we now need to introduce some notation. We denote a Poisson process by $\{X(t) : t \geq 0\}$ where the random variable $X(t)$ represents the state of the process at time t. Time can evolve via discrete time jumps in which case t belongs to the finite positive set $\{t_0, t_1, \ldots, t_n\}$. Alternatively it can evolve continuously for $0 \leq t \leq \infty$. A Poisson process has a positive rate λ and a discrete state space S which is a set of non-negative integers. In addition, it has the following properties:

i $X(0) = 0$

ii For any times $t_0 < t_1 < t_2 < \ldots < t_n$ the time increments are independent, that is

$$X(t_1) - X(t_0), X(t_2) - X(t_1), \ldots, X(t_n) - X(t_n - 1)$$

are independent random variables. The time increments are also stationary such that the probability distribution of an increment $X(t) - X(s)$ depends only on the length of the time interval $t - s$ for any $0 \leq s < t$.

iii The random variable $X(s + t) - X(s)$ for any $0 \leq s < t$ has a Poisson probability distribution:

$$\mathbb{P}\{X(s+t) - X(s) = k\} = \frac{(\lambda t)^k e^{-\lambda t}}{k!}. \tag{3.1}$$

The first moment (mean μ) and second moment (variance σ^2) of a Poisson process of rate λ are

$$\mathrm{E}[X(t)] = \lambda t \quad \text{and} \quad \mathrm{Var}[X(t)] = \lambda t,$$

respectively. By this definition, for a small waiting time τ we see that the probability of a single event occuring in the time interval $[t, t+\tau]$ is

$$\begin{aligned}\mathbb{P}\{X(t+\tau) - X(t) = 1\} &= \frac{(\lambda\tau)^1 e^{-\lambda\tau}}{1!} \\ &= \lambda\tau + O(\tau^2). \end{aligned} \tag{3.2}$$

The likelihood of more than two events occurring in an interval $[t, t+\tau]$ corresponds to a probability

$$\mathbb{P}\{X(t+\tau) - X(t) \geq 2\} = O(\tau^2), \tag{3.3}$$

which implies that two or more events cannot occur simultaneously.

A Poisson process can be classed as either *homogeneous* or *non-homogeneous*. In a homogeneous Poisson process the rate λ is constant over time. When the rate λ varies over time such that $\lambda = \lambda(t)$ we call this a non-homogeneous Poisson process. In this case the expected number of events occurring in a time interval $(s, t]$ is Poisson distributed and the mean given by

$$E(X(t) - X(s)) = \int_s^t \lambda(u)\mathrm{d}u.$$

Finally, the combination of n independent Poisson processes with associated rates λ_j for $j = 1, \ldots, n$ is itself a Poisson process with rate $\lambda = \sum_{j=1}^n \lambda_j$.

A Poisson process is the simplest form of *Markov process* which is a stochastic process that satisfies the *Markov property*. This is a memoryless property in which the future behaviour of a process depends only on the current state and not on any states that may have occurred in the past. A Markov process $\{X(t) : t \geq 0\}$ has a state space S which may be either discrete or continuous but we will proceed with our definition using the notation for a continuous state space where $S \subseteq \mathbb{R}$. The Markov process satisfies the Markov property if for all $s, t \geq 0$

$$\begin{aligned}&\mathbb{P}(X(s+t) = y | X(s) = x, \{X(u) : 0 \leq u < s\}) \\ &= \mathbb{P}(X(s+t) = y | X(s) = x) \end{aligned} \tag{3.4}$$

for all $x \in S$, $y \in S$ [93].

3.1.2 The Gillespie Algorithm

The Gillespie algorithm [94] or stochastic simulation algorithm (SSA) is a computational method for numerically simulating the time evolution of a stochastic system. In general, the algorithm can be applied to a system defined at a time t by state variables $\mathbf{x}(t) = (x_1(t), x_2(t), \ldots, x_n(t))^T$, where $\mathbf{x}(t) \in S$. The system is updated to a new state by undergoing one of m possible transitions with associated transition rates λ_i, $i = 1, \ldots, m$. The algorithm can be broken down into four steps as follows:

1. The first step is initialisation, where time is set to $t = 0$ and the initial state $\mathbf{x}(0)$ of the system is defined.

2. For the current system state $\mathbf{x}(t)$ a transition rate $\lambda_i(\mathbf{x})$ is calculated for each possible type of transition, $i = 1, \ldots, m$, and the total transition rate

$$\lambda = \sum_{i}^{m} \lambda_i(\mathbf{x}),$$

   is found.

3. Monte Carlo techniques are then employed to generate two random numbers. Firstly a random number τ, which is the waiting time to the next transition event, is drawn from an exponential distribution with mean $1/\lambda$. The next transition to occur is then chosen with a probability proportional to the rate of that transition. To achieve this a random number u is drawn from a continuous uniform distribution on $[0, \lambda]$. If $0 \leq u \leq \lambda_1(\mathbf{x})$ then the transition with associated rate $\lambda_1(\mathbf{x})$ is the next to occur; if $\lambda_1(\mathbf{x}) < u \leq \lambda_2(\mathbf{x})$ then the transition with rate $\lambda_2(\mathbf{x})$ occurs; and continuing in a similar fashion until, if $\lambda_{m-1}(\mathbf{x}) < u \leq \lambda_m \mathbf{x})$ then the transition with rate $\lambda_m(\mathbf{x})$ occurs.

4. Time is advanced by the random time increment, from t to $t + \tau$. The state of the system is then updated to reflect the occurrence of the chosen event.

5. Finally, the method is repeated from the second step until the time limit of the simulation has been reached or until an update to the system state prevents the occurrence of further events.

The method was originally developed for the simulation of spatially homogeneous chemical systems involving reactive molecular species. In this case, the next chemical reaction to occur

would be chosen at step 3 and the system would be updated in step 4 by altering the molecule count of the species involved in that reaction accordingly.

3.2 One-Dimensional Individual-Based Model

We consider the collective movement of n individuals through a one-dimensional continuous finite domain $\Omega = \{x_l \leq x \leq x_r, x \in \mathbb{R}\}$. Our IBM is a continuous time Markov process model, thus the future state of the system depends only on the given current state and is not influenced by any states that may have arisen in the past, as described in section 3.1.1. The state of the system $\mathbf{x}(t)$ at a particular time t is

$$\mathbf{x}(t) = (x_1(t), x_2(t), \ldots, x_n(t))^T, \tag{3.5}$$

where $x_i \in \Omega$ is a coordinate representing the location of an individual cell i. As time advances, a cell updates its location in the space by undergoing a movement event. This leads to a transition in the state of the system which can be expressed as follows:

$$\mathbf{x} \mapsto \mathbf{x} + \mathbf{e}_i r, \tag{3.6}$$

where $\mathbf{e}_i$ denotes a basis vector with 1 at the i^{th} coordinate and zeros elsewhere. The rate density (which hereafter we will refer to simply as a rate) at which the system undergoes the transition denoted by Eq. (3.6) is $\psi_i(\mathbf{x})\mu(x_i, x_i + r)$, where $\psi_i(\mathbf{x})$ is the movement rate per unit time of cell i when the system is in a state $\mathbf{x}$. From here on we will drop the $\mathbf{x}$ argument from ψ_i for notational ease. The function $\mu(x_i, x_i + r)$ is a probability density function (PDF) for movement by a distance r.

Periodic boundaries are implemented at $x = x_l$ and $x = x_r$ such that when a cell moves over one boundary it reappears at the opposite boundary. This means that for an individual at x attempting to take a step r, the new location y of that individual will be

$$y = \begin{cases} x + r & \text{if } x_l \leq x + r \leq x_r \\ x + r + |x_r - x_l| & \text{if } x + r < x_l \\ x + r - |x_r - x_l| & \text{if } x + r > x_r \ . \end{cases} \tag{3.7}$$

3.2.1 Unbiased Movement

We simulate this stochastic process using the Gillespie algorithm, in which the time increment τ between events is exponentially distributed with mean $1/\lambda(t)$, where $\lambda(t)$ is the sum of all transition rates ψ_i for $i = 1, \ldots, n$ at time t. One of n possible transitions occurs at time $t+\tau$ and is chosen with a probability proportional to the rate of that transition. The movement distance r can then be determined according to $\mu(x_i, x_i + r)$. In the following description we make choices for the functions ψ_i and $\mu(x_i, x_i + r)$, however these can be easily adapted to suit different experimental situations.

The movement rate ψ_i comprises two terms: an intrinsic motility rate m, i.e. the rate at which an isolated cell would move, and a neighbour-dependent component. The latter term sums a contribution $w(z)$ from each of the other cells in the population, where $w(z)$ is a kernel weighting the strength of interaction between a pair of cells displaced by z. Therefore the movement rate for an individual at x_i with n neighbours at x_j is

$$\psi_i = \max(0,\ m + \sum_{\substack{j=1 \\ i \neq j}}^{n} w(x_j - x_i)). \tag{3.8}$$

This definition ensures that $\psi_i \geq 0$. Notice that we remove the effect of the individual i interacting with itself by excluding the case where $i = j$. For simplicity, we choose the interaction kernel $w(z)$ to be a Gaussian function

$$w(z) = \alpha \exp\left(-\frac{z^2}{2\sigma_w^2}\right), \tag{3.9}$$

where α and σ_w^2 determine strength and range of interaction respectively. The intrinsic rate m, interaction kernel $w(z)$ and overall motility rate ψ_i have dimensions T^{-1}.

This choice of kernel means that cells interact strongly with near neighbours but are not influenced by those further afield. For $\alpha > 0$, cell i's motility ψ_i is increased by the presence of close-lying neighbours. This type of interaction is relevant from a biological perspective, for example in collective movement involving cell types which release motility-enhancing diffusible signalling factors into their environment. The high concentrations of signals found at high cell densities can result in increased motility rates for cells in crowded regions [23,95]. On the other hand if $\alpha < 0$ then the presence of close-lying neighbours will reduce ψ_i. For instance, crowding effects such as contact inhibition of locomotion reduce motility at high

local cell densities [13, 32].

When a cell undergoes a movement event it takes a step of displacement r from x to y, drawn from a movement PDF $\mu(x, y)$. In the unbiased case where an individual's direction of movement is not affected by the presence of neighbouring cells, we define $\mu(x, y)$ to be a Laplace distribution

$$\mu(x, y) = \frac{\lambda_\mu}{2} \exp\left(-\lambda_\mu |y - x|\right), \tag{3.10}$$

where the mean step length taken by a cell is $1/\lambda_\mu$. This means cells are more likely to take short steps than undergo large jumps across the space and so is biologically reasonable [12].

3.2.2 Biased Movement

The model described so far allows simulation of collective movement in which an individual's motility is influenced by the cell density in its local neighbourhood, as can be observed experimentally [1, 12]. We now incorporate a local directional bias $b(x)$ such that the presence of neighbouring cells affects the direction of movement of an individual at x. The neighbour-dependent bias $b(x)$ sums a contribution $u(x, y)$ from each of the other cells in the population as follows:

$$b(x) = \sum_{j=1}^{n} u(x, x_j) \ , \tag{3.11}$$

where

$$u(x, y) = v'(y - x)\hat{h}(x). \tag{3.12}$$

and $v'(z)$ denotes the derivative of $v(z)$ with respect to z. Here, $\hat{h}(x)$ is a spatially-dependent weight placed on the cell at x, for example the effect of a spatially-variable environmental factor on direction of movement. In a case where the distribution of cells is homogeneous across space, we have $\hat{h}(x) = 1$. In equation (3.12), $v'(y - x)$ determines the effect of a cell at y on the direction of movement for an individual at x. In theory $v'(z)$ could be replaced by any real-valued kernel which weights the strength of interaction between a cell pair displaced by z. We choose $v(z)$ to be a Gaussian function

$$v(z) = \beta \exp\left(-\frac{z^2}{2\sigma_v^2}\right). \tag{3.13}$$

This means $v'(z)$ has positive and negative values across its domain and the distinction in sign determines direction of movement. The dimensions for β and $v(x, y)$ are L.

In order to visualise the total neighbour-dependent effect in $b(x)$ more easily, consider the example in Figure 3.1 where $\beta > 0$ (and $\hat{h}(x) = 1$). It shows the total effect of interactions $\sum_{j=1}^{10} v(x_j - x)$, from 10 neighbours located at x_j on a cell at x. To understand why we take a derivative of the interaction kernel $v(z)$ it helps to think of the total weighting function as a 'crowding surface' which a cell at x can use as a means of measuring the extent of crowding in its neighbourhood. In Figure 3.1, $-b(x)$ is the gradient of this 'surface' and cells are biased to move down the gradient in the direction of reduced crowding. Consider, for example, the arrangement of cells shown in Figure 3.1(a). Say the cell indicated by the arrow is about to undergo a movement event. At this location the gradient is positive so $b(x) < 0$ which corresponds to a bias for movement in the left direction, away from the crowded region on the individual's right. Thus, the sign of the gradient holds information about the direction in which crowded regions exist. In addition, steep gradients occur at locations on the edges of clusters while shallow or zero gradients occur either within clusters or in sparsely occupied regions. Therefore the magnitude of the gradient provides a measure for the degree of crowding in a location x. The bias $b(x)$ allows us to tap into the information held by the gradient of a 'crowding surface', for a particular arrangement of cells, and use it to determine the direction of movement for an individual at x.

Owing to our choice of $v(z)$, the effect of a neighbour located at y on the direction of movement for a cell at x is greater for small distances $|y - x|$, while for larger distances the effect is negligible. The strength of interaction is determined by the constant β. The variance σ_v^2 is a measure of spread for $v(z)$, affecting the spatial range of displacements over which a pair of cells interact. In Figure 3.1 we consider two different values of σ_v^2. When σ_v^2 is large, $v(z)$ has a wide spread that will influence outlying cells as shown in Figure 3.1(a). On the other hand for small σ_v^2, $v(z)$ is a narrow kernel and only neighbouring cells in close proximity to the individual will be affected by its presence (Fig. 3.1(b)).

As a means of relating the bias to an individual's direction of movement, we use $b(x)$ to determine the probability of moving right $p_r(b)$ for a cell at x. Its complement $(1 - p_r(b))$ determines the probability of moving left. For simplicity, we define $p_r(b)$ to be a logistic function

$$p_r(b) = \frac{1}{1 + \exp(-b)}, \tag{3.14}$$

so that for large $b(x) > 0$ a cell at x is strongly biased to move right, while for large $b(x) < 0$

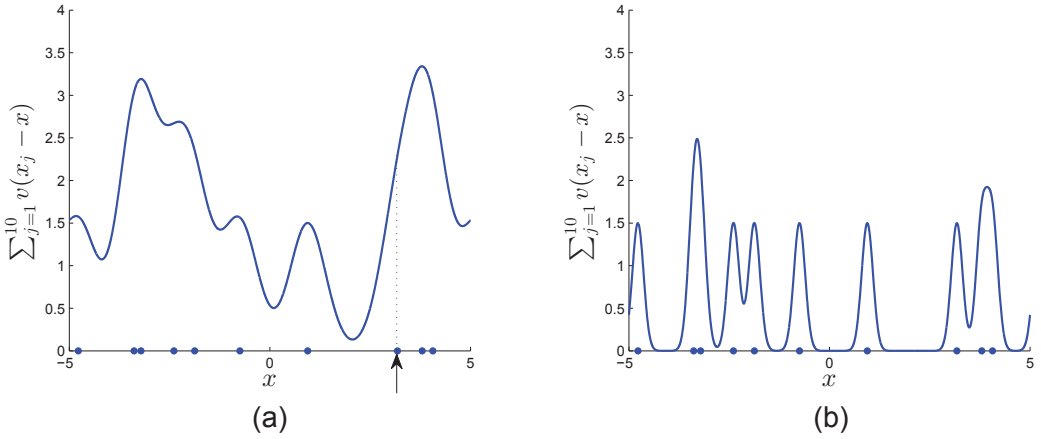


Figure 3.1: A function $\sum_{j=1}^{10} v(x_j - x)$ (blue line) for the total weighting on a cell at x, of interactions from 10 neighbours located at x_j for $j = 1, \ldots, 10$ (blue dots). The gradient of this function is $-b(x)$ and cells are biased to move down the gradient. The interaction strength is $\beta = 1.5$ and we compare σ_v^2 for different values (a) $\sigma_v^2 = 0.2$, and (b) $\sigma_v^2 = 0.02$. The arrow in (a) marks the location of an individual that is biased to move left.

the bias to move left is strong. When $b(x) = 0$ there is no bias from neighbours (i.e. the cell is either isolated or in the centre of a cluster).

Finally, we incorporate the directional bias into the movement PDF $\mu(x, y)$ to give a piecewise function

$$\mu(x,y) = \begin{cases} \lambda_\mu \exp\left(-\lambda_\mu |y-x|\right)\ p_r(b(x)) & \text{if } y - x > 0 \\ \lambda_\mu \exp\left(-\lambda_\mu |y-x|\right)\ (1 - p_r(b(x))) & \text{if } y - x < 0. \end{cases} \tag{3.15}$$

In a biological context $v(z)$ could be representative of, say, the extent to which an individual responds to a concentration of chemical signal secreted by a neighbouring cell. Then $b(x)$ would describe the total strength of a cell's response to signals from all neighbours and $p_r(b)$ the mechanism by which these interactions change the cell's direction of movement. The sign of β determines the nature of the directional bias. When $\beta > 0$, as shown in Figure 3.1, cells are biased to move away from close-lying neighbours. This type of behaviour facilitates movement of individuals out of crowded regions. For example some cell types release chemorepellents which have a repulsive effect on neighbouring cells [24]. Conversely, when $\beta < 0$ the directional bias will drive cells towards one another as may occur in the presence of a cell-secreted chemoattractant [25]. If we set $\beta = 0$ the resulting probability of moving right is $1/2$ and the direction of movement is unbiased. The dimensions of $\mu(x, y)$ are L^{-1} and as it is a PDF we have the constraint that $\int \mu(x, y) \mathrm{d}y = 1$.

With this description of the IBM in place, we are ready to proceed with a derivation of the population-level model. However, it will first be useful to discuss how a pair-correlation function is computed from a configuration of agents arising in an IBM realisation at a particular time.

3.3 Pair Correlation Function

The second spatial moment, the average density of pairs of cells, can be expressed as a pair-correlation function (PCF) $C(r)$, written in terms of a separation distance r [72]. We compute a PCF $C(r)$ from a particular configuration of agents in a domain of length L, in one spatial dimension, or of width L_x and height L_y in two spatial dimensions. First, a reference agent at x_i is chosen and the distance between x_i and a neighbour at x_j is calculated for $n-1$ neighbours. We measure across periodic boundaries such that in one-dimensional space the distance between a pair of agents displaced by $\xi = x_j - x_i$ is

$$r = \begin{cases} |\xi| & \text{if } |\xi| < \frac{L}{2} \\ L - |\xi| & \text{if } |\xi| > \frac{L}{2}. \end{cases} \tag{3.16}$$

A different reference agent is then chosen and the process repeated until each agent in the population has been selected as a reference once. Once all possible pair distances excluding self-pairs have been measured, $C(r)$ is constructed by counting the distances that fall into an interval $[r - \frac{\delta r}{2}, r + \frac{\delta r}{2}]$, i.e. binning distances using a bin width δr. Normalising by $n(n-1)(2\delta r)/L$ in one spatial dimension ensures that $C(r) = 1$ in the complete absence of spatial structure. In Chapter 6 we extend the IBM to two-dimensional space, for which the normalisation $n(n-1)(2\pi r \delta r)/(L_x L_y)$ is used.

The choice of δr is important because very small values can yield a PCF dominated by fluctuations while values that are too large result in an overly-smooth function which may mask spatial structure [74].

Chapter 4

Spatial Moments Model

The local interactions taking place between cells at the level of individuals give rise to larger scale effects at the population level. In the following sections we introduce a description of the first, second and third moments in terms of the probabilities of individuals being found in given regions. The definitions for the moments given here are equivalent to those given by Illian, et al [72]. We then use our IBM to derive a population-level model in terms of the dynamics of the first and second spatial moments. The following notation and method are consistent with the generalised derivation proposed by Plank and Law [61], however we have derived new terms to describe the effect of a neighbour-dependent directional bias.

4.1 Spatial Moments

The first, second and third spatial moments are the average densities of single cells, pairs and triplets, respectively. The concept is better explained by considering the geometry of three small regions δx, δy and δz centred on x, y and z, respectively. Each region has size h (length h in one dimension, area h in two dimensions, and volume h in three dimensions) and it is assumed that the probability of finding multiple cells within a single region is $O(h^2)$. These regions are depicted in Figure 4.1 in two-dimensional (Fig. 4.1(a)) and one-dimensional space (Fig. 4.1(b)). The probability of a cell being present in a small region δx at time t is the first spatial moment $Z_1(x,t)$. The probability of a cell being present in δx and a cell being present in δy at time t is the second spatial moment $Z_2(x,y,t)$. The third moment $Z_3(x,y,z,t)$ is the probability of cells being present in all three regions δx, δy and δz at time t. For a spatially homogeneous distribution of agents, the moments can be expressed in terms of the displacements between pairs of cells, as shown in Figure 4.1.

The spatial moments are functions of time as well as space but for ease of notation we will omit the argument t. The first spatial moment $Z_1(x)$ is expressed in terms of the probability

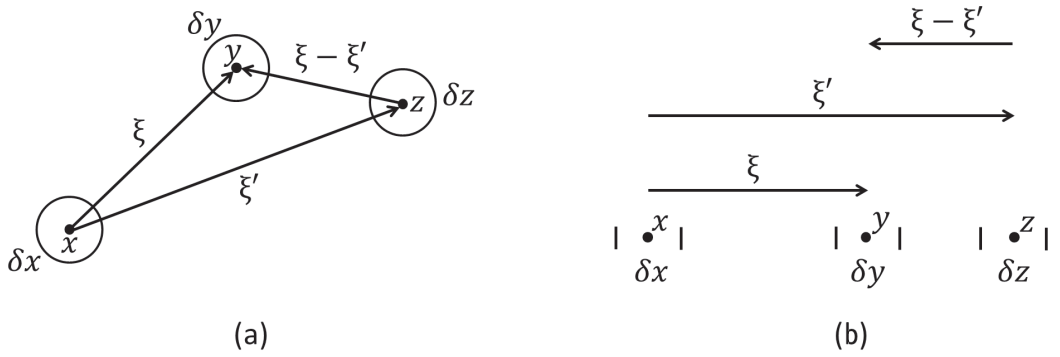


Figure 4.1: First, second and third spatial moments in (a) two-dimensional space, and (b) one-dimensional space. Small regions δx, δy and δz, of size h (length h in one dimension and area h in two dimensions), are centred on locations x, y and z, respectively. For a case where cells are distributed homogeneously throughout space, the second and third moment can be expressed in terms of the displacements $\xi = y - x$ and $\xi' = z - x$.

of a cell being found in a small region δx, centred on x and of size h, at time t as follows:

$$Z_1(x) = \lim_{h \to 0} \frac{P(I(x) = 1)}{h} \ . \tag{4.1}$$

$I(x)$ is an indicator variable such that $I(x) = 1$ if there is a cell in δx centred on x and $I(x) = 0$ if there is no cell in δx.

The second spatial moment $Z_2(x, y)$, the average density of cell pairs, involves the probability of cells being found in the small regions δx and δy as follows:

$$Z_2(x, y) = \lim_{h \to 0} \frac{P(I(x) = 1 \ \& \ I(y) = 1)}{h^2}. \tag{4.2}$$

For simplicity, we assume that δx and δy cannot overlap and so (4.2) excludes the case where $x = y$. A more rigorous definition which accounts for and removes the effect of such self-pairs (that would otherwise create a Dirac-delta peak in $Z_2(x, y)$ at $x = y$) is discussed by Plank and Law [61]. The second spatial moment has a two-fold symmetry such that $Z_2(x, y) = Z_2(y, x)$ [61].

The third spatial moment, the density of triplets in the small regions δx, δy and δz, is defined as

$$Z_3(x, y, z) = \lim_{h \to 0} \frac{P(I(x) = 1 \ \& \ I(y) = 1 \ \& \ I(z) = 1)}{h^3}, \tag{4.3}$$

excluding the cases where $x = y$, $x = z$ and $y = z$ as we assume that δx, δy and δz cannot overlap. Again, a more detailed description which allows for such non-distinct triplets is given

by Plank and Law [61]. The third moment has been shown to have a six-fold symmetry [79]. Similarly, the nth spatial moment Z_n is defined as the expected number of n-tuples of cells per unit (length)^{Dn}, for a D-dimensional space.

It is also useful to define some conditional probabilities. The probability of a cell being found in δy conditional on the presence of a cell in δx is $P(I(y) = 1|I(x) = 1)$. We use the fact that $P(A|B) = P(A\&B)/P(B)$ along with Eqs. (4.1), (4.2) and (4.3) to rewrite this conditional probability as follows:

$$P(I(y) = 1|I(x) = 1) = \frac{P(I(y) = 1 \ \& \ I(x) = 1)}{P(I(x) = 1)} = \frac{Z_2(x, y)h}{Z_1(x)} + O(h^2). \tag{4.4}$$

Similarly, we can write the probability of a cell being found in δz conditional on the presence of a cell in δx and a cell in δy as

$$\begin{aligned} P(I(z) = 1|I(x) = 1 \ \& \ I(y) = 1) &= \frac{P(I(x) = 1 \ \& \ I(y) = 1) \ \& \ I(z) = 1)}{P(I(x) = 1) \ \& \ I(y) = 1)} \\ &= \frac{Z_3(x, y, z)h}{Z_2(x, y)} + O(h^2). \end{aligned} \tag{4.5}$$

4.2 First Spatial Moment

The following derivation can be used to describe moment dynamics in a non-homogeneous space, where the first moment is dependent on x. We derive corresponding descriptions for movement rate ψ_i and PDF $\mu(x, y)$ in terms of spatial moments. In the IBM, the movement rate of individual i, given by (3.8), comprises an intrinsic component and a neighbour-dependent component which describes the contribution of neighbouring cells to individual i's motility. In the spatial moment dynamics this corresponds to an integration over y of the probability of a cell at y conditional on a cell being present at x. Using the conditional probability in Eq. (4.4), the expected movement rate (from hereafter simply referred to as movement rate) for a single cell at x is

$$M_1(x) = m + \int w(y - x)\frac{Z_2(x, y)}{Z_1(x)}\mathrm{d}y. \tag{4.6}$$

The maximum formula which ensured a non-negative movement rate in (3.8) is not incorporated in the spatial moment description because we only consider solutions in which negative expected movement rates do not arise.

In one spatial dimension, when a cell moves, it travels from an original location x to a destination y drawn from the PDF

$$\mu_1(x,y) = \begin{cases} \lambda_\mu \exp\left(-\lambda_\mu |y-x|\right)\ p_r(b_1(x)) & \text{if } y-x>0 \\ \lambda_\mu \exp\left(-\lambda_\mu |y-x|\right)\ (1-p_r(b_1(x))) & \text{if } y-x<0. \end{cases} \tag{4.7}$$

The neighbour-dependent bias term $b_1(x)$ sums a contribution $u(x,y)$ from all possible neighbours to the direction of movement of a cell at x. Therefore, we write $b_1(x)$ as an integration over y of the probability of a neighbour at y conditional on the presence of a cell at x, weighted by an interaction kernel $u(x,y)$:

$$b_1(x) = \int u(x,y)\frac{Z_2(x,y)}{Z_1(x)}\mathrm{d}y. \tag{4.8}$$

The movement rate $M_1(x)$, movement PDF $\mu_1(x,y)$ and neighbour-dependent bias $b_1(x)$ are functions of spatial moments and thus are also functions of time, however we have dropped the argument t from the notation for brevity. The same is true for the corresponding functions for the second spatial moment dynamics given below in (4.14)-(4.16). When solving for a spatially homogeneous case, $M_1(x)$ and $b_1(x)$ are independent of x and $\mu_1(x,y)$ can be expressed in terms of the displacement from x to y.

4.3 Dynamics of the First Spatial Moment

For the dynamics of the first spatial moment $Z_1(x)$ we consider the probability that a cell will be present in the small region δx centred on x at a time $t+\delta t$, where δt is a short period of time. For this situation to arise, a cell could have been present in δx at time t and remained there. Alternatively, a cell located elsewhere in the space could have arrived in δx via a movement event. Movement events occur over time as an inhomogeneous Poisson process and so the probability of more than one event occurring in δt is $O(\delta t^2)$. We can combine these possibilities into a single statement

$$\begin{aligned} P(\text{cell in } \delta x \text{ at } t+\delta t) \ &= P(\text{cell in } \delta x \text{ at } t)P(\text{cell remains in } \delta x \text{ in } [t, t+\delta t]) \\ &\quad + P(\text{cell absent in } \delta x \text{ at } t)P(\text{cell arrives in } \delta x \text{ in } [t, t+\delta t]). \end{aligned} \tag{4.9}$$

The probability that a cell remains in δx in $[t, t+\delta t]$ is

$$\begin{aligned} P(\text{cell remains in } \delta x \text{ in } [t,t+\delta t]) &= 1 - P(\text{cell moves out of } \delta x \text{ in } [t,t+\delta t]) \\ &= 1 - M_1(x)\delta t + O(\delta t^2). \end{aligned} \tag{4.10}$$

The probability that a cell arrives in δx in $[t, t+\delta t]$ can be written as a probability that a cell moved from u into δx, integrated over all possible starting locations u as follows:

$$P(\text{cell arrives in } \delta x \text{ in } [t,t+\delta t]) = h\delta t \int \mu_1(u,x) M_1(u) Z_1(u,t) \mathrm{d}u + O(\delta t^2), \tag{4.11}$$

where $M_1(u)Z_1(u,t)$ is the movement rate per unit area at location u. Now, using equations (4.1), (4.10) and (4.11), we can write (4.9) as

$$\begin{aligned} Z_1(x,t+\delta t)h \;=\; & Z_1(x,t)h\,[1 - M_1(x)\delta t] \\ & + [1 - Z_1(x,t)h]\; h\delta t \int \mu_1(u,x) M_1(u) Z_1(u,t) \mathrm{d}u \\ & + O(h^3) + O(\delta t^2). \end{aligned} \tag{4.12}$$

Making use of the Taylor expansion of $Z_1(x, t+\delta t)$ then taking the limit $h, \delta t \to 0$ leads us to a differential equation for the rate of change of the first spatial moment:

$$\frac{\mathrm{d}Z_1(x,t)}{\mathrm{d}t} \;=\; -M_1(x)Z_1(x,t) + \int \mu_1(u,x) M_1(u) Z_1(u,t) \mathrm{d}u. \tag{4.13}$$

This equation depends on the second spatial moment, incorporated in the movement rate term $M_1(x)$. The first term in (4.13) describes movement out of x while movement into x is accounted for in the second term as an integration over all possible starting locations u.

4.4 Second Spatial Moment

By considering only the first spatial moment, valuable information about small-scale spatial structure that may arise in the system is lost. Therefore, it is necessary to advance to the second spatial moment, the average density of pairs of cells.

For the second moment dynamics, we describe a movement rate function $M_2(x,y)$ for a cell at x conditional on the presence of a cell at y. Recall that the neighbour-dependent

component of movement rate for a single cell $M_1(x)$ was conditional on the presence of a second cell. Similarly, the neighbour-dependent component for $M_2(x, y)$ is conditional on the presence of a third cell at z and requires the third spatial moment. We use Eq. (4.5) to write $M_2(x, y)$ as follows:

$$M_2(x, y) = \ m + \int w(z - x)\frac{Z_3(x, y, z)}{Z_2(x, y)}\mathrm{d}z + w(y - x). \tag{4.14}$$

The third term here accounts for the direct effect of the cell at y on the cell at x. Because the regions δy and δz do not overlap, the third moment does not account for the case where $y = z$ and we add this interaction as a separate term.

In the dynamics of the second moment, a cell at x moves to a new location at y drawn from a PDF $\mu_2(x, y, z)$, where the third argument accounts for the fact that the cell at x is in a pair with a cell at z:

$$\mu_2(x, y, z) = \begin{cases} \lambda_\mu \exp\left(-\lambda_\mu |y - x|\right)\ p_r(b_2(x, z)) & \text{if } y - x > 0 \\ \lambda_\mu \exp\left(-\lambda_\mu |y - x|\right)\ (1 - p_r(b_2(x, z))) & \text{if } y - x < 0. \end{cases} \tag{4.15}$$

The neighbour-dependent bias term $b_2(x, y)$ represents the contribution of all possible neighbours to the direction of movement of the cell at x in a pair with a cell at y. It is an integration over z of the probability of a third neighbour at z conditional on the presence of a cell at x and a cell at y, weighted by the kernel $u(x, z)$. Thus, we have

$$b_2(x, y) = \int u(x, z)\frac{Z_3(x, y, z)}{Z_2(x, y)}\mathrm{d}z + u(x, y). \tag{4.16}$$

As in Eq. (4.14), the direct effect of a cell at y on a cell at x must be added as a separate term because the third moment does not account for the degenerate case where $y = z$. In a spatially homogeneous setting, $M_2(x, y)$, $b_2(x, y)$ and $\mu_2(x, y, z)$ can be expressed in terms of displacements rather than physical locations, as illustrated in Figure 4.1 and described in Section 4.6.

4.5 Dynamics of the Second Spatial Moment

We now derive an equation for the rate of change of the second spatial moment. Consider the probability of finding a cell present in the region δx and a cell in the region δy at a time

$t + \delta t$. The following four cases could have led to this outcome:

i Regions δx and δy each contained a cell at time t and both cells remained there in $[t, t+\delta t]$.

ii There was a cell present in δy but not in δx at time t. In $[t, t+\delta t]$, the cell in δy remained there and a cell arrived in δx.

iii There was a cell present in δx but not in δy at time t. In $[t, t+\delta t]$, the cell in δx remained there and a cell arrived in δy.

iv Neither region δx or δy contained a cell at time t. In $[t, t+\delta t]$, a cell arrived in δx and a cell arrived in δy.

We can summarise these scenarios as follows:

$$\begin{aligned} P\left(\begin{array}{c}\text{cell in } \delta x \text{ \& cell}\\ \text{in } \delta y \text{ at } t+\delta t\end{array}\right) &= P\left(\begin{array}{c}\text{cell in } \delta x \text{ \&}\\ \text{cell in } \delta y \text{ at } t\end{array}\right) P\left(\begin{array}{c}\text{cells remain}\\ \text{in } \delta x \text{ and } \delta y\end{array}\right)\\ &\quad + P\left(\begin{array}{c}\text{cell in } \delta y \text{ but}\\ \text{not in } \delta x \text{ at } t\end{array}\right) P\left(\begin{array}{c}\text{cell remains in } \delta y\\ \text{\& cell arrives in } \delta x\end{array}\right)\\ &\quad + P\left(\begin{array}{c}\text{cell in } \delta x \text{ but}\\ \text{not in } \delta y \text{ at } t\end{array}\right) P\left(\begin{array}{c}\text{cell remains in } \delta x\\ \text{\& cell arrives } \delta y\end{array}\right)\\ &\quad + P\left(\begin{array}{c}\text{cell absent at}\\ \delta x \text{ and } \delta y \text{ at } t\end{array}\right) P\left(\begin{array}{c}\text{cell arrives in } \delta x\\ \text{\& cell arrives in } \delta y\end{array}\right). \end{aligned} \tag{4.17}$$

The probability of cells being present in both δx and δy can be written in terms of $Z_2(x, y)$ from equation (4.2):

$$P(\text{cell in } \delta x \text{ \& cell in } \delta y \text{ at } t) = Z_2(x, y, t)h^2 + O(h^3). \tag{4.18}$$

Using (4.1) and (4.2), the probability of a cell being present in δy and absent from δx is

$$P(\text{cell in } \delta y \text{ but not in } \delta x \text{ at } t) = Z_1(y, t)h - Z_2(x, y, t)h^2 + O(h^3). \tag{4.19}$$

The probability of both cells remaining in δx and δy in $[t, t+\delta t]$ is written in terms of (4.14):

$$P\left(\begin{array}{c}\text{both cells remain}\\ \text{in } \delta x \text{ and } \delta y\end{array}\right) = 1 - P\left(\begin{array}{c}\text{cell moves out of}\\ \delta x \text{ in } [t, t+\delta t]\end{array}\right) - P\left(\begin{array}{c}\text{cell moves out of}\\ \delta y \text{ in } [t, t+\delta t]\end{array}\right) + O(\delta t^2)$$
$$= 1 - (M_2(x, y) + M_2(y, x))\delta t + O(\delta t^2). \tag{4.20}$$

This is comparable to Eq. (4.10) for the first moment dynamics.

The probability of a cell remaining in δy and a cell arriving in δx via a movement event in $[t, t+\delta t]$ is equivalent to the conditional probability that a cell arrives in δx given that there is a cell in δy. As in (4.11), we integrate over all possible starting locations u for the cell moving into δx. However, the probability of a cell being located at u is conditional on the presence of a cell at y. Therefore we have

$$P\left(\begin{array}{c}\text{cell remains in } \delta y\\ \text{\& cell arrives in } \delta x\end{array}\right) = h\delta t \int \mu_2(u, x, y) M_2(u, y) \frac{Z_2(u, y, t)}{Z_1(y, t)} \mathrm{d}u + O(\delta t^2). \tag{4.21}$$

Finally, the probability of a cell arriving in δx and a cell arriving in δy is $O(\delta t^2)$ because this would involve two Poisson events occurring during $[t, t+\delta t]$. Similarly, the higher order terms in Eqs. (4.20) and (4.21) arise from probabilities involving more than one cell undergoing a movement event during a time δt.

We substitute equations (4.18)-(4.21) into (4.17) and make use of the 2-fold symmetry of $Z_2(x, y, t)$ to give

$$\begin{aligned} Z_2(x, y, t+\delta t)h^2 \;=\; & Z_2(x, y, t)h^2[1 - (M_2(x, y) + M_2(y, x))\delta t] \\ & + [Z_1(y, t)h - Z_2(x, y, t)h^2]h\delta t \int \mu_2(u, x, y) M_2(u, y) \frac{Z_2(u, y, t)}{Z_1(y, t)} \mathrm{d}u \\ & + [Z_1(x, t)h - Z_2(x, y, t)h^2]h\delta t \int \mu_2(u, y, x) M_2(u, x) \frac{Z_2(u, x, t)}{Z_1(x, t)} \mathrm{d}u \\ & + O(h^4) + O(\delta t^2). \end{aligned} \tag{4.22}$$

Using a Taylor expansion of $Z_2(x, y, t+\delta t)$, expanding terms, then letting $h, \delta t \to 0$ which

removes the higher order terms, leads us to

$$\begin{aligned}\frac{\mathrm{d}Z_2(x,y,t)}{\mathrm{d}t} &= -(M_2(x,y)+M_2(y,x))Z_2(x,y,t)\\ &\quad+\int \mu_2(u,x,y)M_2(u,y)Z_2(u,y,t)\mathrm{d}u\\ &\quad+\int \mu_2(u,y,x)M_2(u,x)Z_2(u,x,t)\mathrm{d}u. \end{aligned} \tag{4.23}$$

Here, the first negative term describes movement out of x, conditional on the presence of a cell at y. The first integral term represents movement into x from a starting location u, conditional on the presence of a cell at y. The remainder are symmetric terms for movement out of and into y. For notational simplicity, from here on we will drop the t from the spatial moment notation.

4.5.1 Moment Closure for the Third Spatial Moment

Equation (4.23) depends on the third spatial moment. Similarly the third moment dynamics will depend on the fourth moment and so on. Therefore, we need to close the system before solving. To achieve this we use the power-3 closure (also known as Kirkwood superposition approximation) given by

$$\tilde{Z}_3(x,y,z) = \frac{Z_2(x,y)Z_2(x,z)Z_2(y,z)}{Z_1(x)Z_1(y)Z_1(z)}, \tag{4.24}$$

This is only an approximation to the third moment and other choices of closure may perform better. The power-2 closure for example has been shown to perform well, however this closure contains a term that may cause the third moment to become negative [79]. For a Poisson spatial pattern the third moment is $Z_3(x,y,z) = Z_1^3$ and the approximation in (4.24) has perfect accuracy.

4.6 Spatially Homogeneous Moment Dynamics

To solve Eq. (4.23) for the dynamics of the second moment numerically it is beneficial to reduce the number of variables. For a spatially homogeneous distribution of cells, the first spatial moment is independent of x. The second moment $Z_2(x,y)$ depends only on the displacement $y-x$ which can now be treated as a single variable. As shown in Figure 4.1, the displacement from x to y is denoted ξ and the displacement from x to z is denoted ξ'.

For the movement PDF $\mu_2(u,x,y)$, we denote the displacement from u to x as ξ''. The spatially-dependent weight $\hat{h}(x)$ in (3.12) and (6.3) is set to $\hat{h}(x)=1$.

Equation (4.23) is rewritten in terms of the displacements between pairs as follows:

$$\begin{aligned}\frac{\mathrm{d}Z_2(\xi)}{\mathrm{d}t} &= -(M_2(\xi)+M_2(-\xi))Z_2(\xi)\\ &\quad+\int \mu_2(\xi'',\xi''+\xi)M_2(\xi''+\xi)Z_2(\xi''+\xi)\mathrm{d}\xi''\\ &\quad+\int \mu_2(\xi'',\xi''-\xi)M_2(\xi''-\xi)Z_2(\xi''-\xi)\mathrm{d}\xi''.\end{aligned} \tag{4.25}$$

The movement rate $M_2(x,y)$ of a cell at x in a pair with a cell at y given in (4.14) is now expressed in terms of the displacement ξ between x and y:

$$M_2(\xi) = m+\int w(\xi')\frac{Z_3(\xi,\xi')}{Z_2(\xi)}\mathrm{d}\xi'+w(\xi). \tag{4.26}$$

The movement PDF given in (4.15) becomes

$$\mu_2(\xi,\xi')=\begin{cases}\lambda_\mu \exp\left(-\lambda_\mu|\xi|\right)\ p_r(b_2(\xi')) & \text{if } \xi>0\\ \lambda_\mu \exp\left(-\lambda_\mu|\xi|\right)\ (1-p_r(b_2(\xi'))) & \text{if } \xi<0,\end{cases} \tag{4.27}$$

with neighbour-dependent bias

$$b_2(\xi)=\int v'(\xi')\frac{Z_3(\xi,\xi')}{Z_2(\xi)}\mathrm{d}\xi'+v'(\xi). \tag{4.28}$$

The interaction kernels were previously expressed in terms of a single variable in (3.9) and (3.13) and these definitions still hold here. The closure for the third moment is

$$\tilde{Z}_3(\xi,\xi')=\frac{Z_2(\xi)Z_2(\xi')Z_2(\xi'-\xi)}{Z_1^3}. \tag{4.29}$$

The boundary condition is as follows:

$$Z_2(\xi)\to Z_1^2 \text{ as } \xi\to\infty. \tag{4.30}$$

Chapter 5

Numerical Results for One-dimensional Model

We now compare some numerical results to measure how effectively our spatial moment model approximates the behaviour predicted by the IBM in one spatial dimension. We consider a simple case in which the distribution of cells is homogeneous throughout space. Therefore, the spatially-dependent weight $\hat{h}(x)$ in (3.12) is $\hat{h}(x) = 1$. The first spatial moment is independent of x and the second and third moments can be expressed in terms of pair displacements as described in Section 4.6. Because our model does not incorporate cell proliferation or death the first moment is also stationary in time, i.e. its rate of change given in (4.13) is zero.

To obtain spatial information from the IBM we calculate a PCF $C_{\mathrm{IBM}}(\xi)$ by averaging the results of repeated realisations. The PCF predicted by the spatial moment model is given by $C_{\mathrm{SM}}(\xi) = Z_2(\xi)/Z_1^2$ such that $C_{\mathrm{SM}}(\xi) = 1$ in the absence of spatial structure. The second moment is isotropic (i.e. it has symmetry about the origin) and therefore we only show $C_{\mathrm{SM}}(\xi)$ for $\xi \geq 0$.

In each realisation of the IBM we distribute the cells at $t = 0$ according to a spatial Poisson process on $[x_l, x_r]$ with intensity n/L. Therefore, initially there is no spatial structure present. The corresponding initial condition for the spatial moment model is to set $Z_2(\xi) = Z_1^2$ at $t = 0$. Results from both models are compared at time $t = 25$, by which point the system has converged to steady state in the majority of cases.

5.1 Non-Interacting Model

We begin by considering a simple case of cell migration in the absence of any neighbour-dependent interactions. In the complete absence of interactions, movement rate is determined by the intrinsic component alone and direction of movement is unbiased (i.e. $\mu(x, y)$ is a Laplace distribution). Under these conditions, Equation (4.25) reduces to a simpler form and it is straightforward to show analytically that the steady state solution for $Z_2(\xi)$ is a constant.

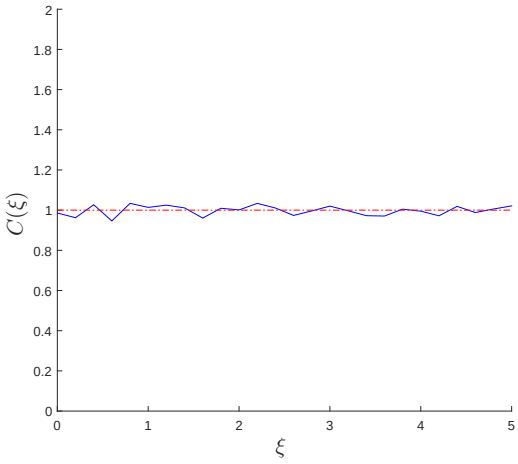


Figure 5.1: Migration in the absence of neighbour-dependent motility and neighbour-dependent directional bias ($\alpha = 0$, $\beta = 0$). Collective movement of 200 cells in a domain of length $L = 500$ therefore $Z_1 = 0.4$. The pair correlation function for 100 averaged realisations of the IBM (blue solid) is plotted against the second moment (red broken) at time $t = 25$. $\sigma_w^2 = \sigma_v^2 = 0.08, \lambda_\mu = 5, m = 10$.

Numerical solutions and IBM simulations confirm this, as shown in Figure 5.1. As we would expect, in the absence of interactions cells merely undergo diffusion-like movement and no spatial structure develops over time. For this simple case the spatial moment model provides a good approximation to the underlying IBM.

5.2 Interacting Model

5.2.1 Neighbour-Dependent Motility

We first consider a case with neighbour-dependent motility but in the absence of neighbour-dependent directional bias. Figure 5.2 shows results for different values of α where interaction strength increases from left to right. We choose $\alpha < 0$ to be sufficiently small such that the sum of the motility rate's intrinsic and neighbour-dependent components will give rise to $\psi_i > 0$ with high probability. Owing to the stochastic nature of the IBM it is possible that negative motility rates may occur by chance; the definition of ψ_i given in Eq. (3.8) ensures that $\psi_i = 0$ for such rare chance events. This does mean, however, that for large magnitudes of $\alpha < 0$ there is an equilibrium where all individuals have motility $\psi_i = 0$ and this is reached with probability 1 if the IBM is simulated for a sufficiently long time.

In Figures 5.2(a)-(c) for $\alpha > 0$, $C(\xi) < 1$ at short displacements which corresponds to a regular spatial pattern. When $\alpha < 0$ (Figs. 5.2(d)-(f)), $C(\xi) > 1$ at short displacements

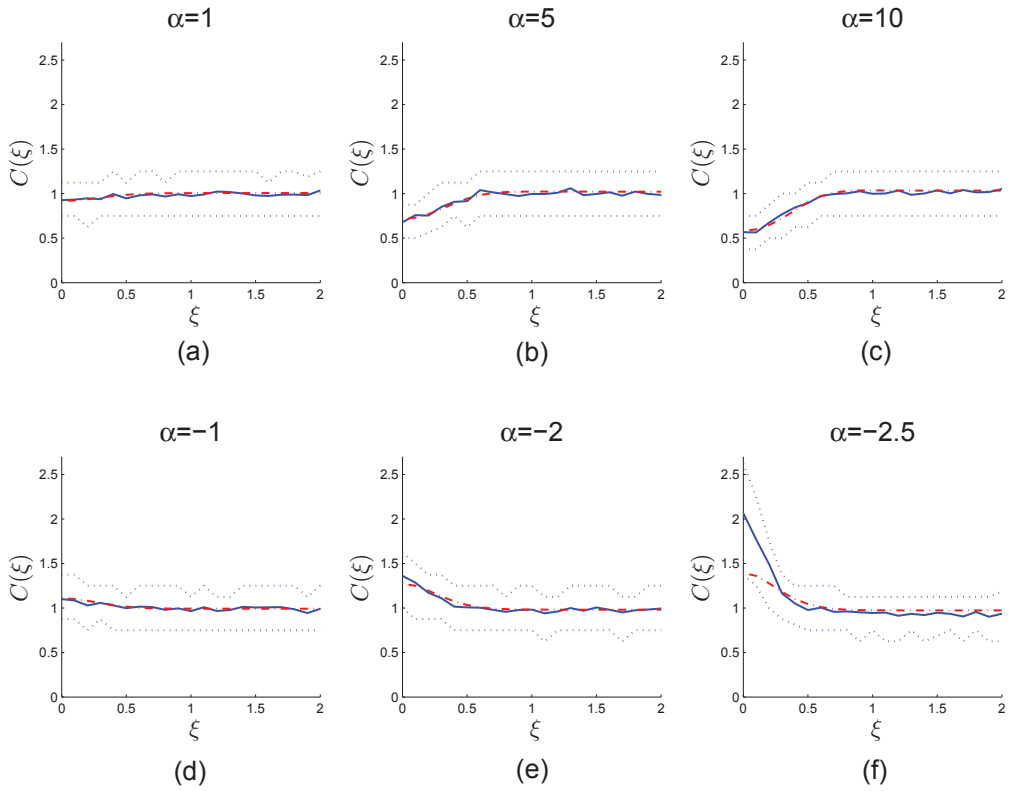


Figure 5.2: Migration with neighbour-dependent motility but in the absence of neighbour-dependent directional bias ($\beta = 0$). Collective movement of 200 cells in a domain of length $L = 500$ therefore $Z_1 = 0.4$. The PCF $C_{\mathrm{IBM}}(\xi)$ (blue solid) for 300 averaged realisations of the IBM is plotted against the PCF $C_{\mathrm{SM}}(\xi)$ (red broken) predicted by the spatial moment model at time $t = 25$. Blue dotted lines indicate the interquartile range for the PCF across the realisations performed, i.e. 50 % of realisations yield a PCF in the region between the blue dotted lines. Parameters are $\sigma_w^2 = \sigma_v^2 = 0.08, \lambda_\mu = 5, m = 10$. (a) $\alpha = 1$, (b) $\alpha = 5$, (c) $\alpha = 10$, (d) $\alpha = -1$, (e) $\alpha = -2$, and (f) $\alpha = -2.5$.

indicating a cluster spatial pattern. Increasing the magnitude of α (i.e. the strength of interaction) increases the extent of spatial structure.

$C_{\mathrm{SM}}(\xi)$ provides a good approximation to $C_{\mathrm{IBM}}(\xi)$ except for $\alpha < 0$ when $|\alpha|$ is large. For $\alpha = -2$ and $\alpha = -2.5$, $C_{\mathrm{SM}}(\xi)$ has converged to a steady state by $t = 25$ but $C_{\mathrm{IBM}}(\xi)$ continues to increase over time at short displacements. The discrepancy between $C_{\mathrm{SM}}(\xi)$ and $C_{\mathrm{IBM}}(\xi)$ can likely be attributed to the increased occurrence of negative motility rates (set to $\psi_i = 0$ as previously discussed), which can accumulate during simulation of the IBM when the magnitude of $\alpha < 0$ is sufficiently large. The chance occurrence of many pairs being found at short displacements, while reasonably rare for the chosen values of α, may cause a positive feedback reaction whereby the motility rate is reduced for these pairs to an extent where they are very unlikely to undergo further movements. Any cells that move into the

resulting cluster will also have their motility rates drastically reduced causing the effect to propagate. The spatial moment model does not account for these rare events as it deals only with average behaviour.

For instance, in the stochastic simulations for $\alpha = -1$ none of the motility rates that arose were negative and $C_{\text{SM}}(\xi)$ matched $C_{\text{IBM}}(\xi)$ well. However for $\alpha = -2$ and $\alpha = -2.5$ at $t = 25$ negative motility rates represented 0.1% and 1% of all motility rates respectively. Increasing t beyond this time caused the incidences to further increase. In contrast, the average motility rate $M_2(\xi)$ predicted by the spatial moment model remained positive for all time. While the 0.1% incidence when $\alpha = -2$ was sufficiently low as to be of little or no consequence for $C_{\text{IBM}}(\xi)$, Figure 5.2(f) shows that even a relatively low incidence of 1% can lead to a mis-match between $C_{\text{IBM}}(\xi)$ and $C_{\text{SM}}(\xi)$. Further increasing the magnitude of $\alpha < 0$ causes a significant increase in the incidences of $\psi_i < 0$ and the fit between $C_{\text{SM}}(\xi)$ and $C_{\text{IBM}}(\xi)$ deteriorates to an even greater extent.

5.2.2 Neighbour-Dependent Directional Bias

We now consider a case of migration in the absence of neighbour-dependent motility but in the presence of neighbour-dependent directional bias. Figures 5.3(a)-(c) show that setting $\beta > 0$, such that cells are biased to move away from crowded regions, generates a regular spatial pattern. In Figure 5.3(c) for $\beta = 10$ there is a peak in both the $C_{\text{IBM}}(\xi)$ and $C_{\text{SM}}(\xi)$ around $\xi = 1$. This peak arises because the strong directional bias is forcing cells to be displaced as far as possible from their nearest neighbours. This leads to an extreme case of regular spatial pattern in which nearly all cells are separated by approximately the same displacement; the peak in $C(\xi)$ corresponds to this common displacement. Setting $\beta < 0$, such that cells are biased to move towards their neighbours, gives rise to clustering as shown in Figures 5.3(d)-(f). Increasing the magnitude of interaction strength β increases the extent of spatial structure.

In Figure 5.3, $C_{\text{SM}}(\xi)$ provides a good approximation to $C_{\text{IBM}}(\xi)$. However, greater magnitudes of $\beta < 0$ lead to disparities between $C_{\text{SM}}(\xi)$ and $C_{\text{IBM}}(\xi)$. For example when $\beta = -0.5$, $C_{\text{IBM}}(0) \approx 3.6$ at $t = 25$ while $C_{\text{SM}}(0) \approx 11.5$ and neither PCF has reached steady state. Over time, the cluster pattern becomes stronger and the disparity between the two approximations deteriorates because $C_{\text{SM}}(0)$ is increasing at a faster rate than $C_{\text{IBM}}(0)$.

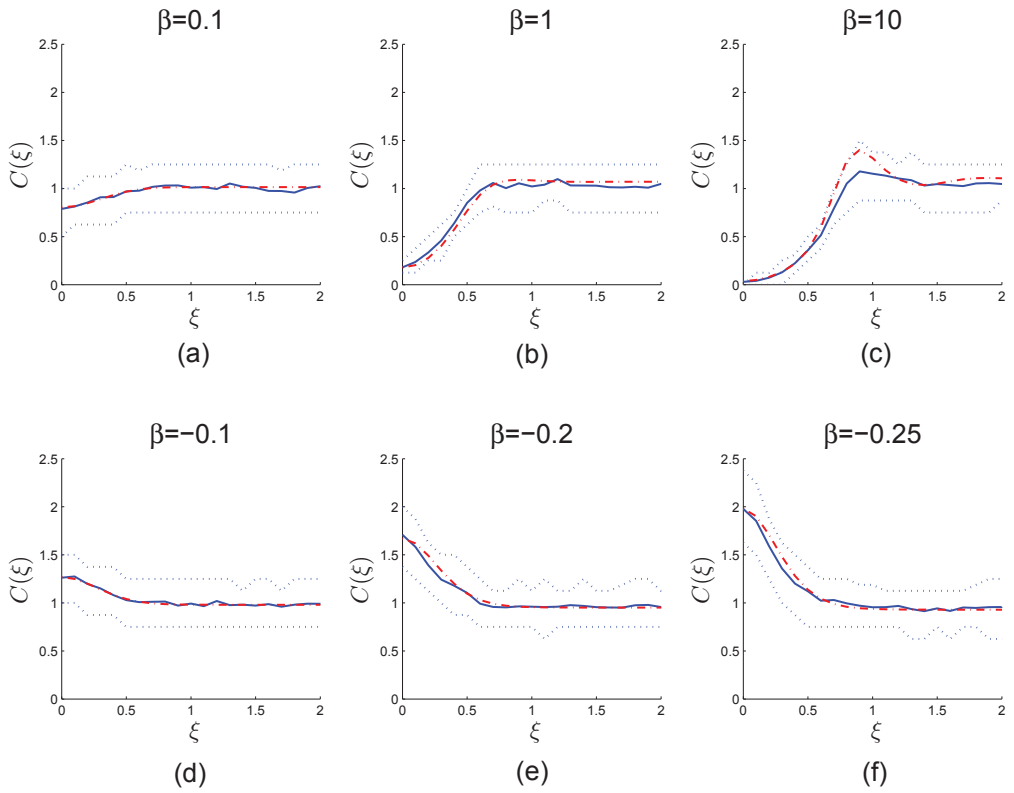


Figure 5.3: Migration with neighbour-dependent directional bias but in the absence of neighbour-dependent motility ($\alpha = 0$). Collective movement of 200 cells in a domain of length $L = 500$ therefore $Z_1 = 0.4$. The PCF $C_{\mathrm{IBM}}(\xi)$ (blue solid) for 300 averaged realisations of the IBM is plotted against the PCF $C_{\mathrm{SM}}(\xi)$ (red broken) predicted by the spatial moment model at time $t = 25$. Blue dotted lines indicate the interquartile range for the PCF across the realisations performed, i.e. 50 % of realisations yield a PCF in the region between the blue dotted lines. Parameters are $\sigma_w^2 = \sigma_v^2 = 0.08, \lambda_\mu = 5, m = 10$. (a) $\beta = 0.1$, (b) $\beta = 1$, (c) $\beta = 10$, (d) $\beta = -0.1$, (e) $\beta = -0.2$, (f) $\beta = -0.25$.

5.2.3 Neighbour-Dependent Motility and Directional Bias

Now that we have a better understanding of the independent effects of neighbour-dependent motility and directional bias we will consider the case where both are incorporated together. Figure 5.4 shows results for four different combinations of α and β. Figures 5.4(a) and 5.4(d) show that when the neighbour-dependent motility and directional bias are working cooperatively to promote spatial structure this results in a greater magnitude of departure from a Poisson spatial pattern than would occur when considering either interaction in isolation. However, when the neighbour-dependent interactions are working in opposition (Figs. 5.4(b)-(c)), they counteract one another and very little spatial structure develops over time as indicated by $C_{\mathrm{SM}}(\xi) \approx 1$. $C_{\mathrm{SM}}(\xi)$ is a good approximation to $C_{\mathrm{IBM}}(\xi)$ except in Figure 5.4(d). In this case, the slight mis-match near $\xi = 0$ is likely due to the fact that the two

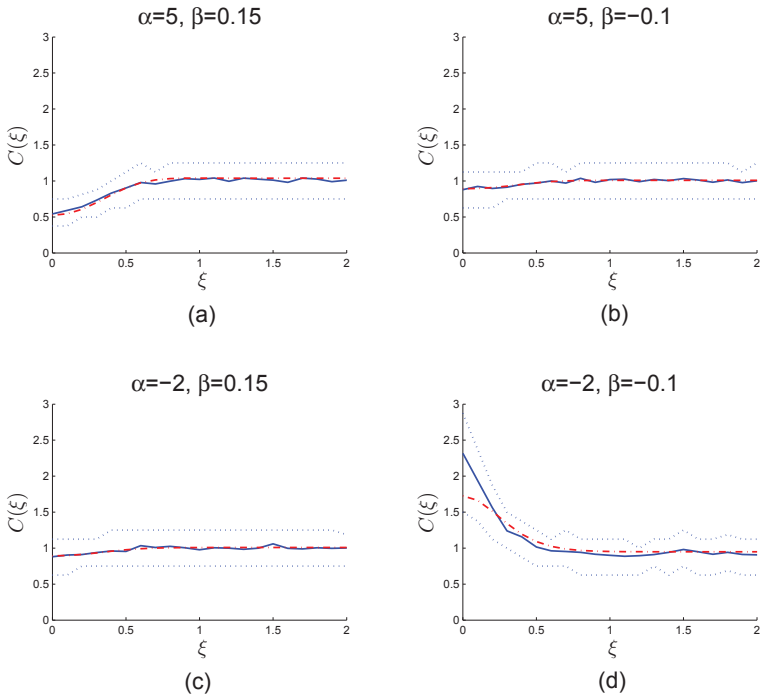


Figure 5.4: Migration with both neighbour-dependent motility and a directional bias. Collective movement of 200 cells in a domain of length $L = 500$ therefore $Z_1 = 0.4$. The PCF $C_{\mathrm{IBM}}(\xi)$ (blue solid) for 300 averaged realisations of the IBM is plotted against the PCF $C_{\mathrm{SM}}(\xi)$ (red broken) predicted by the spatial moment model at time $t = 25$. Blue dotted lines indicate the interquartile range for the PCF across the realisations performed, i.e. 50 % of realisations yield a PCF in the region between the blue dotted lines. Parameters are $\sigma_w^2 = \sigma_v^2 = 0.08, \lambda_\mu = 5, m = 10$. (a) $\alpha = 5$, $\beta = 0.15$, (b) $\alpha = 5$, $\beta = -0.1$, (c) $\alpha = -2$, $\beta = 0.15$, (d) $\alpha = -2$, $\beta = -0.1$.

forms of interaction working together promote clustering to such a degree that incidences of negative motility rate (approximately 0.8%) in the IBM become important.

5.2.4 Moment Closure

The numerical results above employ a power-3 closure to truncate the system of spatial moment dynamics at second order. This closure provides a reasonable approximation to the third spatial moment, except for a strongly clustered spatial pattern. Another closure that is discussed in the literature [79, 96] is the power-2 closure, given by

$$\tilde{Z}_3(\xi,\xi') = \frac{1}{\alpha+\gamma}\left(\alpha\frac{Z_2(\xi)Z_2(\xi')}{Z_1} + \beta\frac{Z_2(\xi)Z_2(\xi'-\xi)}{Z_1} + \gamma\frac{Z_2(\xi')Z_2(\xi'-\xi)}{Z_1} - \beta Z_1^3\right). \tag{5.1}$$

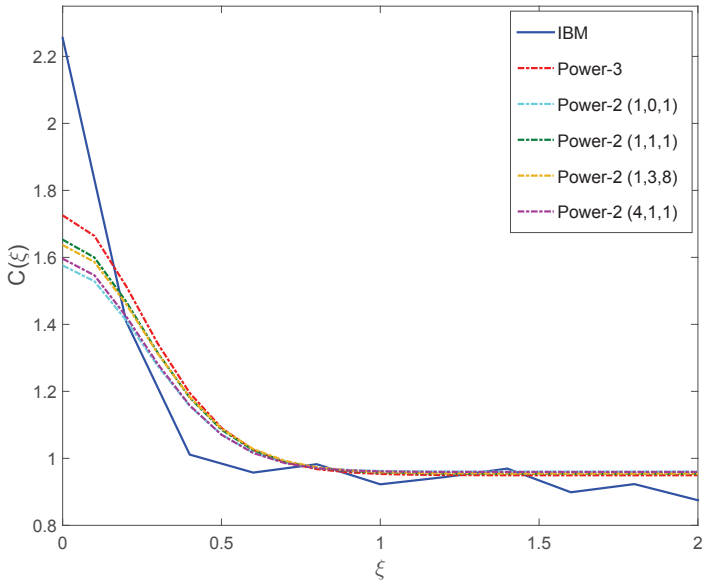


Figure 5.5: Comparison of closures for the third spatial moment. Collective movement of 200 cells in a domain of length $L = 500$ therefore $Z_1 = 0.4$. The PCF $C_{\text{IBM}}(\xi)$ (blue solid) for 300 averaged realisations of the IBM is plotted against the PCF $C_{\text{SM}}(\xi)$ predicted by the spatial moment model at time $t = 25$, using a power-3 closure (red dashed line), power-2 closure with weighting {1,0,1} (cyan dashed line), power-2 closure with weighting {1,1,1} (green dashed line), power-2 closure with weighting {1,3,8} (yellow dashed line) and power-2 closure with weighting {4,1,1} (purple dashed line). Parameters are $\alpha = -2$, $\beta = -0.1$, $\sigma_w^2 = \sigma_v^2 = 0.08$, $\lambda_\mu = 5$, $m = 10$.

Here, α, β and γ are weights that can be applied to each corner of the triplet. Setting $\alpha = \beta = \gamma$ gives the symmetric form of the closure, while setting different weights creates asymmetric versions. Three asymmetric power-2 closures that have proven particularly successful in other models employ the weightings $\{\alpha, \beta, \gamma\} = \{4, 1, 1\}$ [65], $\{\alpha, \beta, \gamma\} = \{1, 3, 8\}$ [79] and $\{\alpha, \beta, \gamma\} = \{1, 0, 1\}$ [70].

Figure 5.5 compares the predictions of the power-3 closure with these four power-2 closures for a cluster spatial pattern, where the performance of the approximations tends to be most limited. The closure with the best performance in this case is the power-3 closure and is therefore an appropriate choice for our model. The asymmetric power-2 closure with weighting {1,0,1} provides the poorest approximation overall.

5.2.5 Average Cell Density

The results discussed so far have explored collective movement for $Z_1 = 0.4$. However, the spatial structure that arises due to short-range interactions will depend largely on this first spatial moment. If we interpret 2σ (two standard deviations) as the approximate range over

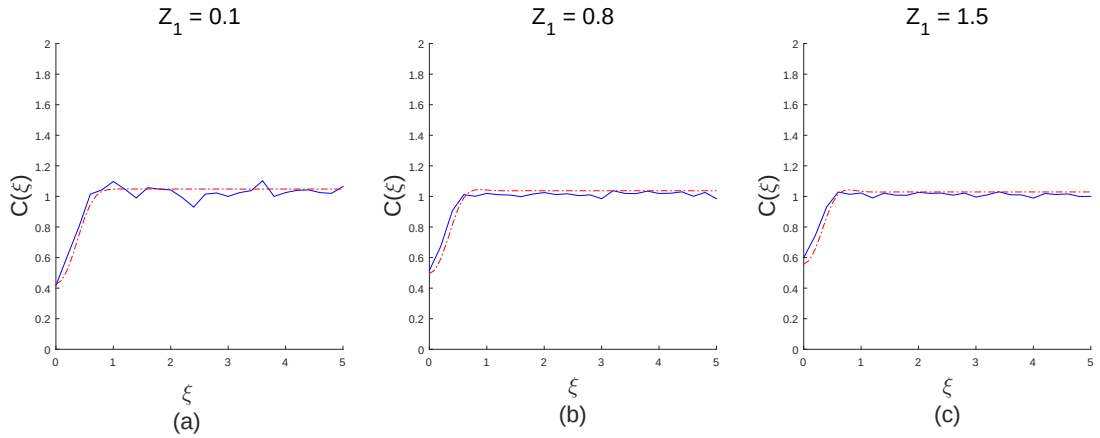


Figure 5.6: Spatial structure at three different average cell densities for 200 cells undergoing collective movement. The PCF $C_{\mathrm{IBM}}(\xi)$ (blue solid) for 100 averaged realisations of the IBM is plotted against the PCF $C_{\mathrm{SM}}(\xi)$ (red broken) predicted by the spatial moment model at time $t = 25$. Parameters are $\alpha = -2$, $\beta = 0.5$, $\sigma_w^2 = \sigma_v^2 = 0.08$, $\lambda_\mu = 5$, $m = 10$. (a) Low average density, $Z_1 = 0.1$, (b) moderate average density, $Z_1 = 0.8$, and (c) high average density, $Z_1 = 1.5$.

which a cell interacts, this can be used as a proxy for the space occupied by a cell giving a sense of scale to the spatial domain. We assume that the theoretical maximum cell density corresponds to the case where individuals are regularly spaced along the domain, each cell separated from its nearest neighbours by a displacement corresponding to the average cell diameter. Under this assumption, the theoretical maximum cell density for n cells is $n/2\sigma n$.

Figure 5.6 shows the PCF for three different values of Z_1. In this case, $\sigma_w^2 = \sigma_v^2 = 0.08$ giving an average cell diameter 0.57 and theoretical maximum density 1.75. Increasing Z_1 (by holding n fixed and decreasing L) leads to a decrease in the magnitude of local spatial structure. For very high densities (above the theoretical maximum density), $C(\xi) \approx 1$ for all values of ξ indicating an absence of spatial structure.

5.2.6 Width of Interaction Kernels

We now consider what effect changing the width of the interaction kernels has on the second spatial moment. We first use the moment equations to give an analytical description of the effect and then show some corroborative numerical results.

First, consider an interaction kernel $w(c\xi)$ which is a horizontal stretching of $w(\xi)$ by a

factor c when $0 < c < 1$ and a horizontal compression for $c > 1$. We now define the following:

$$\begin{aligned} \hat{w}(\xi) &= w(c\xi), & \hat{Z}_1 &= cZ_1, \\ \frac{\mathrm{d}}{\mathrm{d}\xi}\hat{v}(\xi) &= \frac{\mathrm{d}}{\mathrm{d}\xi}\left(\frac{1}{c}v(c\xi)\right), & \hat{Z}_2(\xi) &= c^2 Z_2(c\xi), \\ \hat{\mu}(\xi'', \xi'' + \xi) &= c\mu(c\xi'', c\xi'' + c\xi), & \hat{Z}_3(\xi, \xi') &= c^3 Z_3(c\xi, c\xi'), \end{aligned} \tag{5.2}$$

At steady state, the rate of change equation for the second moment in (4.25) gives us

$$0 = -M_2(\xi)Z_2(\xi) + \int \mu(\xi'', \xi'' + \xi)M_2(\xi'' + \xi)Z_2(\xi'' + \xi)\mathrm{d}\xi'' + \langle \xi \to -\xi \rangle, \tag{5.3}$$

where for sake of conciseness we have only shown terms for movement involving a pair displaced by ξ. The additional terms for movement involving a pair displaced by $-\xi$ are found by making a substitution ξ for $-\xi$ and we have denoted this $\langle \xi \to -\xi \rangle$. We can rewrite (5.3) in terms of the newly-defined functions given in (5.2) as follows:

$$0 = -\hat{M}_2\left(\frac{\xi}{c}\right)\frac{1}{c^2}\hat{Z}_2\left(\frac{\xi}{c}\right) + \int \frac{1}{c}\hat{\mu}\left(\frac{\xi''}{c}, \frac{\xi'' + \xi}{c}\right)\hat{M}_2\left(\frac{\xi'' + \xi}{c}\right)\frac{1}{c^2}\hat{Z}_2\left(\frac{\xi'' + \xi}{c}\right)\mathrm{d}\xi'' + \langle \xi \to -\xi \rangle. \tag{5.4}$$

Making a change of variables from ξ'' to $c\xi''$ in the integral then simplifying terms gives the steady state equation for $\hat{Z}_2\left(\frac{\xi}{c}\right)$:

$$0 = -\hat{M}_2\left(\frac{\xi}{c}\right)\hat{Z}_2\left(\frac{\xi}{c}\right) + \int \hat{\mu}\left(\xi'', \xi'' + \frac{\xi}{c}\right)\hat{M}_2\left(\xi'' + \frac{\xi}{c}\right)\hat{Z}_2\left(\xi'' + \frac{\xi}{c}\right)\mathrm{d}\xi'' + \langle \xi \to -\xi \rangle, \tag{5.5}$$

where

$$\begin{aligned} \hat{M}_2\left(\frac{\xi}{c}\right) &= m + \int \hat{w}\left(\frac{\xi'}{c}\right)\frac{\frac{1}{c^3}\hat{Z}_3\left(\frac{\xi}{c}, \frac{\xi'}{c}\right)}{\frac{1}{c^2}\hat{Z}_2\left(\frac{\xi}{c}\right)}\mathrm{d}\xi' + \hat{w}\left(\frac{\xi}{c}\right) \\ &= m + \int \hat{w}(\xi')\frac{\hat{Z}_3\left(\frac{\xi}{c}, \xi'\right)}{\hat{Z}_2\left(\frac{\xi}{c}\right)}\mathrm{d}\xi' + \hat{w}\left(\frac{\xi}{c}\right), \end{aligned} \tag{5.6}$$

and

$$\begin{aligned}\hat{b_2}\left(\frac{\xi}{c}\right) &= \int \frac{\mathrm{d}}{\mathrm{d}\xi'} c\hat{v}\left(\frac{\xi'}{c}\right) \frac{\frac{1}{c^3}\hat{Z}_3\left(\frac{\xi}{c},\frac{\xi'}{c}\right)}{\frac{1}{c^2}\hat{Z}_2\left(\frac{\xi}{c}\right)} \mathrm{d}\xi' + \frac{\mathrm{d}}{\mathrm{d}\xi} c\hat{v}\left(\frac{\xi}{c}\right) \\ &= \int \frac{\mathrm{d}}{\mathrm{d}\xi'} \hat{v}(\xi') \frac{\hat{Z}_3\left(\frac{\xi}{c},\xi'\right)}{\hat{Z}_2\left(\frac{\xi}{c}\right)} \mathrm{d}\xi' + \frac{\mathrm{d}}{\mathrm{d}\xi} c\hat{v}\left(\frac{\xi}{c}\right),\end{aligned} \tag{5.7}$$

involving a change of variables from ξ' to $c\xi'$. Setting $\eta = \frac{\xi}{c}$ in (5.5)-(5.7), it becomes clear that the steady state equation for $\hat{Z}_2(\eta)$ is of the same form as the steady state equation for $Z_2(\eta)$. It follows that when $\hat{Z}_2(\xi)$ is normalised by ${\hat{Z}_1}^2$ such that $\hat{Z}_2(\xi) = 1$ in the absence of spatial structure, its steady state solution will be a horizontal stretch or compression by a factor c of the normalised steady state solution $Z_2(\xi)$.

Thus, there is an equivalence between varying kernel width σ^2 and varying Z_1. A horizontal stretch in the kernels by a factor c leads to the same spatial structure as would increasing Z_1 by a factor c. The second moment $Z_2(\xi)$ is increased by a factor c^2 and horizontally stretched by a factor c. Thus, increasing the range of cell-cell interactions is equivalent to increasing the average density.

We can verify this numerically as shown in Figure 5.7. Figure 5.7(a) shows the average pair density at steady state when $c = 1$ such that $\hat{w}(\xi) = w(\xi)$ and $\frac{\mathrm{d}}{\mathrm{d}\xi}\hat{v}(\xi) = \frac{\mathrm{d}}{\mathrm{d}\xi}v(\xi)$. In comparison when $c = 0.5$ the width of the interaction kernels $\hat{w}(\xi)$ and $\hat{v}(\xi)$ are double that of $w(\xi)$ and $v(\xi)$. As expected, this results in a solution $\hat{Z}_2(\xi)$ which is a horizontal stretching of $Z_2(\xi)$ by a factor $c = 0.5$ as seen in Figure 5.7(b).

5.3 Numerical Methods

In this section we describe methods we have implemented in order to solve the equation for the second spatial moment (4.25) numerically.

5.3.1 Discretisation of Pair Displacement ξ

Equation (4.25) was solved using the method of lines with MATLAB's in-built ode45 solver, which is based on an explicit Runge-Kutta (4,5) formula, the Dormand-Prince pair [97]. Ode45 is a one-step solver, i.e. in computing the solution at a time point t_n, it requires only

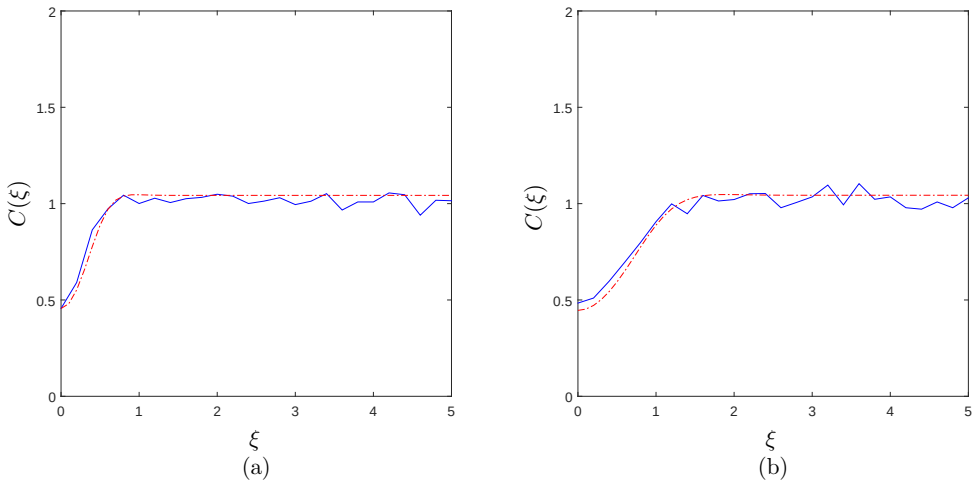


Figure 5.7: Effect of changing interaction kernel width on the steady state solution of the second moment. The PCF $C_{\mathrm{IBM}}(\xi)$ (blue solid) for 100 averaged realisations of the IBM is plotted against the PCF $C_{\mathrm{SM}}(\xi)$ (red broken) predicted by the spatial moment model at time $t = 25$ for two values of c. (a) $c = 1$. Collective movement of 200 cells in a domain of length $L = 500$ therefore $\hat{Z}_1 = 0.4$. Parameters are $\alpha = -2$, $\beta = 0.5$, $\sigma_1^2 = \sigma_2^2 = 0.08$, $\lambda_\mu = 5$, $m = 10$. (b) $c = 0.5$. Collective movement of 200 cells in a domain of length $L = 1000$ therefore $\hat{Z}_1 = 0.2$. Parameters are $\alpha = -2$, $\beta = 1$, $\sigma_1^2 = \sigma_2^2 = 0.32$, $\lambda_\mu = 2.5$, $m = 10$.

the solution at the immediately preceding time point t_{n-1}. The method of lines involved a discretisation of ξ with grid spacing Δ over the domain $|\xi| \leq \xi_{max}$, where ξ_{max} was large enough so that $Z_2(\xi) \approx Z_1^2$ at $|\xi| = \xi_{max}$. Required values of $Z_2(\xi)$ that lay outside of the computable domain were set to the value of $Z_2(\xi)$ at the boundary. The integral terms in (4.25) were approximated using the trapezoidal rule with the same discretisation.

The numerical results above used a grid spacing $\Delta = 0.1$ and $\xi_{max} = 5$ and were insensitive to a reduction in Δ. Some numerical error can arise in $Z_2(\xi)$ due to this discretisation of ξ and will accumulate over time. At steady state, any numerical error is most noticeable for large ξ where we know the exact solution is $Z_2(\xi) = 1$ but in the numerical solution $Z_2(\xi)$ the accumulated error can at worst lead to a deviation from 1 of the order 10^{-1}. The magnitude of this error depends on the choice of ξ_{max}. Increasing ξ_{max} incurs a loss of computational efficiency but reduces this numerical error. For instance increasing ξ_{max} from 5 to 10 reduces $Z_2(\xi)$ at $\xi = 5$ from 1.04 to 1.02. Decreasing Δ to give a finer grid does not significantly affect $Z_2(\xi)$ or the numerical error. In the above results we have chosen values of ξ_{max} and Δ which give a reasonably good match between $C_{\mathrm{IBM}}(\xi)$ and $C_{\mathrm{SM}}(\xi)$ at relatively low computational cost. However, if computational efficiency is not of concern then a more accurate solution can be found by choosing a greater ξ_{max} and smaller Δ.

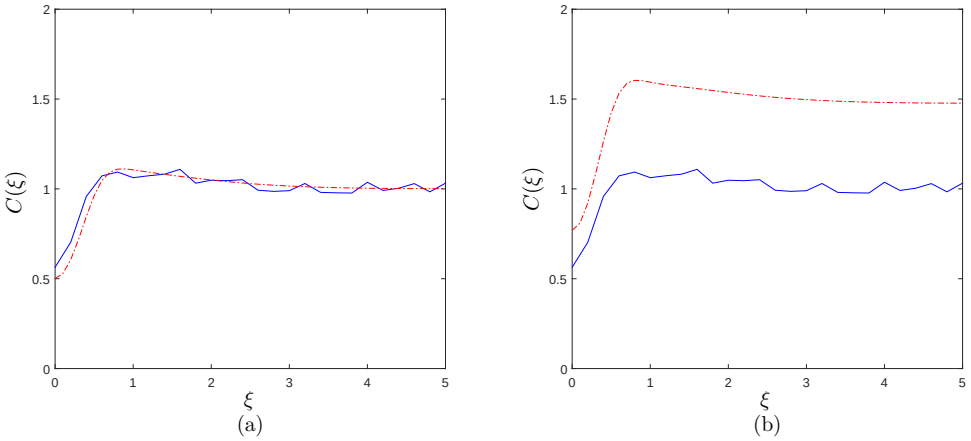


Figure 5.8: Effect of numerical normalisation on $Z_2(\xi)$. Collective movement of 200 cells in a domain of length $L = 500$ therefore $Z_1 = 0.4$. The PCF $C_{\text{IBM}}(\xi)$ (blue solid) for 100 averaged realisations of the IBM is plotted against the PCF $C_{\text{SM}}(\xi)$ (red broken) predicted by the spatial moment at time $t = 1$ for $\alpha = -2$, $\beta = 0.5$, $\sigma_1^2 = \sigma_2^2 = 0.08$, $\lambda_\mu = 5$, $m = 10$, (a) with numerical normalisation, and (b) without numerical normalisation of $\mu_2(\xi, \xi')$.

5.3.2 Numerical Normalisation of $\mu_2(\xi, \xi')$ in the Moment Equations

The PDF for movement $\mu_2(\xi, \xi')$ is normalised numerically using the trapezium rule such that $\int \mu_2(\xi, \xi')\mathrm{d}\xi = 1$. This involves division of $\mu_2(\xi, \xi')$ by the numerical normalisation constant

$$N = \frac{h\lambda}{2} \sum_k \left(\exp(\lambda\xi_k) + \exp(\lambda\xi_{k+1})\right), \tag{5.8}$$

where ξ_k denotes the k^{th} grid point in the discretised ξ domain. Figure 5.8 demonstrates the effect that this normalisation has on the accuracy of the numerical solution $Z_2(\xi)$. When $\mu_2(\xi, \xi')$ is normalised by N, Figure 5.8(a) shows $C_{SM}(\xi) \approx 1$ for large ξ at $t = 1$ as expected. Under the same conditions but in the absence of numerical normalisation, $C_{SM}(\xi) \approx 1.5$ for large ξ at $t = 1$ (Fig. 5.8(b)). Over time the solution deteriorates further as $Z_2(\xi)$ continues to increase across all ξ which induces other irregularities such as $M_2(\xi) < 0$.

5.3.3 Truncation of Kernels in the Moment Equations

The interaction kernels $w(\xi)$ and $v(\xi)$, and movement PDF $\mu_2(\xi, \xi')$ are defined over an infinite domain, therefore truncation of the kernels is necessary for numerical integration of the moment equations. We truncate the kernel tails such that $w(\xi) = v(\xi) = \mu_2(\xi, \xi') = 0$ over the domain $\{\xi_k \mid 1 \leq k \leq 5,\ n-4 \leq k \leq n\}$, where ξ_k denotes the k^{th} grid point and n is the number of grid points in the discretised ξ domain. We ensure that $1/\lambda_\mu < 4\sigma_w, 4\sigma_v < \xi_{n-4}$.

5.4 Discussion

The IBM enables us to simulate the stochastic behaviour of cells undergoing collective movement and our numerical results demonstrate how individual-level interactions give rise to the development of spatial structure in the population. To obtain an accurate description of average behaviour we either simulate movement for a large number of cells or average the results from many realisations of the model. This approach becomes computationally intensive when cell abundance is high because the interactions between each individual and all of its neighbours must be calculated before every movement event. In addition, IBMs are not directly amenable to mathematical analysis. Therefore there is good motivation for a population-level model in terms of spatial moment dynamics which provides mechanistic insight into how local directional bias gives rise to spatial structure and creates scope for a more formal analysis of the underlying stochastic process [61].

Unlike models which employ the mean-field assumption, our spatial moment model takes into account the effects of local spatial structure on the dynamics of the population. The second moment predicted by the model, expressed as a PCF, provides a measure of this structure and can be directly compared to PCFs calculated from images of *in vitro* cell migration experiments [38]. Our numerical results show that the moment model can provide a good approximation of the spatial structure predicted by the IBM when the distribution of cells is homogeneous throughout space. In the case where interactions affect neighbour-dependent motility but not the direction of movement, the two models mostly match very well both when motility rate is increased in close proximity to neighbours or when it is decreased. However, when interactions that decrease cell motility are strong the moment model tends to under-predict the second moment. This is likely due to it not accounting for the higher incidence of negative motility rates that can arise by chance in the IBM. One possible solution to this issue would be to use an alternative form of movement rate in which the neighbourhood effects are non-linear, for example an exponential decay of movement rate with local crowding. However, deriving the spatial moment approximation for a non-linear movement rate would be far from straightforward.

When interactions determine only direction of movement and do not affect movement rate the two models again correspond well except when cells are strongly biased to move towards one another. In this case, the spatial moment model over-predicts the second moment. As the

motility rate is constant, a rise in negative motility rates can not be causing the disparity in this case. Instead, it is likely that the performance of the power-3 closure as an approximation to the third moment breaks down when the spatial pattern becomes strongly clustered. If this is the case, the spatial moment model may also be over-predicting the second moment in the case where strong interactions with close neighbours reduce motility rate and generate clustering. However, it is possible that the high average pair densities that arise in the IBM, due to the increased incidences of negative motilities, could be masking the effect.

When interactions influence both motility rate and directional bias, the results from the IBM and the spatial moments still correspond well in most cases. The two models only start to disagree when the second moment is large for short displacements. In this case the mechanisms that we have seen cause disparity between $C_{\mathrm{SM}}(\xi)$ and $C_{\mathrm{IBM}}(\xi)$ when interactions affecting motility rate and directional bias are considered in isolation, may both contribute to the mis-match in results. However, we have shown that in general our spatial moment model provides a good approximation to the underlying IBM and only starts to break down when the spatial pattern becomes strongly clustered.

In the cell movement model proposed by Middleton et al. [71], neighbour interactions are affected by both speed and direction of movement. This differs from our model where average speed and direction are treated as independent effects. Despite these differences, we see similar qualitative patterns of spatial structure due to local repulsive or attractive forces as in Middleton et al [71].

The closure for the third spatial moment is only an approximation and different closures may have better performance under different conditions. A number of asymmetric power-2 closures have performed well for other models discussed in the literature [65,70,79], although they do have the potential to violate the positivity constraint which is required because an average density of triplets can never be negative [79]. For our model, the power-3 closure outperformed the four different power-2 closures that were tested for a strongly clustered pattern, however it is possible that other weightings or closures may improve the approximation. While it is important to keep this in mind, an exhaustive analysis of moment closures is outside the scope of this work.

As our model does not incorporate volume exclusion, for example through the representation of cells as hard objects, there is the possibility that cell locations may arise in very close proximity in the IBM. The use of an interaction kernel which is concentrated around

short pair displacements provides a mechanism for generating a regular spatial pattern and thus allows us to reduce the likelihood of two cells being found close together. However, this approach is probabilistic and does not altogether exclude the possibility of such an occurrence.

It is appealing to consider the collective movement of cells in one dimension from a theoretical perspective, in particular because it simplifies the derivation and numerical solution of the spatial moments description. Solving the differential equation in two dimensions is considerably more computationally intensive as it involves double integrations in both the x- and y- direction. While the majority of experimental data is two- or three-dimensional, our results suggest that a one-dimensional model could still prove useful for quantifying the behaviour of moving cells. In one dimension we observe traits in the second moment that we would expect to see in a live population of cells, namely the development of clusters or regular spatial patterns. However, as a cell moving through two-dimensional space is interacting with neighbours in all directions, not just those on either side, it is possible that this could have a more profound impact on spatial structure than is predicted by a one-dimensional model. Therefore, in the next chapter we present numerical results for the model extended to two spatial dimensions.

Chapter 6

Two-dimensional Model and *In Vitro* Data

In this chapter, we take a combined experimental and modelling approach to explore collective behaviour in two-dimensional space. We first extend our original IBM to an equivalent two-dimensional model, then provide descriptions for the corresponding terms in the spatial moment dynamics. As for the one-dimensional model, the dynamics of the second spatial moment are solved for a homogeneous space. Using imaging data from *in vitro* experiments, we estimate parameters for the model and show that it can generate similar spatial structure to that observed in a 3T3 fibroblast cell population.

6.1 Two-Dimensional Individual-Based Model

In two spatial dimensions we define a continuous bounded domain $\Omega \in \mathbb{R}^2$, with periodic conditions at the boundaries. The location of a cell i is a coordinate $\mathbf{x}_i \in \mathbb{R}^2$. The movement rate ψ_i is still expressed by (3.8), however the kernel $w(\mathbf{z})$, which weights the strength of interaction between a pair of cells displaced by $\mathbf{z} = (z_1, z_2)^T$, is a two-dimensional Gaussian

$$w(\mathbf{z}) = \alpha \exp\left(-\frac{|\mathbf{z}|^2}{2\sigma_w^2}\right). \tag{6.1}$$

We now describe a mechanism for neighbour-dependent directional bias which is comparable to that outlined in Section 3.2.2 but with some differences that are required for extension to two spatial dimensions. The neighbour-dependent bias $\mathbf{b}(\mathbf{x})$ accounts for the effect of n neighbouring cells located at $\mathbf{x}_j$ on the direction of movement of an individual at $\mathbf{x}$:

$$\mathbf{b}(\mathbf{x}) = \sum_{j=1}^{n} \mathbf{u}(\mathbf{x}, \mathbf{x_j}), \tag{6.2}$$

where

$$\mathbf{u}(\mathbf{x}, \mathbf{y}) = \nabla v(\mathbf{y} - \mathbf{x})\hat{h}(\mathbf{x}). \tag{6.3}$$

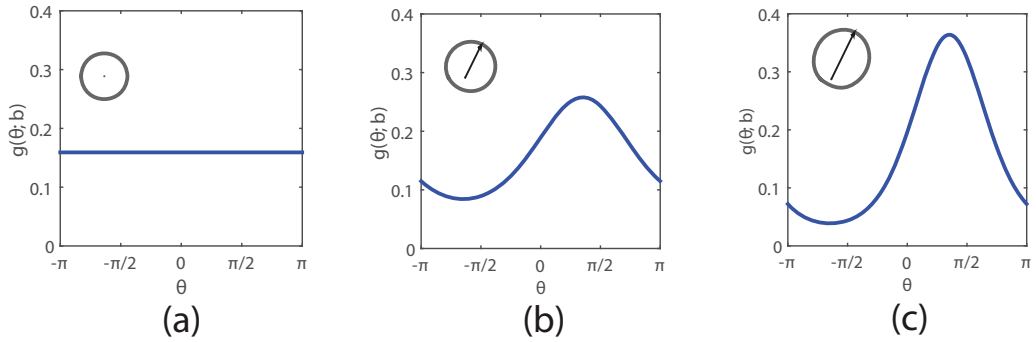


Figure 6.1: Examples of probability density function $g(\theta; \mathbf{b})$ (blue solid line) for movement in a direction $\theta \in [0, 2\pi]$. The neighbour-dependent bias $\mathbf{b}$ is a vector indicating the direction ($\arg(\mathbf{b})$) in which the greatest/lowest degree of crowding arises in a cell's neighbourhood, as well as the extent to which it occurs ($|\mathbf{b}|$). Insets are schematics illustrating $g(\theta; \mathbf{b})$ (grey solid line), where black arrows indicate the direction ($\arg(\mathbf{b})$) in which an individual (black dot) is most biased to move. (a) Unbiased movement; (b) weak directional bias $\mathbf{b} = (0.25, 0.5)^T$; (c) strong directional bias $\mathbf{b} = (0.5, 1)^T$.

Here, $\nabla v(\mathbf{z})$ denotes the gradient vector $\nabla v(\mathbf{z}) = (\frac{\partial v}{\partial z_1}, \frac{\partial v}{\partial z_2})^T$. The kernel $v(\mathbf{z})$ weights the strength of interaction between a pair of cells displaced by $\mathbf{z}$ and is a two-dimensional Gaussian function

$$v(\mathbf{z}) = \beta \exp\left(-\frac{|\mathbf{z}|^2}{2\sigma_v^2}\right), \tag{6.4}$$

so, as before, the interaction will be strong for a pair of cells located close together and negligible if they are far apart. In two dimensions, the neighbour-dependent bias $\mathbf{b}(\mathbf{x})$ is a vector holding information about both the extent and direction of crowded regions in the neighbourhood of a cell at x. We use the angle $\arg(\mathbf{b}(\mathbf{x}))$ to describe the direction of $\mathbf{b}(\mathbf{x})$. When $\beta > 0$, $\arg(\mathbf{b}(\mathbf{x}))$ is the direction in which the lowest degree of cell crowding arises locally. Conversely for $\beta < 0$, $\arg(\mathbf{b}(\mathbf{x}))$ is the direction of greatest local crowding. The magnitude $|\mathbf{b}(\mathbf{x})|$ provides a measure of the extent of crowding.

When a cell moves, its direction of movement $\theta \in [0, 2\pi]$ is drawn from a PDF $g(\theta; \mathbf{b})$ which depends on the neighbour-dependent bias $\mathbf{b}(\mathbf{x})$. The function $g(\theta; \mathbf{b})$ is a von Mises distribution with mean $\arg(\mathbf{b})$ and concentration $|\mathbf{b}|$:

$$g(\theta; \mathbf{b}) = \frac{\exp\Big(|\mathbf{b}|\cos\big(\theta - \arg(\mathbf{b})\big)\Big)}{2\pi I_0(|\mathbf{b}|)}, \tag{6.5}$$

where I_0 is the modified Bessel function of order 0. Thus, a cell is most likely to move in the direction $\arg(\mathbf{b})$ and the strength of this directional bias increases with $|\mathbf{b}|$. The PDF $g(\theta; \mathbf{b})$ is illustrated in Figure 6.1 for different examples of bias vector $\mathbf{b}(\mathbf{x})$.

In our one-dimensional IBM, the distance moved by a cell was drawn from an exponential distribution with mean step length $1/\lambda_\mu$. However, in two spatial dimensions this choice of distribution causes the spatial moment model to under-predict the extent of spatial structure generated by the directional bias, as shown in Figure 6.2(a). If, instead, the distance moved by a cell is drawn from a non-negative normal distribution with mean step length $1/\lambda_\mu$ and variance σ_μ^2, the spatial moment model provides a much better approximation to the IBM (Fig. 6.2(b)). It is likely that this discrepancy will arise for choices of movement PDF $\mu(\mathbf{x}, \mathbf{y})$ that are concentrated around very short displacements $\mathbf{y} - \mathbf{x}$. We discuss this limitation in Section 6.5.

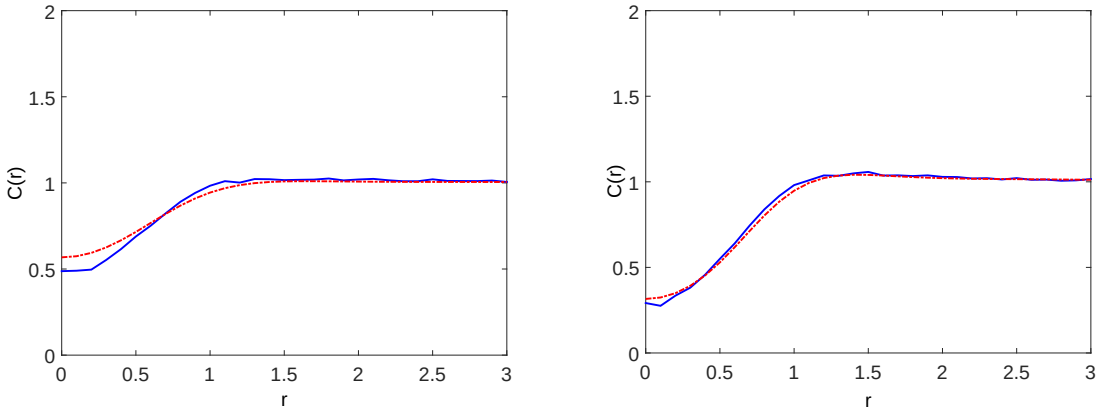
Therefore, in the two-dimensional model, the distance moved by a cell is drawn from a non-negative normal distribution with mean step length $1/\lambda_\mu$ and variance σ_μ^2. Thus, the probability of an individual at $\mathbf{x}$ moving to a new location at $\mathbf{y}$ is distributed according to

$$\mu(\mathbf{x}, \mathbf{y}) = N \exp\left(-\frac{\left(|\mathbf{y} - \mathbf{x}| - \frac{1}{\lambda_\mu}\right)^2}{2\sigma_\mu^2} \right) g\big(\arg(\mathbf{y} - \mathbf{x}); \mathbf{b}(\mathbf{x})\big). \tag{6.6}$$

This means that a cell at $\mathbf{x}$ is biased to move away from close-lying neighbours when $\beta > 0$. From a biological perspective this repulsive force could correspond to, for example, movement in response to a cell-released chemorepellant [24] or physical forces due to deformation of the cell membrane under direct contact with other cells [20]. When $\beta < 0$ the bias is towards crowded regions, such as might arise in the presence of a cell-released chemoattractant [25]. The bias strength increases with increasing neighbourhood cell density. In the absence of neighbourhood interactions (i.e. setting $\beta = 0$), $g\big(\arg(\mathbf{y} - \mathbf{x}); \mathbf{b}(\mathbf{x})\big) = 1/(2\pi)$ and the cell is equally likely to move in any direction, i.e. movement is unbiased. The PDF $\mu(\mathbf{x}, \mathbf{y})$ has dimensions L^{-2} and normalising by the constant N satisfies the constraint $\int \mu(\mathbf{x}, \mathbf{y}) \mathrm{d}\mathbf{y} = 1$ for any fixed $\mathbf{x}$.

6.2 Spatial Moments in Two Dimensions

The equations for the spatial moment dynamics can easily be extended to consider movement in two-dimensional space. However, solving the differential equation numerically in two dimensions is more computationally intensive because, for a location $\mathbf{x} = (x_1, x_2)^T$, each integration over $\mathbf{x}$ involves an integration over both x_1 and x_2. In two spatial dimensions, the



(a) (b)

Figure 6.2: Performance of spatial moment approximation using two choices of movement PDF $\mu(\mathbf{x}, \mathbf{y})$. Spatial structure for 200 cells undergoing collective movement with neighbour-dependent bias ($\alpha = 0$) in a 20 x 20 domain at time $t = 5$. A PCF $C_{IBM}(r)$ (blue solid line) is computed (using a bin width $\delta r = 0.1$) by averaging results from 500 simulations of the IBM. This is compared to the PCF $C_{SM}(r)$ (red dashed line) approximated by the spatial moment model, solved using $\Delta = 0.1$ and $\xi_{max} = 4$. The result was insensitive to a finer grid spacing Δ. Parameters are $\beta = 0.2$, $\sigma_w = \sigma_v = 0.5$, $m = 10$, $\lambda_\mu = 5$, $\sigma_\mu = 0.05$; (a) Distance moved by a cell is drawn from an exponential distribution with mean step length $1/\lambda_\mu$; (b) Distance moved by a cell is drawn from a non-negative normal distribution with mean step length $1/\lambda_\mu$ and variance σ_μ^2.

only expression that differs from the one-dimensional description is the movement PDF. We use equation (6.6) from the two-dimensional IBM to derive $\mu_1(\mathbf{x}, \mathbf{y})$ and $\mu_2(\mathbf{x}, \mathbf{y}, \mathbf{z})$, required for the dynamics of the first and second spatial moments respectively.

For the dynamics of the first spatial moment, when a cell at $\mathbf{x}$ moves, its new location $\mathbf{y}$ is drawn from a PDF

$$\mu_1(\mathbf{x}, \mathbf{y}) = N \exp\left(-\frac{\left(|\mathbf{y} - \mathbf{x}| - \frac{1}{\lambda_\mu} \right)^2}{2\sigma_\mu^2} \right) g\big(\arg(\mathbf{y} - \mathbf{x}); \mathbf{b}_1(\mathbf{x})\big). \tag{6.7}$$

The neighbour-dependent bias $\mathbf{b}_1(\mathbf{x})$ for a cell at $\mathbf{x}$ is

$$\mathbf{b}_1(\mathbf{x}) = \int u(\mathbf{x}, \mathbf{y}) \frac{Z_2(\mathbf{x}, \mathbf{y})}{Z_1(\mathbf{x})} \mathrm{d}\mathbf{y}. \tag{6.8}$$

The rate of change of the first spatial moment is given by

$$\frac{\mathrm{d}Z_1(\mathbf{x}, t)}{\mathrm{d}t} = -M_1(\mathbf{x}) Z_1(\mathbf{x}, t) + \int \mu_1(\mathbf{u}, \mathbf{x}) M_1(\mathbf{u}) Z_1(\mathbf{u}, t) \mathrm{d}\mathbf{u}. \tag{6.9}$$

For the dynamics of the second spatial moment, when a cell at $\mathbf{x}$ moves, its new location

$\mathbf{y}$ is drawn from a PDF $\mu_2(\mathbf{x}, \mathbf{y}, \mathbf{z})$. The third argument in $\mu_2(\mathbf{x}, \mathbf{y}, \mathbf{z})$ accounts for the fact that $\mathbf{x}$ is in a pair with a cell at $\mathbf{z}$:

$$\mu_2(\mathbf{x}, \mathbf{y}, \mathbf{z}) = N \exp\left(-\frac{\left(|\mathbf{y} - \mathbf{x}| - \frac{1}{\lambda_\mu}\right)^2}{2\sigma_\mu^2} \right) \; g\big(\arg(\mathbf{y} - \mathbf{x}); \mathbf{b}_2(\mathbf{x}, \mathbf{z})\big). \tag{6.10}$$

The neighbour-dependent bias $\mathbf{b}_2(x, y)$ for a cell at $\mathbf{x}$ in a pair with a cell at $\mathbf{y}$ is

$$\mathbf{b}_2(\mathbf{x}, \mathbf{y}) = \int u(\mathbf{x}, \mathbf{z}) \frac{Z_3(\mathbf{x}, \mathbf{y}, \mathbf{z})}{Z_2(\mathbf{x}, \mathbf{y})} \mathrm{d}\mathbf{z} + u(\mathbf{x}, \mathbf{y}). \tag{6.11}$$

Finally, the rate of change for the second moment in two dimensions is

$$\begin{aligned} \frac{\mathrm{d}Z_2(\mathbf{x}, \mathbf{y}, t)}{\mathrm{d}t} \;=\; & -\left(M_2(\mathbf{x}, \mathbf{y}) + M_2(\mathbf{y}, \mathbf{x})\right) Z_2(\mathbf{x}, \mathbf{y}, t) \\ & + \int \mu_2(\mathbf{u}, \mathbf{x}, \mathbf{y}) M_2(\mathbf{u}, \mathbf{y}) Z_2(\mathbf{u}, \mathbf{y}, t) \mathrm{d}\mathbf{u} \\ & + \int \mu_2(\mathbf{u}, \mathbf{y}, \mathbf{x}) M_2(\mathbf{u}, \mathbf{x}) Z_2(\mathbf{u}, \mathbf{x}, t) \mathrm{d}\mathbf{u}. \end{aligned} \tag{6.12}$$

6.2.1 Spatially Homogeneous Moment Dynamics

In two dimensions, (6.12) can also be expressed in terms of (4.25), (4.26) and (4.29) with the displacements $\boldsymbol{\xi}, \boldsymbol{\xi}'$ and $\boldsymbol{\xi}'' \in \mathbb{R}^2$.

The movement PDF given in (6.10) becomes

$$\mu_2(\boldsymbol{\xi}, \boldsymbol{\xi}') = N \exp\left(-\frac{\left(|\boldsymbol{\xi}| - \frac{1}{\lambda_\mu}\right)^2}{2\sigma_\mu^2} \right) \; g\big(\arg(\boldsymbol{\xi}); \mathbf{b}_2(\boldsymbol{\xi}')\big). \tag{6.13}$$

where

$$\mathbf{b}_2(\boldsymbol{\xi}) = \int \nabla v(\boldsymbol{\xi}') \frac{Z_3(\boldsymbol{\xi}, \boldsymbol{\xi}')}{Z_2(\boldsymbol{\xi})} \mathrm{d}\boldsymbol{\xi}' + \nabla v(\boldsymbol{\xi}). \tag{6.14}$$

The boundary condition is given by

$$Z_2(\boldsymbol{\xi}) \to Z_1^2 \text{ as } |\boldsymbol{\xi}| \to \infty. \tag{6.15}$$

6.3 Numerical Results

To explore whether our model is capable of generating spatial structure in a simulated cell population we average results from repeated simulations of the IBM and compute a periodic PCF $C_{IBM}(r)$. We compare this to numerical solutions of our spatial moment model to examine whether it provides a good approximation to the underlying stochastic process. We consider a case where the distribution of cells is spatially homogeneous, therefore the spatially-dependent weight $\hat{h}(x)$ in (6.3) is $\hat{h}(x) = 1$. The equation for the dynamics of the second spatial moment can be expressed in terms of displacements between pairs of cells, as outlined in Section 4.6. The PCF $C_{SM}(\boldsymbol{\xi})$ is given by $Z_2(\boldsymbol{\xi})/Z_1^2$ such that $C_{SM}(\boldsymbol{\xi}) = 1$ in the complete absence of spatial structure. The second spatial moment is radially symmetric about the origin of $\boldsymbol{\xi}$. Therefore, in the results below we show only a radial section of $C_{SM}(\boldsymbol{\xi})$ which we denote $C_{SM}(r)$, where $r = |\boldsymbol{\xi}|$. Cells are initially distributed across a domain of width L_x and height L_y, according to a spatial Poisson process with intensity $n/(L_x L_y)$. In the spatial moment model this corresponds to $Z_2(\boldsymbol{\xi}) = Z_1^2$ at $t = 0$. The system is allowed to reach steady state before results from each model are compared.

6.3.1 Numerical Methods

Equation (4.25) for the dynamics of the second spatial moment was solved numerically using the method of lines with MATLAB's in-built ode23 one-step solver, which is an implementation of an explicit Runge-Kutta (2,3) pair of Bogacki and Shampine [98]. The method of lines involved a discretisation of $\boldsymbol{\xi} = (\xi_1, \xi_2)^T$ with grid spacing Δ over the domain $\{-\xi_{max} \leq \xi_1, \xi_2 \leq \xi_{max}\}$, where ξ_{max} was large enough so that $Z_2(\boldsymbol{\xi}) \approx Z_1^2$ at the boundary. Required values of $Z_2(\boldsymbol{\xi})$ that lay outside of the computable domain were set to the value of $Z_2(\boldsymbol{\xi})$ at a corner of the boundary, i.e. $Z_2(\xi_{max}, \xi_{max})$. The integral terms in (4.25) were approximated using the trapezium rule with the same discretisation. In addition, the PDF for movement $\mu_2(\boldsymbol{\xi}, \boldsymbol{\xi'})$ was normalised numerically using the trapezium rule such that $\int \mu_2(\boldsymbol{\xi}, \boldsymbol{\xi'}) \mathrm{d}\boldsymbol{\xi} = 1$ for any fixed $\boldsymbol{\xi'}$. The results were insensitive to a reduction in grid spacing Δ.

The interaction kernels $w(\boldsymbol{\xi})$ and $v(\boldsymbol{\xi})$, and movement PDF $\mu_2(\boldsymbol{\xi}, \boldsymbol{\xi'})$ are defined over an infinite domain, therefore truncation of the kernels is necessary for the numerical integration of the moment equations. We truncate the tails of the kernels such that $w(\boldsymbol{\xi}) = v(\boldsymbol{\xi}) =$

$\mu_2(\boldsymbol{\xi}, \boldsymbol{\xi'}) = 0$ over the domain $|\boldsymbol{\xi}| > \xi_{max}/2$ and we ensure that $(1/\lambda_\mu + 4\sigma_\mu) \leq 4\sigma_w, 4\sigma_v \leq \xi_{max}/2$.

6.3.2 Moment Closure

In our one-dimensional model we employed a power-3 closure to truncate the system of spatial moment dynamics at second order. This closure performed well as an approximation to the third spatial moment, except for a strongly clustered spatial pattern. We now compare the performance of the power-3 closure with a number of others for our two-dimensional model, to consider whether this choice is still appropriate for describing movement in two-dimensional space. In addition to the power-2 closure given previously in equation (5.1), we explore the performance of the symmetric power-1 closure, given by

$$\tilde{Z}_3(\boldsymbol{\xi}, \boldsymbol{\xi'}) = Z_1 Z_2(\boldsymbol{\xi'} - \boldsymbol{\xi}) + Z_1 Z_2(\boldsymbol{\xi'}) + Z_1 Z_2(\boldsymbol{\xi}) - 2Z_1^3. \tag{6.16}$$

Figure 6.3 compares the predictions of five different closures for a cluster spatial pattern, where the performance of the approximations tends to be most limited. The closures that have the best performance in this case are the power-3 closure and the asymmetric power-2 closure with weighting {1,3,8}, while the power-1 closure provides the poorest approximation overall. In the following numerical results we choose to continue employing the power-3 closure. The power-2 (1,3,8) closure outperforms the power-3 very slightly for this cluster spatial pattern, however its negative term means that it can violate the positivity constraint of the third spatial moment [79].

6.3.3 Non-Interacting Model

In the complete absence of interactions, an individual's direction of movement is unbiased and its movement rate is solely determined by the intrinsic component. It is straightforward to show analytically that the steady-state solution for $Z_2(\boldsymbol{\xi})$ is a constant under these conditions. Numerical solutions and averaged IBM simulations confirm this.

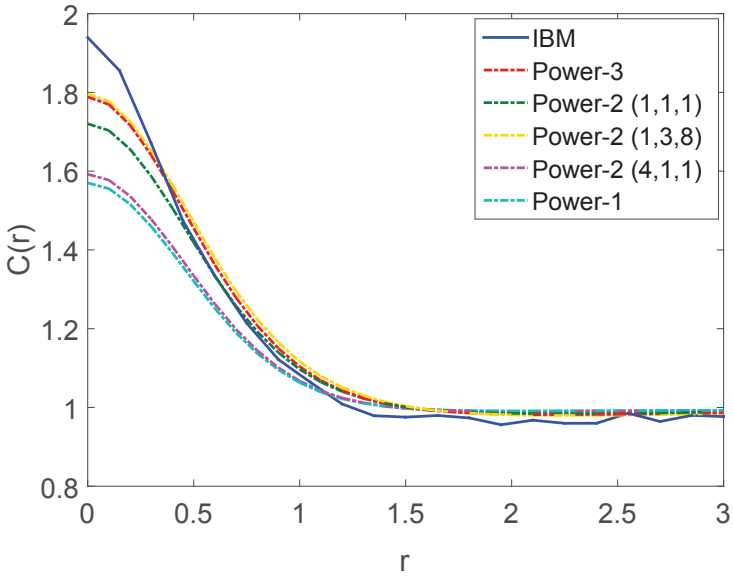


Figure 6.3: Comparison of closures for the third spatial moment. Spatial structure for 200 cells undergoing collective movement with neighbour-dependent directional bias ($\alpha = 0$) in a 20 x 20 domain at time $t = 25$. A PCF $C_{IBM}(r)$ (blue solid line) is computed (using a bin width $\delta r = 0.15$) by averaging results from 500 simulations of the IBM. This is compared to the PCF $C_{SM}(r)$ approximated by the spatial moment model ($\Delta = 0.1$ and $\xi_{max} = 4$) solved using a power-3 closure (red dashed line), power-2 closure with weighting {1,1,1} (green dashed line), power-2 closure with weighting {1,3,8} (yellow dashed line), power-2 closure with weighting {4,1,1} (purple dashed line) and power-1 closure (cyan dashed line). Parameters are $\beta = -0.05$, $\sigma_w = \sigma_v = 0.5$, $m = 10$, $\lambda_\mu = 5$, $\sigma_\mu = 0.05$.

6.3.4 Interacting Model

6.3.4.1 Neighbour-Dependent Directional Bias

The effect of the neighbour-dependent directional bias, in the absence of neighbour-dependent motility, is shown in Figure 6.4. Regular spatial patterns are generated by the directional bias when $\beta > 0$ while $\beta < 0$ gives rise to clustering. The spatial moment model performs very well as an approximation to the IBM except when there is strong clustering (Fig 6.4(d)). This can likely be attributed to limitations of the moment-closure assumption. The power-3 closure provides a reasonable approximation to the third moment for Poisson spatial patterns and regular patterns, but performs quite poorly for cluster spatial patterns where it can cause the model to underestimate the second moment [79, 85, 96].

6.3.4.2 Neighbour-Dependent Motility

Figure 6.5 shows the spatial structure generated by the mechanism for neighbour-dependent motility when there is no local directional bias (i.e. $\beta = 0$). Neighbourhood interactions

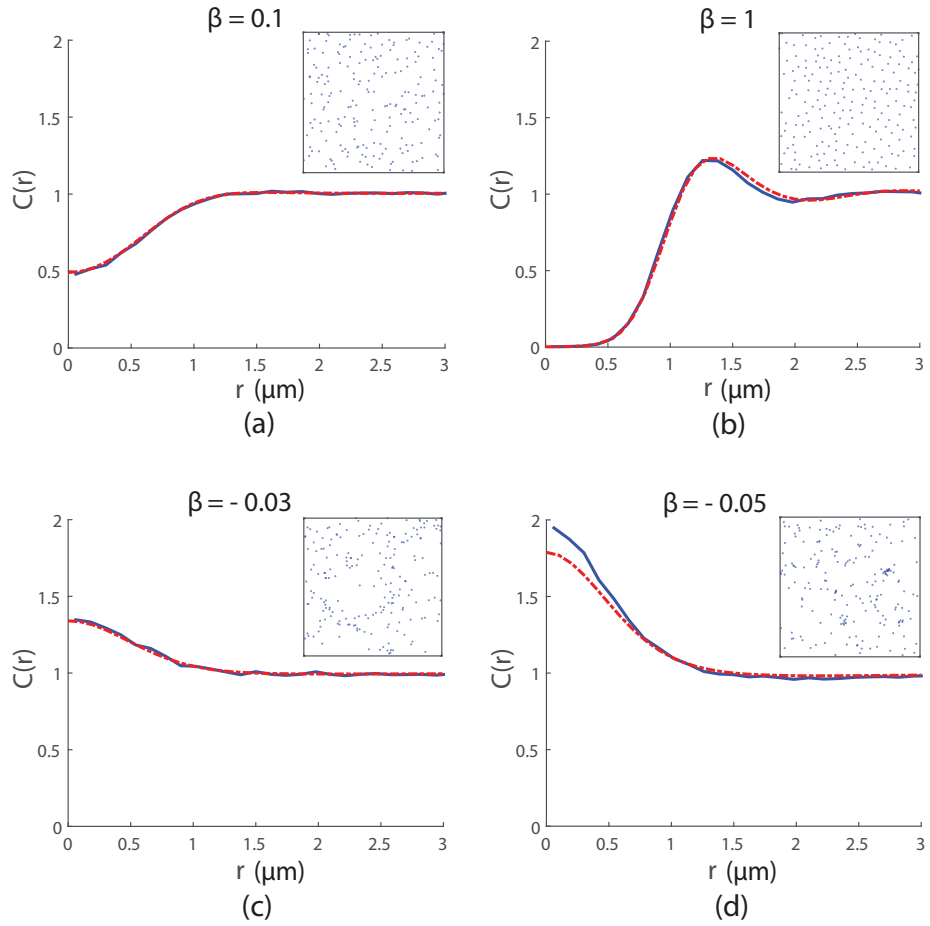


Figure 6.4: Spatial structure for 200 cells undergoing collective movement with neighbour-dependent directional bias ($\alpha = 0$ hours^{-1}) in a 20 μm x 20 μm domain at time $t = 25$ hours. Insets show IBM simulation snapshots of cell locations (blue dots). A PCF $C_{IBM}(r)$ (blue solid line) is computed (using a bin width $\delta r = 0.12$ μm) by averaging results from 500 simulations of the IBM. This is compared to the PCF $C_{SM}(r)$ (red dashed line) approximated by the spatial moment model, solved using $\Delta = 0.1$ μm and $\xi_{max} = 4$ μm. Parameters are $\alpha = 0$ hours^{-1}, $\sigma_w = \sigma_v = 0.5$ μm, $m = 10$ hours^{-1}, $\lambda_\mu = 5$ μm^{-1}, $\sigma_\mu = 0.05$ μm; (a) $\beta = 0.1$ μm; (b) $\beta = 1$ μm; (c) $\beta = -0.03$ μm; (d) $\beta = -0.05$ μm.

give rise to regular spatial patterns when $\alpha > 0$ and cluster spatial patterns when $\alpha < 0$. Again, we see good agreement between $C_{SM}(r)$ and $C_{IBM}(r)$ except for large magnitudes of $\alpha < 0$ where the pattern is clustered and the moment model under-predicts spatial structure (Fig. 6.5(d)). While the limitations associated with the moment closure may play a role, there is another factor that could also be contributing to the poor fit here. We have chosen values of α such that the probability of $\psi_i > 0$ is high. However $\psi_i = 0$ can arise by chance in an IBM simulation and while such occurrences are relatively rare they can have a self-propagating effect, leading to strong clustering. The spatial moment model does not account for these chance events so this might explain why spatial structure is underestimated more dramatically even for relatively weak clustering. As in the one-dimensional case, for large

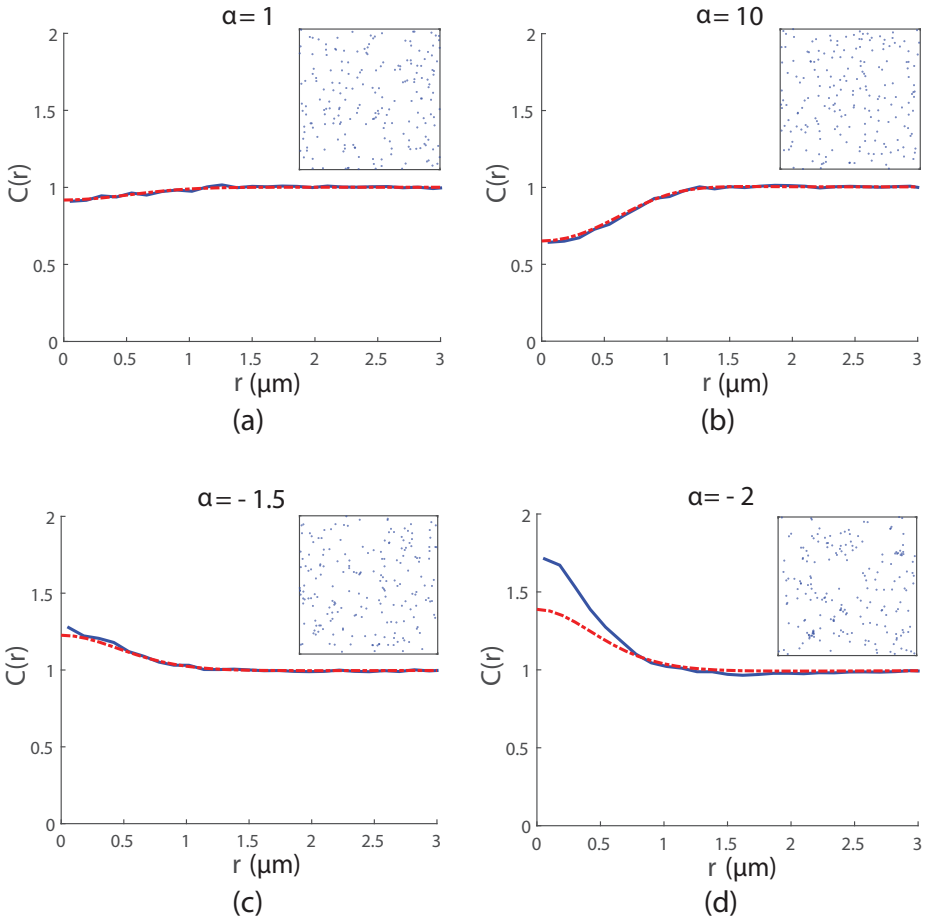
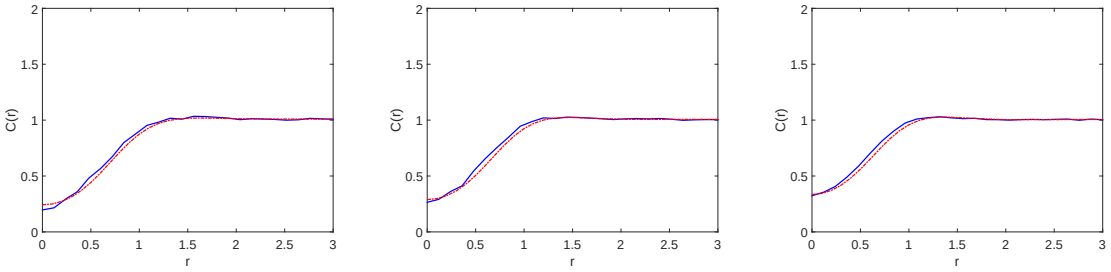


Figure 6.5: Spatial structure for 200 cells undergoing collective movement with neighbour-dependent motility ($\beta = 0$ μm) in a 20 μm x 20 μm domain at time $t = 25$ hours. Insets show IBM simulation snapshots of cell locations (blue dots). A PCF $C_{IBM}(r)$ (blue solid line) is computed (using a bin width $\delta r = 0.12$ μm) by averaging results from 500 simulations of the IBM. This is compared to the PCF $C_{SM}(r)$ (red dashed line) approximated by the spatial moment model, solved using $\Delta = 0.1$ μm and $\xi_{max} = 4$ μm. Parameters are $\beta = 0$ μm, $\sigma_w = \sigma_v = 0.5$ μm, $m = 10$ hours^{-1}, $\lambda_\mu = 5$ μm^{-1}, $\sigma_\mu = 0.05$ μm; (a) $\alpha = 1$ hours^{-1}; (b) $\alpha = 10$ hours^{-1}; (c) $\alpha = -1.5$ hours^{-1}; (d) $\alpha = -2$ hours^{-1}.

magnitudes of $\alpha < 0$ there is an equilibrium where all individuals have motility $\psi_i = 0$ which is reached with probability 1 if the IBM is simulated for a sufficiently long time.

6.3.4.3 Average Cell Density

We now consider what effect varying the average cell density has on spatial structure, in the presence of both neighbour-dependent motility and directional bias. Figure 6.6 compares the spatial structure for a low (Fig. 6.6(a)), moderate (Fig. 6.6(b)) and high average cell density (Fig. 6.6(c)). Here, directional bias is the stronger structure-generating mechanism and promotes a regular spatial pattern. However, as average cell density is increased (Fig. 6.6(a)-(c) from left to right), the spatial structure weakens. Further increase in average cell



(a) (b) (c)

Figure 6.6: Spatial structure at three different average cell densities for 200 cells undergoing collective movement at time $t = 25$. A PCF $C_{IBM}(r)$ (blue solid line) is computed (using a bin width $\delta r = 0.12$) by averaging results from 500 simulations of the IBM. This is compared to the PCF $C_{SM}(r)$ (red dashed line) approximated by the spatial moment model, solved using $\Delta = 0.1$ and $\xi_{max} = 4$. Parameters are $\alpha = -1$, $\beta = 0.2$, $\sigma_w = \sigma_v = 0.5$, $m = 10$, $\lambda_\mu = 5$, $\sigma_\mu = 0.05$; (a) Low average cell density $Z_1 = 0.25$; (b) Moderate average cell density $Z_1 = 0.5$; (c) High average cell density $Z_1 = 0.75$.

density causes the structure to tend towards a Poisson spatial pattern.

6.4 *In Vitro* Experimental Data

6.4.1 Experimental Methods

6.4.1.1 Cell Culture

Murine fibroblast 3T3 cells were cultured in Dulbecco's modified Eagle medium (Invitrogen, Australia) with 5% foetal calf serum (FCS) (Hyclone, New Zealand), 2 mM L-glutamine (Invitrogen), 50U/ml penicillin and 50μg/ml streptomycin (Invitrogen), in 5% CO_2 and 95% air at 37 °C. Monolayers of 3T3 cells were cultured in T175 cm^2 tissue culture flasks (Nunc, Thermo Scientific, Denmark). Prior to confluence, cells were lifted with 0.05% trypsin (Invitrogen). Viable cells were counted using the trypan blue exclusion test and a haemocytometer.

Two cell suspensions were created at approximate average cell densities of 20,000 cells/ml and 30,000 cells/ml. The experiments were performed in triplicate for each initial cell density. Cells were seeded in a 24 well tissue culture plate (each well of diameter 15.6 mm) and incubated overnight in 5% CO_2 and 95% air at 37 °C to allow them to attach to the base of the plate. Initially, cells were approximately uniformly distributed in each well.

6.4.1.2 Imaging Techniques and Analysis

Time-lapse images of the cells were captured, over a period of 12 hours at 3 hour intervals, using a light microscope and Eclipse TIS software at 100x magnification. For each sample, a 4500 μm x 450 μm image was reconstructed from overlapping adjacent images captured at approximately the centre of the well. The locations of the n cells in each image were manually determined by superimposing markers onto cells and recording the Cartesian coordinates of markers using ImageJ image analysis software. These coordinates were used to calculate a pair-correlation function (PCF) for each image following the method in Section 3.3.

6.4.2 Model Validation Using Experimental Data

We will now use *in vitro* experimental data to validate our model. We begin by exploring whether the directional bias mechanism is capable of generating spatial structure that is qualitatively similar to that observed in 3T3 fibroblast cell populations studied *in vitro* and aim to estimate parameters which yield a reasonable qualitative match to our data.

6.4.2.1 Parameter Estimation

Movement rates for 3T3 fibroblast cells are discussed in the literature [12, 23]. We choose a biologically relevant rate of 50 μm/hour for the speed at which an isolated cell moves (i.e. in the absence of neighbourhood interactions). Cell speed is not itself a parameter of our model, but can be decomposed into two constituent parts for input into the model: a mean step length $1/\lambda_\mu = 10$ μm and an intrinsic movement rate $m = 5$ hour^{-1}. For the movement PDF $\mu(\mathbf{x}, \mathbf{y})$ we set $\sigma_\mu = 2.5$ μm which is biologically reasonable as it ensures cells are more likely to take short steps than undergo large jumps across the space. We employ the directional bias mechanism to incorporate volume exclusion effects by interpreting $2\sigma_v$ as the approximate range over which a cell interacts with neighbours and treating this as a proxy for the average diameter of a cell. From the literature, the average cell diameter for 3T3 fibroblast cells is approximately 20 μm which yields $\sigma_v = 10$ μm [12, 99]. Here, we consider the directional bias mechanism in the absence of neighbour-dependent motility (i.e. we set $\alpha = 0$). With these parameter choices in place, interaction strength β is the only parameter that we need to estimate.

Images are taken at the centre of the well to avoid edge effects and when analysing our *in*

vitro data, we assume that cells are distributed homogeneously across this region. An average cell density is estimated from each image, by dividing the number of cells in an image (which ranged between 80 and 318 cells) by the image area. Previously, we implemented periodic boundary conditions in our IBM simulations such that cells located near a boundary of the domain could interact with those at an opposite boundary. Therefore it was reasonable to calculate a periodic PCF from the configurations of cells that arose. However, for our experimental data, the motility of a cell located near the edge of an image will not be affected by a cell at an opposite edge. Therefore, to calculate an accurate average pair density for the short displacements we are primarily interested in, we choose to generate a non-periodic PCF $C_{exp}(r)$ from the experimental images.

To obtain an estimate for β we consider a single experimental image of dimensions 4500 μm x 450 μm with 286 cells, as shown in Figure 6.7(a) with markers superimposed over cell locations. We use our IBM to simulate movement in this 4500 μm x 450 μm region using the parameters discussed above and explore different values of β. In each simulation repeat, 286 cells are initially distributed according to a spatial Poisson process and we compute a PCF once the system has converged to steady state. Figure 6.7(b) shows a snapshot from an IBM simulation at $t = 15$ hours. The presence of spatial structure is not obvious from visual inspection of Figures 6.7(a)-(b) alone but calculating a PCF (Fig. 6.7(c)) indicates a regular spatial pattern over displacements < 50 μm. We find that for $\beta = 1000$ μm the PCFs predicted by our IBM and spatial moment model provide a very good visual match to that computed from the *in vitro* data for this sample. Unlike $C_{IBM}(r)$ and $C_{SM}(r)$, the PCF computed from each experimental image does not tend to 1 for large displacements because it is computed from non-periodic distances and owing to the image dimensions. However, we see good agreement at short to moderate displacements. To validate our estimate, we compare PCFs obtained using the same parameter choices and $\beta = 1000$ μm for the average cell densities in each of the other images (Figures given in the Appendix). For all samples we see a reasonable qualitative agreement between the PCFs predicted by the model and the PCF generated from the *in vitro* data.

The PCFs $C_{exp}(r)$ and $C_{IBM}(r)$ employ a bin width δr which provides a reasonably smooth function for the majority of experimental samples yet contains sufficient information about spatial structure to allow us to carry out our analysis. Smaller values of δr give a better match to $C_{SM}(r)$, however $C_{exp}(r)$ becomes dominated by fluctuations.

6.4.2.2 Average Cell Density and Spatial Structure

From our numerical results we know that both the mechanisms for neighbour-dependent motility and directional bias are capable of generating spatial structure. In the absence of directional bias, large values of α are required to generate the extent of spatial structure observed in the *in vitro* data. When carrying out IBM simulations under these conditions, individuals experience strong neighbourhood interactions and, as a result, movement rates ψ_i are often considerably higher than the average movement rates of fibroblast cells discussed in the literature [12,23]. For example, using the same parameter choices as for Figure 6.7 but in the absence of directional bias ($\beta = 0$), an interaction strength of $\alpha = 1000$ generates spatial structure which is a reasonable qualitative match to the *in vitro* data. However, 23% of individuals undergo movement with a rate $\psi_i > 100$ hour^{-1}, which corresponds to a biologically unreasonable cell speed of 1000μm/hour. Therefore, we do not consider neighbour-dependent motility in isolation. When both mechanisms are acting together, numerous combinations of α and β exist that would give rise to similar spatial structure.

Numerical and analytical results suggest that there is a relationship between the average cell density and the extent of spatial structure in the moving cell population. Increasing the average cell density causes a decrease in the extent of spatial structure, i.e. for a regular spatial pattern average pair densities at short displacements increase towards 1. However, for the average cell densities studied here, it is not immediately obvious whether our *in vitro* experimental data supports the suggestion that a significant relationship exists. We now explore this idea in more depth by using the area between the PCF to calculate a summary statistic which quantifies the extent of spatial structure. We consider two metrics, the first of which measures spatial structure as $\int_0^R (1 - C(r))\mathrm{d}r$ (Fig. 6.8(c)). For this metric, positive values indicate a regular spatial pattern while negative values indicate a cluster spatial pattern. The second is given by $\int_0^R |1 - C(r)|\mathrm{d}r$ (Fig. 6.8(d)). Both metrics are calculated for $R = 80$ μm and have units μm. We compute each metric for PCFs generated from the IBM, spatial moment model and *in vitro* data, as shown in Figure 6.8. The average cell densities obtained from the *in vitro* data lie within a relatively small range and so the overall change in the metric is small. Nevertheless, for both metrics our model predicts that increasing average cell density decreases the extent of spatial structure. To investigate whether our *in vitro* data supports this we carry out a simple linear regression, yielding

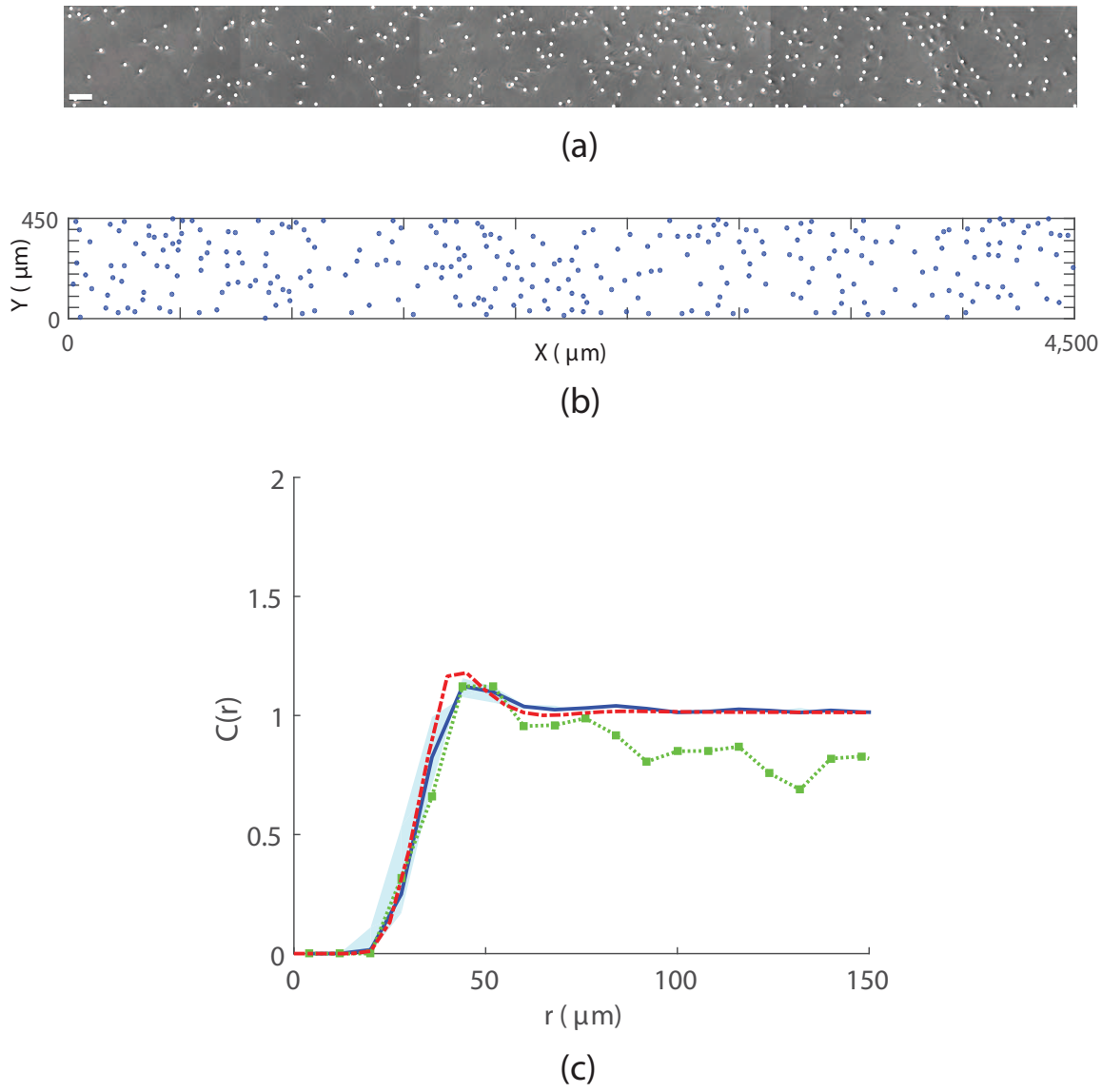


Figure 6.7: Spatial structure in 3T3 fibroblast cells for 286 cells in a 4500μm x 450μm region. (a) Sample image (obtained from a well containing cell suspension of approximate initial density 30,000 cells/ml) showing superimposed markers (white dots). Scale bar corresponds to 100μm; (b) Cell locations (blue dots) at $t = 15$ hours from a single IBM simulation. Parameters are $\alpha = 0$ hours^{-1}, $\beta = 1000$ μm, $\sigma_w = \sigma_v = 10$ μm, $m = 5$ hours^{-1}, $\lambda_\mu = 0.1$ μm^{-1}, $\sigma_\mu = 2.5$ μm; (c) PCF $C_{IBM}(r)$ (blue solid line) obtained from averaging results from 200 simulations of the IBM at $t = 15$ hours. PCFs computed from the IBM using values of β within the range $\pm 75\%$ of $\beta = 1000$ μm, lie within the region indicated by the blue shaded area. PCF $C_{exp}(r)$ (green squares-dotted line) generated from experimental image, for $\delta r = 8$ μm. PCF $C_{SM}(r)$ (red dashed line) approximated by spatial moment model at $t = 15$ hours, for $\Delta = 5$ μm and $\xi_{max} = 150$ μm.

p-values of 0.0211 and 0.0435 for the first (Fig. 6.8(a)) and second metric (Fig. 6.8(b)), respectively. Thus, using either metric and despite the noise in our *in vitro* data, the results suggest that a significant relationship does indeed exist between average cell density and spatial structure.

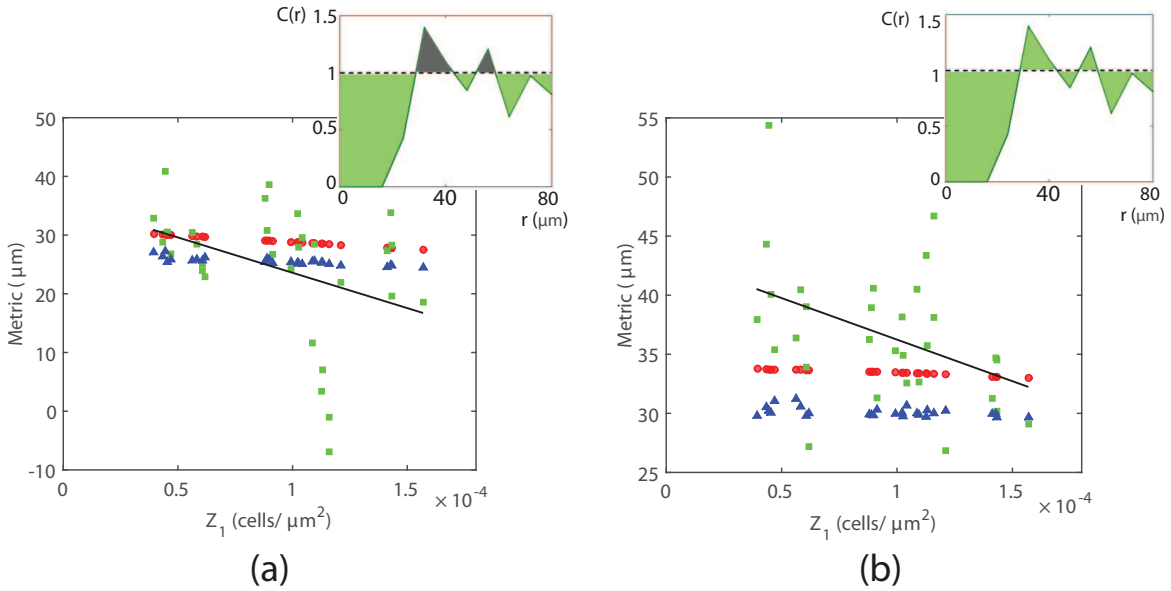


Figure 6.8: Relationship between average cell density and the extent of spatial structure. (a)-(b) Metrics calculated from IBM (blue triangles), spatial moment model (red circles) and *in vitro* data (green squares) for the average cell densities in each of the images. A regression line (black line) is fitted to the experimental data. (c) The metric in (a) is calculated by integrating $(1 - C(r))$ over displacements $0 \leq r \leq 80$ μm, i.e. summing the green-shaded area and subtracting the grey-shaded area. (d) The metric in (b) is calculated by integrating $|1 - C(r)|$ over displacements $0 \leq r \leq 80$ μm, i.e. summing the green-shaded area.

6.5 Discussion

IBMs of collective movement allow us to explore how interactions between individuals give rise to spatial structure and how, in turn, this self-generated spatial structure affects the population dynamics. However, IBMs are limited when it comes to explaining population-level behaviour as they can be difficult to analyse mathematically. To move beyond these limitations, population-level models can be derived from IBMs but often employ a mean-field assumption which neglects spatial correlations between cells. We have derived a population-level description in terms of spatial moment dynamics to account for spatial correlations and give insight into how neighbour-dependent directional bias generates spatial structure in a moving cell population. Extending our original model [14] from one to two spatial dimensions makes it more amenable for use alongside experimental data. Our results verify that the spatial moment model can provide a good approximation to averaged simulations of the underlying IBM when cells are distributed homogeneously through space.

Volume exclusion effects can be incorporated into lattice-free models of interacting agents, for example using a hard sphere approach where neighbours are explicitly excluded from a re-

gion surrounding an individual. Instead, we employ the mechanism for neighbour-dependent directional bias as a means of accounting for crowding effects. Using an interaction kernel concentrated around short pair displacements allows us to reduce the likelihood of two cells being found in very close proximity, although it does not altogether rule out the possibility.

In vitro studies have shown that cell motility can be heavily influenced by the average density of cells, particularly at high densities where crowding effects come into play, affecting the movement rate or direction of individuals [1, 12, 100]. In addition, spatial correlations between cells can have major implications for motility, for example cell populations with clustering exhibit different behaviour to those that adopt regular spatial patterns [35, 39]. We carried out *in vitro* experiments with motile 3T3 fibroblast cells for model validation and to explore the extent to which spatial structure is generated in fibroblast cell populations. It is not obvious from visual inspection alone whether spatial structure is present, however calculating PCFs from the imaging data indicates a regular spatial pattern. The spatial structure arises over displacements < 50 μm and is likely predominantly a consequence of space being excluded by the cells, however chemotactic interactions, such as chemokine signalling, may also contribute to a lesser extent [12]. We obtain the majority of our model parameters by selecting biologically relevant values from the literature and use our data to provide an estimate for the interaction strength β. This parameter was estimated from a single experimental image and for validation we use the same estimate for the average cell densities in each of the other images. A visual comparison of the PCFs suggests that our parameterised model can successfully predict the spatial structure of 3T3 fibroblasts at various average cell densities.

We choose to calculate a non-periodic PCF from each experimental image to obtain an accurate average pair density at short displacements. Because we do not apply edge corrections and owing to the image dimensions, the PCF often has values less than 1 for large displacements. However, we would expect that a PCF calculated either for a very large number of cells (at the same average density) or by averaging results from many identically-prepared repeated experiments, would give $C(r) \approx 1$ for large displacements. A number of methods to account for edge effects are discussed in the literature, for example the use of buffer zones, toroidal edge corrections or employing weighting factors [87, 101]. However, in some cases, applying an edge correction may yield results that do not provide an accurate representation of the spatial structure in the population. For instance, when analysing spatial

patterns that are clustered or regular, the use of a toroidal correction can lead to an unknown extent of bias in the resulting distribution of distances [101]. To avoid this uncertainty, we have chosen to work with the actual pair distances between cells in the experimental images and not correct for edge effects.

We have further validated our model by considering in more detail the relationship between average cell density and the extent of spatial structure in a cell population. Numerical and analytical results from our model suggest that increasing the average cell density decreases the extent of spatial structure. There is considerable noise in the *in vitro* data because we choose to analyse PCFs generated from individual images as opposed to working with averaged results. In addition, the data considers a relatively small range of average cell densities. Nevertheless, our experimental data also supports the idea that such a relationship exists. The most likely explanation for this effect is that as average cell density increases, there is less free space available and cells are forced into closer proximity. Because of their deformable plasma membranes, pairs of cells can arise at displacements less than the average diameter of a cell. This increases the average pair density at short displacements, thus reducing the extent of spatial structure. Because we do not employ a hard sphere volume-exclusion method, instead representing cells by points in space, our model will predict a Poisson spatial pattern for very high average cell densities (far greater than those in our data). In reality, the fact that 3T3 fibroblasts have a minimum area they can occupy means that this would never be observed *in vitro*.

The spatial moment model is only an approximation to the IBM because it invokes a closure assumption which closes the dynamical system at second order and ignores higher order moments. The performance of our model depends on the suitability of this closure as an approximation to the third moment. Different closures are proposed in the literature and we use the power-3 closure (Kirkwood Superposition Approximation), which is a relatively simple closure that is often applied in cell movement models. This closure is known to perform reasonably well for regular and Poisson spatial patterns but causes the model to underestimate the second moment for cluster patterns. A number of other closures also share this limitation. The asymmetric power-2 closure, which expresses the third moment in terms of weighted sums of lower order moments, can prove more successful for cluster spatial patterns. However it is not always obvious which weighting constants are most appropriate and the closure has the potential to predict negative average densities of triplets [79, 85, 96].

We have chosen to use kernels suitable for modelling fibroblast movement but different kernels could be employed for applications in other contexts. However, there is a numerical constraint associated with choosing the movement PDF $\mu(\mathbf{x}, \mathbf{y})$. If using a PDF that has large positive values concentrated at pair displacements very close to zero, the spatial moment model cannot always accurately capture the full extent of the directional bias at these short displacements. This, in turn, causes the model to underestimate the extent of spatial structure. Choosing a movement PDF with positive values at displacements further from zero, such as the PDF employed here, overcomes this issue. Expressing and solving the moment dynamics equations in polar coordinates may also allow for greater flexibility in the choice of movement PDF.

Chapter 7

Model of a Birth-death-movement Process

In this chapter we explore how the collective dynamics are affected when individuals are able to give birth or die as well as move. In the context of cell biology, birth refers to cell proliferation, i.e. the division of a cell to form two daughter cells. *In vitro* studies of cell proliferation have shown that the rate at which cells divide is dependent on local cell density. For instance, at high local cell densities, crowding effects such as contact inhibition of proliferation can reduce the rate at which proliferation occurs [44, 102–104]. To consider what implications this has for the emergent behaviour of a cell population, we extend our two-dimensional IBM to a birth-death-movement process and derive corresponding descriptions for birth and death in the spatial moment model. We allow the rate at which an individual gives birth to depend on the density of cells in its neighbourhood and explore how this affects the growth of the population, both in the presence and absence of directional bias. Similar spatial moment models that incorporate neighbourhood effects into birth or death processes are discussed in the literature (see for example, [80], [82], [65]), with some also accounting for movement [85], [86]. However, moment approximations for birth-death processes incorporating neighbour-dependent movement, in particular directional bias, have received considerably less attention.

7.1 Two-dimensional IBM for Cell Birth, Death and Movement

We now extend our two-dimensional IBM for collective movement to incorporate cell birth and death. With birth, death and movement the state of the system at time t comprises the location $\mathbf{x}_i \in \Omega$ of each individual i in a population of size $N(t)$. In the original model, individual i moved as a Poisson process over time with movement rate per unit time ψ_i. With each movement event the system underwent a transition to a new state with rate $\psi_i \; \mu(\mathbf{x}_i, \mathbf{x}_i + \xi'')$ for a displacement ξ''. We now consider a second type of possible transition

in the state of the system which corresponds to individual i giving birth to a new cell. This type of transition adds an additional coordinate $\mathbf{x}'$ to the current system state $\mathbf{x}(t)$, where $\mathbf{x}' \in \Omega$ represents the location of cell i's offspring. The system undergoes this transition with rate $\hat{B}_i(\mathbf{x})\mu^{(\rho)}(\mathbf{x}_i, \mathbf{x}')$, where $\hat{B}_i(\mathbf{x})$ is the birth rate per unit time of cell i when the system is in state $\mathbf{x}$. The function $\mu^{(\rho)}(\mathbf{x}_i, \mathbf{x}')$ is a PDF for dispersal of cell i's offspring to a location $\mathbf{x}'$. Finally, we allow for the death of cell i by considering a third type of transition in system state which involves the removal of the i^{th} coordinate from $\mathbf{x}(t)$. This transition occurs with rate $\hat{D}_i(\mathbf{x})$ which is the death rate per unit time of individual i.

As before, we employ the Gillespie algorithm to simulate this birth-death-movement process. In this case, the time increment τ between each event is exponentially distributed with mean $1/\lambda(t)$, where

$$\lambda(t) = \sum_{i=1}^{N(t)} \left(\psi_i(\mathbf{x}) + \hat{B}_i(\mathbf{x}) + \hat{D}_i(\mathbf{x}) \right), \tag{7.1}$$

at time t. For brevity, from here on we omit the $\mathbf{x}$ argument from the movement, birth and death rate notation. One of $3N(t)$ possible transitions is chosen to occur at time $t + \tau$, with a probability that is proportional to the rate of that transition. If the transition corresponds to a movement event, a movement displacement ξ'' is determined according to $\mu(\mathbf{x}_i, \mathbf{x}_i + \xi'')$. If a birth event is chosen then the location $\mathbf{x}'$ of the daughter cell is drawn from $\mu^{(\rho)}(\mathbf{x}_i, \mathbf{x}')$. With this general framework in place, we now make particular choices for the rate functions and PDFs in order to simulate cell movement. However it is important to keep in mind that the model is amenable to different choices of function for application in other contexts.

The movement rate ψ_i and movement PDF $\mu(\mathbf{x}, \mathbf{y})$ are the same as those given previously in equations (3.8) and (6.6) respectively. We define a death rate $\hat{D}_i$ that is comprised solely of an intrinsic death rate d and does not depend on other cells in the neighbourhood of individual i:

$$\hat{D}_i = d. \tag{7.2}$$

We consider a birth rate $\hat{B}_i$ comprising an intrinsic component and a component that depends on the presence of other cells in the neighbourhood. The intrinsic birth rate ρ describes the rate at which a cell would give birth in the absence of other cells. A neighbouring cell, displaced from $\mathbf{x}_i$ by $\mathbf{z}$, contributes $w_\rho(\mathbf{z})$ to the birth rate of cell i. Summing the intrinsic rate and contributions from $N(t)$ neighbours at $\mathbf{x}_j$ gives the overall birth rate for an individual

at $\mathbf{x}_i$:

$$\hat{B}_i = \max(0,\ \rho + \sum_{\substack{j=1 \\ i \neq j}}^{N(t)} w_\rho(\mathbf{x}_j - \mathbf{x}_i)), \tag{7.3}$$

where we have implemented the maximum formula to ensure $\hat{B}_i \geq 0$.

The function $w_\rho(\mathbf{z})$ weights the strength of interaction between a pair of cells displaced by $\mathbf{z}$ and we choose $w_\rho(\mathbf{z})$ to be a two-dimensional Gaussian function

$$w_\rho(\mathbf{z}) = \gamma \exp\left(-\frac{|\mathbf{z}|^2}{2\sigma_\rho^2}\right). \tag{7.4}$$

Here, the strength of interaction and spatial range of interaction are determined by γ and σ_ρ respectively. The dimensions of ρ, d, $w_\rho(\mathbf{z})$, $\hat{B}_i$ and $\hat{D}_i$ are T^{-1}. Setting $\gamma < 0$ means that cell i's birth rate $\hat{B}_i$ is reduced by interactions with nearby neighbours. This allows us to consider how crowding effects, such as contact inhibition of proliferation, affect the collective cell dynamics.

When a cell at $\mathbf{x}$ gives birth, its offspring is dispersed to a new location $\mathbf{y}$ with a probability drawn from the PDF $\mu^{(\rho)}(\mathbf{x}, \mathbf{y})$ for dispersal by a displacement $\mathbf{y} - \mathbf{x}$. The dispersal PDF $\mu^{(\rho)}(\mathbf{x}, \mathbf{y})$ is a bivariate normal distribution

$$\mu^{(\rho)}(\mathbf{x}, \mathbf{y}) = \frac{1}{2\pi\sigma_{\mu_\rho}^2} \exp\left(\frac{-|\mathbf{y} - \mathbf{x}|^2}{2\sigma_{\mu_\rho}^2}\right), \tag{7.5}$$

where σ_{μ_ρ} determines the spatial range over which offspring are dispersed from a parent cell at $\mathbf{x}$. Therefore, even if the direction of cell movement is biased, the dispersal of offspring is unbiased (i.e. the direction in which offspring are dispersed is not directly affected by neighbourhood interactions). The function $\mu^{(\rho)}(\mathbf{x}, \mathbf{y})$ satisfies $\int \mu^{(\rho)}(\mathbf{x}, \mathbf{y}) \mathrm{d}\mathbf{y} = 1$ and has dimensions L^{-2}. Periodic conditions are implemented at the boundaries of the spatial domain.

7.2 Spatial Moment Dynamics with Birth, Death and Movement

Extending the IBM to a birth-death-movement process provides insight into how cell proliferation and death give rise to spatial structure in a moving population and how this, in turn, affects the collective dynamics. To gain a deeper understanding of how this individual-level behaviour relates to the emergent population dynamics, we now scale up our extended IBM

by deriving an approximation in terms of the dynamics of spatial moments up to second order.

We make use of the definitions for spatial moments given in Chapter 4 by equations (4.1)-(4.5). As for our original model, we derive a description of the moment dynamics that can be applied for a non-homogeneous spatial pattern. However, in our numerical results we only consider a spatially homogeneous case.

7.2.1 First Spatial Moment

We first describe rate functions and dynamics for the first spatial moment. The expected movement rate $M_1(\mathbf{x})$ for a cell at $\mathbf{x}$ was given previously in (4.6). In a similar manner, we use the birth rate $\hat{B}_i$ in (7.3) to derive an expected birth rate function $B_1(\mathbf{x})$ for a cell at $\mathbf{x}$ in terms of spatial moments. Accounting for contributions from neighbours involves an integration over $\mathbf{y}$ of the probability of a cell being located at $\mathbf{y}$ conditional on the presence of a cell at $\mathbf{x}$, weighted by $w_\rho(\mathbf{y}-\mathbf{x})$. Making use of the conditional probability in (4.4) we obtain the expected birth rate $B_1(\mathbf{x})$ for a cell at $\mathbf{x}$:

$$B_1(\mathbf{x}) = \rho + \int w_\rho(\mathbf{y}-\mathbf{x})\frac{Z_2(\mathbf{x},\mathbf{y})}{Z_1(\mathbf{x})}\mathrm{d}\mathbf{y}. \tag{7.6}$$

The maximum formula in (7.3) is not incorporated in the spatial moment description because we only consider solutions in which negative expected birth rates do not arise. The birth rate $B_1(x)$ and the equivalent birth rate $B_2(x,y)$ for the second moment dynamics (given below in (7.11)) are functions of spatial moments and are therefore also dependent on time, however we omit the argument t for notational ease.

The equivalent expression for death rate $\hat{D}_i$ is the expected death rate for a cell at x given by

$$D = d. \tag{7.7}$$

The PDF for dispersal of offspring $\mu^{(\rho)}(\mathbf{x},\mathbf{y})$, required for the first moment dynamics, is given by (7.5). Unlike the movement PDF $\mu_1(\mathbf{x},\mathbf{y})$, which is a function of spatial moments (contained within the bias function $b_1(x)$ in (6.8)), the dispersal PDF $\mu^{(\rho)}(\mathbf{x},\mathbf{y})$ is independent of time. In a spatially homogeneous setting, $M_1(\mathbf{x})$ and $B_1(\mathbf{x})$ are independent of x, and the PDFs $\mu_1(\mathbf{x},\mathbf{y})$ and $\mu^{(\rho)}(\mathbf{x},\mathbf{y})$ can be expressed in terms of pair displacements.

7.2.2 Dynamics of the First Spatial Moment

We now extend the description of the first moment dynamics developed in Section 4.3 to incorporate cell birth and death. As before, we consider the probability that a cell is present in the small region δx centred on $\mathbf{x}$ at a time $t + \delta t$, given in (4.9). Movement, birth and death events occur as independent Poisson processes over time so the probability of more than one event taking place in a short period of time δt is $O(\delta t^2)$.

A cell can depart δx by either moving or dying, thus the probability that a cell remains in δx in $[t, t + \delta t]$ is

$$P(\text{cell remains in } \delta x \text{ in } [t, t + \delta t]) = 1 - (M_1(\mathbf{x}) + D)\,\delta t + O(\delta t^2). \tag{7.8}$$

A cell arrives in δx via either a movement event or a birth event accompanied by dispersal. Therefore, we write the probability that a cell arrives in δx in $[t, t + \delta t]$ as the sum of two terms. The first component is the probability that a cell moves from $\mathbf{u}$ into δx, integrated over possible starting locations $\mathbf{u}$. The second is the probability that a parent cell at $\mathbf{u}$ gave birth and dispersed a daughter cell into δx, integrated over all possible parent locations $\mathbf{u}$. This gives

$$P(\text{cell arrives in } \delta x \text{ in } [t, t + \delta t]) = h\delta t \int \left(\mu_1(\mathbf{u}, \mathbf{x})M_1(\mathbf{u}) + \mu^{(\rho)}(\mathbf{u}, \mathbf{x})B_1(\mathbf{u})\right) Z_1(\mathbf{u}, t)\mathrm{d}\mathbf{u} + O(\delta t^2), \tag{7.9}$$

where $M_1(\mathbf{u})Z_1(\mathbf{u}, t)$ and $B_1(\mathbf{u})Z_1(\mathbf{u}, t)$ are the movement rate and birth rate per unit area at location $\mathbf{u}$, respectively. Substituting (4.1),(7.8) and (7.9) into (4.9), then using the Taylor expansion of $Z_1(\mathbf{x}, t + \delta t)$ and taking the limit $h, \delta t \to 0$, we get the rate of change for the first spatial moment:

$$\begin{aligned} \frac{\mathrm{d}Z_1(\mathbf{x}, t)}{\mathrm{d}t} &= -(M_1(\mathbf{x}) + D)Z_1(\mathbf{x}, t) \\ &\quad + \int \left(\mu_1(\mathbf{u}, \mathbf{x})M_1(\mathbf{u}) + \mu^{(\rho)}(\mathbf{u}, \mathbf{x})B_1(\mathbf{u})\right) Z_1(\mathbf{u}, t)\mathrm{d}\mathbf{u}. \end{aligned} \tag{7.10}$$

This equation is of a similar form to the dynamics of the first spatial moment for cell movement without birth and death, given in equation (4.13). As before, the negative term describes movement out of $\mathbf{x}$ but here we also account for departure from $\mathbf{x}$ via cell death. The integral term accounts for arrival at $\mathbf{x}$ which previously occurred only by movement from a starting

location $\mathbf{u}$ but now also arises by dispersal of a daughter cell from a parent at $\mathbf{u}$.

7.2.3 Second Spatial Moment

For the dynamics of the second spatial moment, an expected movement rate function $M_2(\mathbf{x}, \mathbf{y})$ for a cell at $\mathbf{x}$ conditional on the presence of a cell at $\mathbf{y}$, was given previously in (4.14). The expected birth rate $B_2(\mathbf{x}, \mathbf{y})$ for a cell at $\mathbf{x}$ conditional on the presence of a cell at $\mathbf{y}$ is derived in a similar manner to give

$$B_2(\mathbf{x}, \mathbf{y}) = \rho + \int w_\rho(\mathbf{z} - \mathbf{x}) \frac{Z_3(\mathbf{x}, \mathbf{y}, \mathbf{z})}{Z_2(\mathbf{x}, \mathbf{y})} \mathrm{d}\mathbf{z} + w_\rho(\mathbf{y} - \mathbf{x}). \tag{7.11}$$

The PDF for dispersal $\mu^{(\rho)}(\mathbf{x}, \mathbf{y})$ in the second moment dynamics is still given by (7.5). When cells are distributed homogeneously throughout space we can express $M_2(\mathbf{x}, \mathbf{y})$, $B_2(\mathbf{x}, \mathbf{y})$, $\mu_2(\mathbf{x}, \mathbf{y}, \mathbf{z})$ and $\mu^{(\rho)}(\mathbf{x}, \mathbf{y})$ in terms of displacements between the physical locations $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$.

7.2.4 Dynamics of the Second Spatial Moment

To incorporate cell birth and death into the second moment dynamics derived in Section 4.5, we once again consider the probability of cells being present in the regions δx and δy at time $t + \delta t$, given in (4.17). Now that a cell can depart a region via either a movement event or death event, the probability of both cells remaining in δx and δy in $[t, t + \delta t]$ is

$$P\left(\begin{array}{c}\text{both cells remain} \\ \text{in } \delta x \text{ and } \delta y\end{array}\right) = 1 - (M_2(\mathbf{x}, \mathbf{y}) + M_2(\mathbf{y}, \mathbf{x}) + 2D)\,\delta t + O(\delta t^2). \tag{7.12}$$

The probability that a cell remains in δy and a cell arrives in δx in $[t, t + \delta t]$ is equivalent to the conditional probability that a cell arrives in δx given there is a cell in δy. A cell can arrive in each region via a movement event or a birth event accompanied by dispersal. As in (7.9), we consider the probability that a cell moves from $\mathbf{u}$ into $\mathbf{x}$ plus the probability that a cell at $\mathbf{u}$ gives birth to a daughter at $\mathbf{x}$, then integrate over all possible locations of origin $\mathbf{u}$. In this case, the probability of a cell being located at $\mathbf{u}$ is conditional on the presence of

a cell at $\mathbf{y}$. Together, this gives

$$P\left(\begin{array}{c}\text{cell remains in } \delta y \\ \text{\& cell arrives in } \delta x\end{array}\right) = h\delta t\left[\int\left(\mu_2(\mathbf{u},\mathbf{x},\mathbf{y})M_2(\mathbf{u},\mathbf{y}) + \mu^{(\rho)}(\mathbf{u},\mathbf{x})B_2(\mathbf{u},\mathbf{y})\right)\frac{Z_2(\mathbf{u},\mathbf{y},t)}{Z_1(\mathbf{y},t)}\mathrm{d}\mathbf{u} + \mu^{(\rho)}(\mathbf{y},\mathbf{x})B_1(\mathbf{y})\right] + O(\delta t^2). \tag{7.13}$$

Here, an extra term has been included to account for the possibility that the cell at $\mathbf{y}$ gives birth and disperses a daughter cell to $\mathbf{x}$.

We substitute equations (7.12)-(7.13) along with (4.18)-(4.19) into (4.17) and employ the Taylor expansion of $Z_2(\mathbf{x},\mathbf{y},t+\delta t)$. By expanding terms, making use of the two-fold symmetry of $Z_2(\mathbf{x},\mathbf{y},t)$, then taking the limit $h,\delta t \to 0$, we get the rate of change for the second spatial moment:

$$\begin{aligned}\frac{\mathrm{d}Z_2(\mathbf{x},\mathbf{y},t)}{\mathrm{d}t} &= -(M_2(\mathbf{x},\mathbf{y}) + M_2(\mathbf{y},\mathbf{x}) + 2D)Z_2(\mathbf{x},\mathbf{y},t) \\ &\quad + \int\left(\mu_2(\mathbf{u},\mathbf{x},\mathbf{y})M_2(\mathbf{u},\mathbf{y}) + \mu^{(\rho)}(\mathbf{u},\mathbf{x})B_2(\mathbf{u},\mathbf{y})\right)Z_2(\mathbf{u},\mathbf{y},t)\mathrm{d}\mathbf{u} \\ &\quad + \int\left(\mu_2(\mathbf{u},\mathbf{y},\mathbf{x})M_2(\mathbf{u},\mathbf{x}) + \mu^{(\rho)}(\mathbf{u},\mathbf{y})B_2(\mathbf{u},\mathbf{x})\right)Z_2(\mathbf{u},\mathbf{x},t)\mathrm{d}\mathbf{u} \\ &\quad + \mu^{(\rho)}(\mathbf{y},\mathbf{x})B_1(\mathbf{y})Z_1(\mathbf{y}) + \mu^{(\rho)}(\mathbf{x},\mathbf{y})B_1(\mathbf{x})Z_1(\mathbf{x}).\end{aligned} \tag{7.14}$$

Equation (7.14) has a similar form to the second moment dynamics given by (4.23) for movement only. Here, the negative component accounts for departure from $\mathbf{x}$ by movement or death, conditional on the presence of a cell at $\mathbf{y}$, and symmetric terms for departure from $\mathbf{y}$. The first integral term represents arrival at $\mathbf{x}$ via either movement from a starting location $\mathbf{u}$ or dispersal of a daughter cell from a parent at $\mathbf{u}$, conditional on the presence of a cell at $\mathbf{y}$. The second integral contains symmetric terms corresponding to departure from $\mathbf{y}$. The final two terms, which do not appear in (4.23), describe dispersal of a daughter cell to $\mathbf{x}$ from a parent at $\mathbf{y}$, and vice-versa. Equation (7.14) is the same as that derived by Plank and Law (2015) [61], with the exception that the movement kernel $\mu_2(\mathbf{x},\mathbf{y},\mathbf{z})$ incorporates neighbour-dependent directional bias into the dynamics, providing insights into how directional bias affects spatial structure in a cell population undergoing birth, death and movement.

7.2.5 Spatially Homogeneous Moment Dynamics

In a spatially homogeneous setting, the dynamics of spatial moments can be expressed in terms of displacements between physical locations, as described in Section 4.6 for the original model without birth and death. Here, we use the same notation as in Section 4.6 to describe the dynamics of the first and second moments for a spatially homogeneous distribution of cells.

The first moment dynamics are now independent of $\mathbf{x}$ and, using the constraint $\int \mu(\mathbf{u}, \mathbf{x})\mathrm{d}\mathbf{x} = 1$ and $\int \mu^{(\rho)}(\mathbf{u}, \mathbf{x})\mathrm{d}\mathbf{x} = 1$, equation (7.10) simplifies to:

$$\frac{\mathrm{d}Z_1}{\mathrm{d}t} = (B_1 + D)Z_1, \tag{7.15}$$

where

$$B_1 = \rho + \int w_\rho(\boldsymbol{\xi})\frac{Z_2(\boldsymbol{\xi})}{Z_1}\mathrm{d}\boldsymbol{\xi}. \tag{7.16}$$

Therefore, in the spatially homogeneous case the dynamics of the first moment are not directly affected by movement (because the movement terms have been cancelled out in (7.15)). However, movement can influence the dynamics of the second moment, thus the effects of movement can channel to the first moment dynamics indirectly through the $Z_2(\boldsymbol{\xi})$ term in (7.16).

The dynamics of the second moment are given by

$$\begin{aligned}\frac{\mathrm{d}Z_2(\boldsymbol{\xi})}{\mathrm{d}t} = & -(M_2(\boldsymbol{\xi}) + M_2(-\boldsymbol{\xi}) + 2D)Z_2(\boldsymbol{\xi}) \\ & + \int \left(\mu_2(\boldsymbol{\xi}'', \boldsymbol{\xi}'' + \boldsymbol{\xi})M_2(\boldsymbol{\xi}'' + \boldsymbol{\xi}) + \mu^{(\rho)}(\boldsymbol{\xi}'')B_2(\boldsymbol{\xi}'' + \boldsymbol{\xi})\right) Z_2(\boldsymbol{\xi}'' + \boldsymbol{\xi})\mathrm{d}\boldsymbol{\xi}'' \\ & + \int \left(\mu_2(\boldsymbol{\xi}'', \boldsymbol{\xi}'' - \boldsymbol{\xi})M_2(\boldsymbol{\xi}'' - \boldsymbol{\xi}) + \mu^{(\rho)}(\boldsymbol{\xi}'')B_2(\boldsymbol{\xi}'' - \boldsymbol{\xi})\right) Z_2(\boldsymbol{\xi}'' - \boldsymbol{\xi})\mathrm{d}\boldsymbol{\xi}'' \\ & + \left(\mu^{(\rho)}(-\boldsymbol{\xi}) + \mu^{(\rho)}(\boldsymbol{\xi})\right) B_1 Z_1.\end{aligned} \tag{7.17}$$

The movement rate $M_2(\boldsymbol{\xi})$, movement PDF $\mu_2(\boldsymbol{\xi}, \boldsymbol{\xi}')$ and neighbour-dependent bias $\mathbf{b}_2(\boldsymbol{\xi})$ were given previously in (4.26), (6.13) and (6.14) respectively. The birth rate $B_2(\boldsymbol{\xi})$ for a cell, in a pair with a neighbour displaced from it by $\boldsymbol{\xi}$, is

$$B_2(\boldsymbol{\xi}) = \rho + \int w_\rho(\boldsymbol{\xi}')\frac{Z_3(\boldsymbol{\xi}, \boldsymbol{\xi}')}{Z_2(\boldsymbol{\xi})}\mathrm{d}\boldsymbol{\xi}' + w_\rho(\boldsymbol{\xi}). \tag{7.18}$$

The PDF $\mu_\rho(\boldsymbol{\xi})$ for dispersal of offspring a displacement $\boldsymbol{\xi}$ away from the parent cell is

$$\mu^{(\rho)}(\boldsymbol{\xi}) = \frac{1}{2\pi\sigma^2_{\mu_\rho}} \exp\left(\frac{-|\boldsymbol{\xi}|^2}{2\sigma^2_{\mu_\rho}}\right). \tag{7.19}$$

The boundary condition is as follows:

$$Z_2(\boldsymbol{\xi}) \to Z_1^2 \text{ as } |\boldsymbol{\xi}| \to \infty. \tag{7.20}$$

7.3 Numerical Results

In this section we consider to what extent spatial structure affects the collective dynamics of a cell population undergoing birth, death and movement. We are also interested in how neighbourhood interactions affecting birth and movement give rise to spatial structure and the relative importance of each structure-generating mechanism. To approach these problems, we simulate our IBM for cell birth, death and movement and consider the growth of the population by computing average cell density $N(t)/(L_xL_y)$ over time t. We are primarily interested in the role of crowding effects, such as contact inhibition of proliferation, in population growth. Therefore, we only consider values of $\gamma < 0$ such that a cell's birth rate is reduced by interactions with neighbours and the growing population eventually reaches a steady state average cell density. In the context of *in vitro* cell studies, steady state corresponds to the average cell density of a population that has reached confluence (i.e. the cell carrying capacity), beyond which further growth is limited by the availability of space and nutrients [1]. In addition, we choose $\gamma < 0$ to be sufficiently small so as not to violate the positivity constraint of the expected birth rate B_1 and $B_2(\boldsymbol{\xi})$.

The extent of spatial structure arising in a realisation of the IBM at a particular time is quantified by computing a pair-correlation function $C_{IBM}(r)$. Averaged results from repeated IBM realisations are compared to numerical solutions of the spatial moment model to explore how well it approximates the underlying stochastic process. As in Chapters 5 and 6, we consider a spatially homogeneous problem by solving the dynamics of the first and second moment described in Section 7.2.5. Again, we compute a PCF $C_{SM}(\boldsymbol{\xi}) = Z_2(\boldsymbol{\xi})/Z_1^2$ but only show a radial section of $C_{SM}(\boldsymbol{\xi})$ which we denote $C_{SM}(r)$, where $r = |\boldsymbol{\xi}|$. Initially, cells are distributed across the domain according to a spatial Poisson process with intensity $N(t)/(L_xL_y)$, corresponding to $Z_2(\boldsymbol{\xi}) = Z_1^2$ in the spatial moment dynamics. The system is

allowed to reach steady state before simulations are stopped, which is determined to have occurred once there is no longer a noticeable difference in the dynamics of the first and second spatial moments over time.

7.3.1 Numerical Methods

Equation (7.22) for the mean-field dynamics of the first spatial moment (derived below) is solved using MATLAB's in-built ode23 solver. We perform a numerical integration of equation (7.17) for the second moment dynamics following the method outlined in Section 6.3.1. The conditions at the boundary of the discretised $\boldsymbol{\xi}$ domain are approximately mean-field such that $Z_2(\boldsymbol{\xi}) \approx Z_1^2$ at the boundary. Therefore, we compute the first spatial moment over time by taking the square root of the value of $Z_2(\boldsymbol{\xi})$ at a corner of the domain, i.e. $Z_1 = \sqrt{Z_2(\xi_{max}, \xi_{max})}$. Because the computation must be constrained to a finite-size discretised spatial grid, the value of $\sqrt{Z_2(\xi_{max}, \xi_{max})}$ is only an approximation to Z_1 and will not correspond exactly to the value of Z_1 obtained by solving equation (7.15) numerically. Therefore, numerical integration of the full coupled system of equations (7.15) and (7.17) would involve small inconsistencies in the value of Z_1. To avoid this discrepancy we solve only equation (7.17) and obtain the solution to the first moment by computing $\sqrt{Z_2(\xi_{max}, \xi_{max})}$ over time.

The dispersal PDF $\mu^{(\rho)}(\boldsymbol{\xi})$ is normalised numerically using the trapezium rule such that $\int \mu^{(\rho)}(\boldsymbol{\xi}) \mathrm{d}\boldsymbol{\xi} = 1$. We carry out the same truncation of kernels as outlined in Section 6.3.1. Similarly, the tails of the interaction kernel $w_\rho(\boldsymbol{\xi})$ and dispersal PDF $\mu^{(\rho)}(\boldsymbol{\xi})$ are truncated such that $w_\rho(\boldsymbol{\xi}) = \mu^{(\rho)}(\boldsymbol{\xi}) = 0$ over the domain $|\boldsymbol{\xi}| > \xi_{max}/2$ and we ensure that $(1/\lambda_\mu + 4\sigma_\mu) \leq 4\sigma_\rho, 4\sigma_{\mu_\rho} \leq \xi_{max}/2$.

7.3.2 Mean-field Population Dynamics

In order to understand the effects of spatial structure on the collective dynamics, it is useful to first consider how a population behaves in the complete absence of spatial structure, i.e. for a Poisson spatial pattern of cells. In terms of spatial moments this corresponds to $Z_2(\boldsymbol{\xi}) = Z_1^2$ (the mean-field assumption). Applying $Z_2(\boldsymbol{\xi}) = Z_1^2$ to equation (7.16) gives the expected birth rate B_1 of a cell for the mean-field case:

$$B_1 = \rho + Z_1 \int w_\rho(\boldsymbol{\xi}) \mathrm{d}\boldsymbol{\xi}. \tag{7.21}$$

Our choice of interaction kernel $w_\rho(\boldsymbol{\xi})$, given in (7.4), has the property that $\int w_\rho(\boldsymbol{\xi})\mathrm{d}\boldsymbol{\xi} = \gamma(2\pi\sigma_\rho^2)$. Using this property and substituting (7.21) into (7.15) gives the mean-field dynamics of the first spatial moment:

$$\frac{\mathrm{d}Z_1}{\mathrm{d}t} = (\rho - d)Z_1 + \gamma(2\pi\sigma_\rho^2)Z_1^2. \tag{7.22}$$

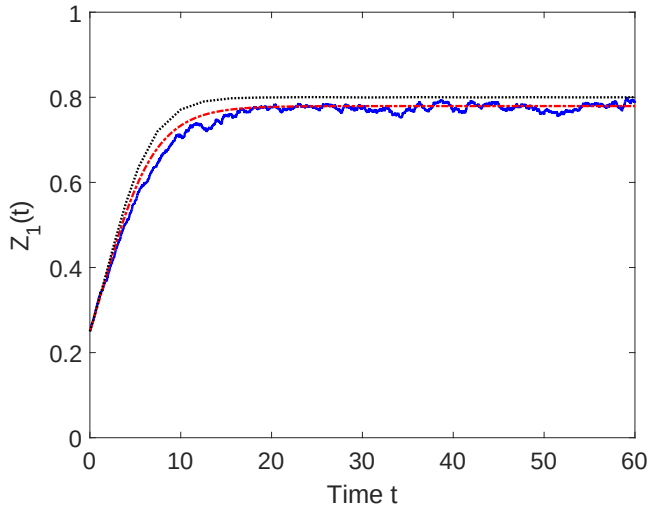
Here, the first term represents growth of the population arising due to intrinsic cell birth and death. The second term accounts for the overall effect of interactions affecting the expected birth rate. In the absence of interactions (i.e. for $\gamma = 0$), the population grows exponentially with rate $(\rho - d)$ and the average cell density over time is described by

$$Z_1(t) = Z_1(0)\exp\left((\rho - d)t\right). \tag{7.23}$$

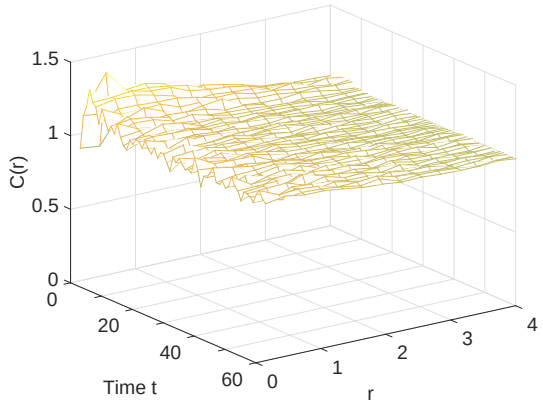
When interactions are present (i.e. for $\gamma < 0$), the population undergoes logistic growth and reaches an equilibrium average cell density

$$Z_1^* = \frac{d - \rho}{\gamma(2\pi\sigma_\rho^2)}. \tag{7.24}$$

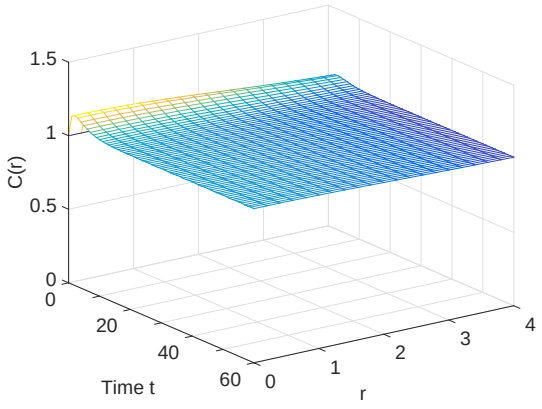
In the original IBM (for movement only) spatial structure was generated both by the neighbour-dependent component of cell movement rate and by the neighbour-dependent directional bias. With cell birth and death there are two additional mechanisms that can affect the spatial structure: short-range dispersal of offspring and short-range neighbourhood interactions affecting cell birth rate. Therefore, in order to simulate mean-field conditions where there is no spatial structure, we use a wide interaction kernel $w_\rho(\boldsymbol{\xi})$ (such that pairs of cells can interact over long spatial ranges) and a wide dispersal kernel. In addition we switch off the structure-generating mechanisms of neighbour-dependent motility and directional bias by setting $\alpha = \beta = 0$. Figure 7.1 shows the first and second spatial moments over time for a cell population undergoing birth, death and intrinsic movement (i.e. neighbourhood interactions do not affect cell movement rate or direction). We consider the growth of the population for a non-spatial case by simulating the IBM using a wide interaction kernel $w_\rho(\boldsymbol{\xi})$ and wide dispersal kernel $\mu^{(\rho)}(\boldsymbol{\xi})$. Under these conditions the population undergoes logistic growth and reaches an equilibrium average cell density of approximately 0.8, as shown in Figure 7.1(a). Initially the average cell density is low so the overall effect of neighbourhood interactions is



(a)



(b)



(c)

Figure 7.1: Mean-field dynamics for a cell population undergoing birth, death and intrinsic movement ($\alpha = 0$, $\beta = 0$) in a 20 x 20 domain. Initially, the size of the cell population is $N(0) = 100$. The IBM is simulated until $t = 60$ and results are averaged over 20 realisations. Interactions and disperal occur over a large spatial range ($\sigma_\rho = \sigma_{\mu_\rho} = 2$). Parameters are $m = 5$, $\rho = 1$, $d = 0.6$, $\gamma = -0.5/(2\pi\sigma_\rho^2) = -0.02$, $\lambda_\mu = 2.5$, $\sigma_\mu = 0.1$. (a) First spatial moment over time predicted by the IBM (solid blue line) is compared to that approximated by the spatial moment model (red broken line), solved using $\Delta = 0.2$, $\xi_{max} = 8$ and a power-2 (4,1,1) closure. The mean-field approximation of population growth (the numerical solution of (7.22) which assumes $Z_2(\boldsymbol{\xi}) = Z_1^2$) is shown as the black dotted line. (b) Time series of the PCF $C_{IBM}(r)$ computed from IBM simulations (with bin width $\delta r = 0.2$). (c) Time series of the PCF $C_{SM}(r)$ predicted by the spatial moment model.

sufficiently weak to allow the population to grow (i.e. $\hat{B}_i > \hat{D}_i$). As the average cell density increases, neighbourhood interactions start to have a greater effect until eventually they are strong enough that $\hat{B}_i \approx \hat{D}_i$ and the population reaches steady state.

The numerical solution to equation (7.22), for the mean-field dynamics of the first spatial moment, describes the growth of the population in the complete absence of spatial structure

and agrees well with the IBM's prediction. However, the mean-field dynamics rely on the assumption that $Z_2(\boldsymbol{\xi}) = Z_1^2$ over time (corresponding to a PCF $C(r) = 1$) while the IBM predicts that a very small degree of spatial structure arises in the population initially (Figure 7.1(b)). The dispersal of offspring that accompanies cell birth generates a weak cluster spatial pattern because even though the dispersal kernel is wide, it is still slightly concentrated at small displacements. In other words, a daughter cell has a slightly higher probability of being dispersed a short distance from the parent cell than a large distance, giving rise to weak clustering. Therefore, spatial structure is most noticeable initially when the overall birth rate $\hat{B}_i$ tends to be high because the combined effect of neighbourhood interactions is weak. The spatial structure dissipates as the population grows and the effect of neighbourhood interactions becomes strong enough to counteract the clustering generated by dispersal. At steady state the PCF $C_{IBM}(r) \approx 1$, however, because the Gaussian kernels are slightly concentrated at short displacements, $C_{IBM}(r)$ will always be slightly greater than 1 at short displacements. The presence of this weak spatial structure explains why the IBM predicts slightly slower population growth initially and a lower steady state average cell density than that of the mean-field dynamics.

The numerical solutions to equations (7.15)-(7.17) for the dynamics of the first and second spatial moments, employing the same wide kernels, are also shown in Figure 7.1. The moment model provides a very good approximation to both the first and second spatial moments over time for this non-spatial case. We now consider how spatial structure is generated by neighbourhood interactions affecting cell birth and by the dispersal of offspring, and explore the effect of this spatial structure on population growth.

7.3.3 Moment Closure

In our original model we employed a power-3 closure in order to close the system of spatial moment dynamics at second order. While this closure performed well when only cell movement was considered, except for strongly clustered spatial patterns, we now compare its performance against other choices of closure for the extended model with cell birth and death. Figure 7.2 shows the dynamics of the first and second spatial moments predicted by the IBM and the spatial moment model for four different closures. We consider a symmetric power-1 closure, symmetric power-2 closure, asymmetric power-2 closure with weighting $\{4, 1, 1\}$, and power-3 closure [79]. These closures tend to perform well when the spatial pattern is

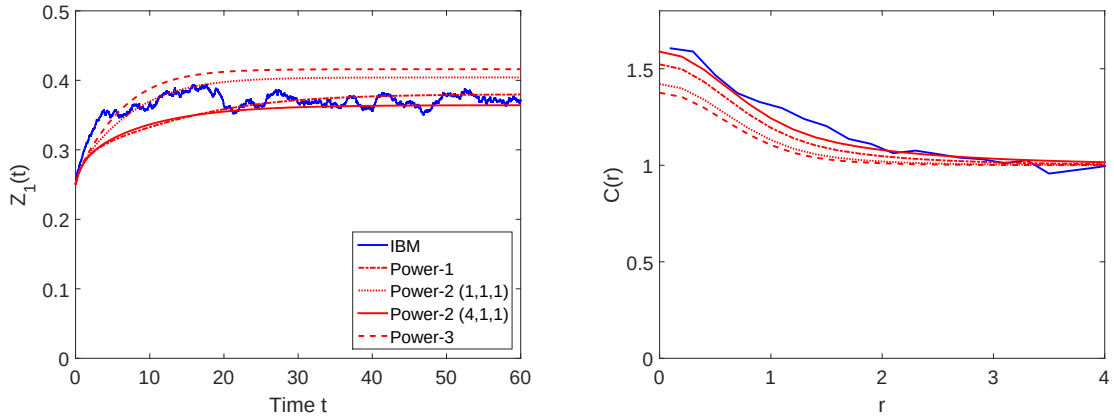


Figure 7.2: Comparison of closures for the third spatial moment. Collective dynamics of a cell population, of initial size $N(0) = 100$, undergoing birth, death and movement in a 20 x 20 domain. Parameters are $m = 5$, $\rho = 1$, $d = 0.6$, $\alpha = \beta = 0$, $\gamma = -0.5$, $\sigma_\rho = \sigma_{\mu_\rho} = 0.5$, $\lambda_\mu = 2.5$, $\sigma_\mu = 0.1$. Dynamics of spatial moments predicted by the IBM (solid blue line) (averaged over 20 realisations) are compared to those approximated by the spatial moment model ($\Delta = 0.2$ and $\xi_{max} = 8$) solved using a power-1 closure (red dashed-dotted line), power-2 closure with weighting $\{1, 1, 1\}$ (red dotted line), power-2 closure with weighting $\{4, 1, 1\}$ (red solid line) and power-3 closure (red dashed line): (a) First spatial moment $Z_1(t)$ over time t; (b) Second spatial moment, expressed as a PCF $C(r)$ (with bin width $\delta r = 0.2$), computed at $t = 60$, by which time the system is at steady state.

Poisson or regular but are more limited for cluster spatial patterns [96], therefore we carry out our comparison under conditions that generate clustering. Figure 7.2(a) shows the first spatial moment over time predicted by the four closures. All closures predict a steady state average cell density below the mean-field density, with the power-1 and asymmetric power-2 closures providing the best fit to the IBM. In Figure 7.2(b) we compare the second spatial moment estimated by each closure once the system has reached steady state. While all the closures underestimate the second moment to some extent, the power-3 closure provides the poorest fit to the IBM. The asymmetric power-2 closure with weighting $\{4, 1, 1\}$ has the best performance under these conditions, therefore we employ this closure in the remainder of our numerical analysis.

7.3.4 Width of Interaction and Dispersal Kernels

We now consider how spatial structure is generated by neighbourhood interactions and dispersal, and how this in turn affects the growth of a population undergoing birth, death and intrinsic movement. In some of the results below we see similar trends to those predicted by the spatial logistic model in studies by Bolker and Pacala (1997) [80] and Law, Murrell and

Dieckmann (2003) [65]. Before considering the combined effect of neighbour-dependent birth rate and dispersal, it is first helpful to consider what kind of spatial structure each mechanism is independently generating. When dispersal of offspring occurs over a short spatial range (i.e. using a narrow dispersal kernel $\mu^{(\rho)}(\boldsymbol{\xi})$) this generates clusters of individuals. However, the independent effects of neighbour-dependent birth cannot be so easily determined because this mechanism's ability to affect spatial structure is dependent upon dispersal. When neighbourhood interactions occur over a short range (narrow interaction kernel $w_\rho(\boldsymbol{\xi})$) the birth rates of individuals in close proximity are reduced. Therefore, if the spatial pattern is clustered, individuals separated by larger displacements give birth at a higher rate than those within clustered regions, thereby filling up the remaining space with dispersed offspring and breaking down the spatial structure. On the other hand, for a regular spatial pattern, individuals are less likely to have close-lying neighbours so the overall effect of short-range neighbourhood interactions is low. Thus, the birth rates for individuals are higher than would be expected in the absence of any spatial structure. This leads to a rapid filling of the space with dispersed offspring which breaks down the regular pattern. Therefore, short-range neighbourhood interactions act to break down spatial structure, promoting a Poisson spatial pattern, but this effect is reliant on the dispersal of offspring.

In the non-spatial case we employed wide kernels such that cell interactions and dispersal of offspring occurred over a long spatial range. This resulted in a Poisson spatial pattern and logistic population growth. Figure 7.3 shows how employing kernels of different widths affects the first and second spatial moments. When varying the width of interaction kernel $w_\rho(\boldsymbol{\xi})$ we set interaction strength $\gamma = -0.5/(2\pi\sigma_\rho^2)$. This ensures that while the range of neighbourhood interactions is varied (by choosing different values of σ_ρ), the overall summed contribution of interactions remains the same (i.e. the integral of $w_\rho(\boldsymbol{\xi})$ over $\boldsymbol{\xi}$ remains fixed). If γ were instead fixed, the steady state average cell density would merely tend to zero with increasing kernel width.

We first consider a case where cells interact over a long spatial range (wide interaction kernel $w_\rho(\boldsymbol{\xi})$) and offspring are dispersed over a short range (narrow dispersal kernel $\mu^{(\rho)}(\boldsymbol{\xi})$) (Fig. 7.3(a)-(d)). In this case, the short-range dispersal of offspring promotes clustering, as indicated by $C(r) > 1$ at short displacements (Fig. 7.3(d)). The cluster spatial pattern is relatively strong because the wide interaction kernel can do little to counteract this spatial structure. In the non-spatial dynamics, the population grew until it reached a steady state

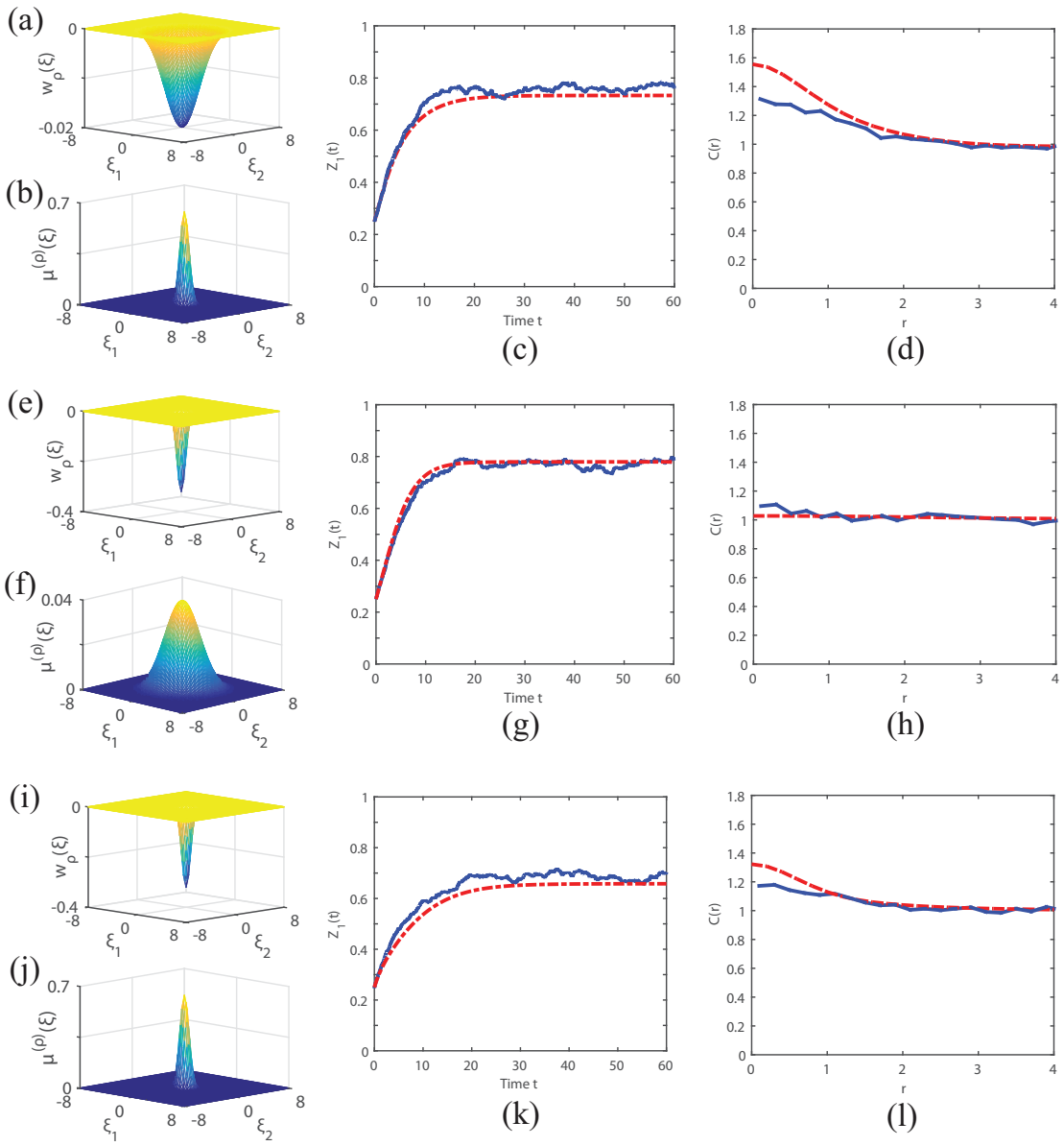


Figure 7.3: Effect of varying width of interaction kernel $w_\rho(\boldsymbol{\xi})$ and dispersal kernel $\mu^{(\rho)}(\boldsymbol{\xi})$. Collective dynamics for a cell population, of initial size $N(0) = 100$, undergoing birth, death and intrinsic movement ($\alpha = 0$, $\beta = 0$) in a 20 x 20 domain. The IBM is simulated until $t = 60$ and results are averaged over 20 realisations. First and second spatial moments predicted by the IBM (solid blue line) are compared to those approximated by the spatial moment model (red broken line), solved using $\Delta = 0.2$ and $\xi_{max} = 8$. Parameters are $m = 5$, $\rho = 1$, $d = 0.6$, $\lambda_\mu = 2.5$, $\sigma_\mu = 0.1$. Row 1: Kernel widths are $\sigma_\rho = 2$, $\sigma_{\mu_\rho} = 0.5$, and $\gamma = -0.5/(2\pi\sigma_\rho^2) = -0.02$. Row 2: Kernel widths are $\sigma_\rho = 0.5$, $\sigma_{\mu_\rho} = 2$, and $\gamma = -0.5/(2\pi\sigma_\rho^2) = -0.32$. Row 3: Kernel widths are $\sigma_\rho = 0.5$, $\sigma_{\mu_\rho} = 0.5$, and $\gamma = -0.5/(2\pi\sigma_\rho^2) = -0.32$. (a), (e), (i) Interaction kernel $w_\rho(\boldsymbol{\xi})$; (b), (f), (j) dispersal kernel $\mu^{(\rho)}(\boldsymbol{\xi})$; (c), (g), (k) first spatial moment $Z_1(t)$ over time; (d), (h), (l) PCF $C(r)$ (with bin width $\delta r = 0.2$) computed at $t = 60$.

average cell density $Z_1^* = 0.8$, by which point the combined effect of neighbourhood interactions was sufficiently strong to limit further growth. However, when the spatial pattern is clustered the steady state average cell density is slightly lower than for the non-spatial case,

as shown in Figure 7.3(c). Despite the weak interaction weight associated with a wide kernel $w_\rho(\boldsymbol{\xi})$, close neighbours interact more strongly than those displaced far apart because the wide Gaussian kernel is still slightly concentrated around short displacements. The presence of a cluster spatial pattern means that cell pairs are more likely to be separated by short displacements where neighbourhood interactions have the strongest effect. Therefore, despite the weak interaction weight, the threshold average cell density at which the combined effect of neighbourhood interactions starts to restrict population growth is lower than if spatial structure were absent. On the other hand, if the interaction kernel was perfectly uniform with respect to separation distance, then the presence of clustering would have no effect on the steady state average cell density. These results agree well with those predicted by the spatial logistic model, in which clustering generated by short-range dispersal and long-range interactions also gave rise to a lower steady state average cell density than that of a non-spatial case [65, 80].

In the absence of cell movement, the population would approach its steady state average cell density at a slower rate than the non-spatial case because the initially uninhabited space could only be filled slowly by the dispersal of offspring over short distances [65]. In contrast, the long-range dispersal in the non-spatial dynamics would facilitate more rapid filling of the available space. However, with motile cells the initial population growth is only slightly slower than for the non-spatial dynamics. This is because movement predominates over dispersal as the space-filling mechanism ($m \gg \rho$), thus altering the dispersal range has little consequence for the initial rate of population growth.

Next, we consider a case where neighbourhood interactions occur over a short spatial range while offspring are dispersed over a wide range (Fig. 7.3(e)-(h)). Here, the wide dispersal kernel can promote only very weak clustering and this spatial structure is broken down by the strong, short-range interactions which greatly reduce the birth rates of any close-lying neighbours. This results in a Poisson spatial pattern (indicated by $C(r) \approx 1$) (Fig. 7.3(h)) and logistic population growth matching the mean-field dynamics (Fig. 7.3(g)).

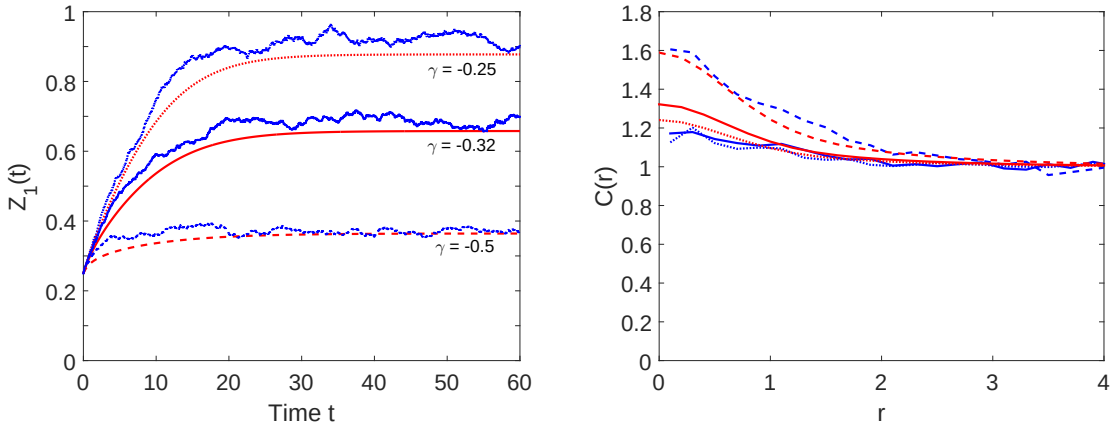
Finally, we explore the collective dynamics when both neighbourhood interactions and dispersal of offspring occur over a short spatial range (Fig. 7.3(i)-(l)). Again, the short-range dispersal promotes clustering in the population. However, in this case the short-range neighbourhood interactions counteract this effect, partially breaking down the spatial structure. The resulting cluster spatial pattern (Fig. 7.3(l)) is therefore weaker than that observed in

Figure 7.3(d) where long-range interactions could do little to counteract the spatial structure. Despite this weaker cluster pattern, the steady state average cell density (Fig. 7.3(k)) is even lower than that shown in Figure 7.3(c). This is likely owing to the fact that the short-range neighbourhood interactions are considerably stronger at short displacements than the long-range interactions. Fixing $\gamma = -0.5/(2\pi\sigma_\rho^2)$ ensures that, for a Poisson spatial pattern, varying σ_ρ does not change the summed contribution from neighbours to the birth rate B_1. However, for a particular cluster spatial pattern the summed contribution from neighbours is greater (in magnitude) for the small kernel width $\sigma_\rho = 0.5$ than the large kernel width $\sigma_\rho = 2$. Therefore, with the same extent of clustering at steady state, the narrow interaction kernel would result in a lower birth rate B_1 and consequently a lower steady state average cell density compared to the wide kernel. In Figure 7.3 the steady state cluster spatial pattern is slightly weaker for the narrow interaction kernel, however it is still sufficiently strong to result in a lower steady state average cell density (Fig. 7.3(k)) than for the wide kernel (Fig. 7.3(c)).

It is also worth noting that the rate at which the population grows to its steady state average cell density is slowest in the third case. As discussed above, this is because the available space is filled more slowly by offspring dispersed over short distances compared to long-range dispersal, even though cell movement facilitates faster filling of the space. The slow growth rate is more noticeable here because, at the initially low average cell densities, short-range dispersal quickly generates clustering and offspring are deposited close enough to parent cells that they are immediately within the range at which interactions are very strong (compared to the wide interaction kernel). Therefore, parents interact strongly with their own offspring and even at early times the overall strength of short-range interactions is already sufficient to reduce birth rate and slow population growth. A similar result was observed with the spatial logistic model, in which short-range interactions and dispersal also gave rise to slow population growth and a particularly low steady state average cell density [65, 80].

7.3.5 Neighbour-dependent Birth

We now consider the effect on the collective dynamics of varying the strength of neighbourhood interactions affecting cell birth rate (by varying interaction strength γ). We employ the same interaction and dispersal kernel widths as in Figure 7.3(i)-(j) such that neighbourhood interactions and the dispersal of offspring occur over a short spatial range. Figure 7.4 com-



(a) (b)

Figure 7.4: Effect of varying interaction strength γ. Collective dynamics for a cell population, of initial size $N(0) = 100$, undergoing birth, death and intrinsic movement ($\alpha = 0$, $\beta = 0$) in a 20 x 20 domain. The IBM is simulated until $t = 60$ and results are averaged over 20 realisations. First and second spatial moments predicted by the IBM (blue lines) against those approximated by the spatial moment model (red lines), solved using $\Delta = 0.2$ and $\xi_{max} = 8$. Parameters are $m = 5$, $\rho = 1$, $d = 0.6$, $\lambda_\mu = 2.5$, $\sigma_\mu = 0.1$, $\sigma_\rho = \sigma_{\mu_\rho} = 0.5$. Three different values of interaction strength γ are considered: $\gamma = -0.25$ (dotted lines), $\gamma = -0.32$ (solid lines) and $\gamma = -0.5$ (dashed lines). The corresponding mean-field equilibrium average cell density (correct to 2 d.p.) is $Z_1^* = 1.02$ for $\gamma = -0.25$, $Z_1^* = 0.80$ for $\gamma = -0.32$, and $Z_1^* = 0.51$ for $\gamma = -0.5$. (a) First spatial moment $Z_1(t)$ over time t; (b) Second spatial moment, expressed as a PCF $C(r)$ (with bin width $\delta r = 0.2$), computed at $t = 60$.

pares the first and second spatial moments for the strength $\gamma = -0.32$ (employed previously in Fig. 7.3(k)-(l)) and two other values of $\gamma < 0$. For each value of $\gamma < 0$, a cluster spatial pattern is generated (Fig. 7.4(b)) and the steady state average cell density is below that predicted by the mean-field dynamics (Fig. 7.4(a)). When kernel widths are fixed, increasing the strength of interactions (i.e. increasing the magnitude of $\gamma < 0$) reduces the steady state average cell density (Fig. 7.4(a)) but results in a greater extent of spatial structure at steady state (Fig. 7.4(b)). To understand why this is the case it helps to consider the growth of the population as it unfolds.

Initially, the spatial pattern is Poisson and average cell density is low so cells are less likely to come into close enough proximity for short-range interactions to have any effect. This means the birth rates of cells in the initial population are high on average and the population starts to grow. The short-range dispersal of offspring that accompanies cell birth quickly generates strong clustering in the regions of space occupied by each parent cell from the initial population, although cell movement diffuses this spatial structure to some extent. Within these clusters, neighbouring cells are more likely to be located close enough together

for short-range neighbourhood interactions to come into play. Owing to the cluster spatial pattern, for high magnitudes of $\gamma < 0$ (e.g. $\gamma = -0.5$) the overall effect of neighbourhood interactions is already sufficiently strong at these low average densities that cell birth rates are significantly reduced. In this case the population grows slowly to a low steady state average cell density beyond which further population growth is restricted by neighbourhood interactions (Fig. 7.4(a)). In Figure 7.3(h) we saw that strong short-range neighbourhood interactions can break down spatial structure to promote a Poisson spatial pattern as average cell density increases. However, because the steady state average cell density is low here, there has not been much opportunity for neighbourhood interactions to break down the cluster spatial pattern and it remains strong at steady state (Fig. 7.4(b)).

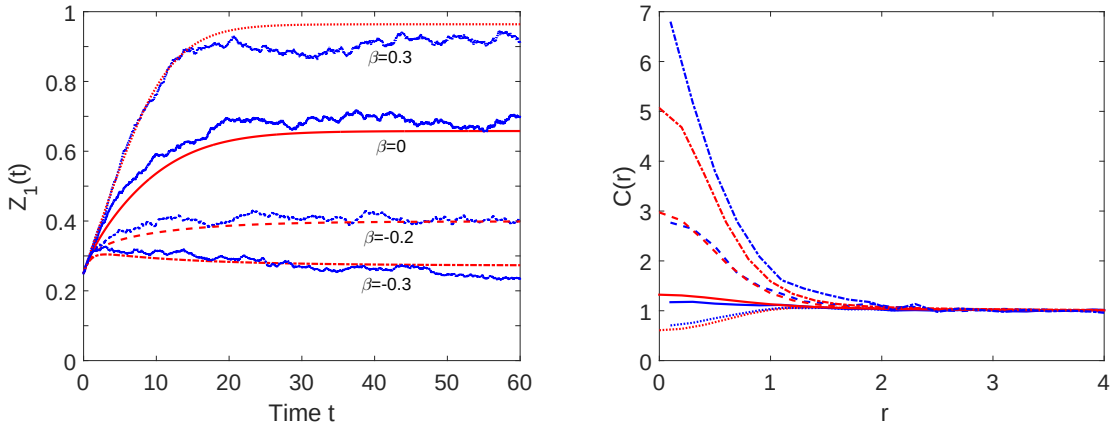
In contrast, when the magnitude of $\gamma < 0$ is small (e.g. $\gamma = -0.25$), the overall effect of neighbourhood interactions is insufficient to slow population growth at early times when average cell density is low, in spite of the cluster spatial pattern. Therefore, the population undergoes rapid growth initially. The cluster pattern weakens over time because short-range interactions reduce the birth rates of cells in close proximity while those separated by larger displacements give birth at high rates and rapidly fill the available space with offspring. This continues until average cell density reaches the threshold at which interactions are strong enough overall to halt population growth. Therefore, at steady state the extent of spatial structure that remains is less than for higher magnitudes of $\gamma < 0$, as shown in Figure 7.4(b).

As previously mentioned, we only consider values of $\gamma < 0$ that do not violate the positivity constraint of the expected birth rate B_1 and $B_2(\boldsymbol{\xi})$. However, owing to the IBM's inherent stochasticity, negative individual birth rates $\hat{B}_i$ can arise at a particular time by chance and the maximum formula in (7.3) ensures these rates are set to $\hat{B}_i = 0$. For the results presented in Figure 7.4, these rates $\hat{B}_i = 0$ represented 1.3% of the birth rates of all cells over time when $\gamma = -0.25$, 2.4% when $\gamma = -0.32$ and 5.6% when $\gamma = -0.5$. Despite the spatial moment model not accounting for this effect, we see a reasonable agreement between the first and second spatial moments predicted by the IBM and moment approximation over time. Employing larger magnitudes of $\gamma < 0$ (e.g. $\gamma = -0.6$) in simulations of the IBM causes the population size to decline rather than grow and can lead to extinction. However, we do not consider such cases here because the required magnitudes of $\gamma < 0$ are sufficiently large to violate the positivity constraint of B_1 and $B_2(\boldsymbol{\xi})$.

7.3.6 Neighbour-dependent Birth with Directional Bias

So far we have explored the collective dynamics for a cell population undergoing neighbour-dependent birth and death with only intrinsic movement. We now consider how the dynamics are affected when movement with neighbour-dependent directional bias is incorporated. Figure 7.5 shows the first and second spatial moments for different values of interaction strength β, where the width of interaction/dispersal kernels and interaction strength γ are all fixed with the same values employed in Figure 7.4. For ease of comparison we also show the results for $\beta = 0$ that were described in the previous section. We previously explained that for values of $\beta > 0$ the directional bias acts as a repulsive force to promote a regular spatial pattern while for values of $\beta < 0$ the force is attractive and promotes clustering. In addition, increasing the magnitude of β leads to stronger spatial structure.

Figure 7.5(b) shows that for $\beta < 0$ the directional bias and dispersal mechanisms work cooperatively to generate a stronger cluster pattern than would occur in the absence of bias. Initially, the overall effect of neighbourhood interactions is low so short-range dispersal is likely to be the predominant mechanism driving clustering. However, as average density increases (Fig. 7.5(a)), neighbourhood interactions become stronger and directional bias plays a greater role as an attractive force, driving cells towards clustered regions and rapidly strengthening the cluster spatial pattern. The steady state average cell density is therefore lower than in the absence of bias (e.g. $\beta = -0.2$ in Fig. 7.5(a)) because with stronger clustering the overall effects of neighbour-dependent birth can halt population growth at lower average cell densities. Figure 7.5(a) shows that a sufficiently strong directional bias towards crowded regions (e.g. $\beta = -0.3$) can actually cause the population to decline. At first, average cell density increases because the combined effect of neighbourhood interactions is weak for the initially Poisson pattern of cells. However, the combination of short-range dispersal and strong directional bias quickly generates clustering to the extent that the expected neighbour-dependent birth rate is reduced below the value of the expected death rate. This causes the population to decline to a steady state average cell density lower than the initial density (Fig. 7.5(a)). The spatial structure at steady state is a particularly strong cluster spatial pattern in this case (Fig. 7.5(b)), so strong in fact that there is some disagreement between the predictions of the IBM and the moment approximation. This can likely be attributed to the deterioration in performance of the power-2 moment closure as an approximation for



(a) (b)

Figure 7.5: Effect of directional bias for different interaction strengths β. Collective dynamics for a cell population, of initial size $N(0) = 100$, undergoing birth, death and movement in a 20 x 20 domain. The IBM is simulated until $t = 60$ and results are averaged over 20 realisations. First and second spatial moments predicted by the IBM (blue lines) are compared to those approximated by the spatial moment model (red lines), solved using $\Delta = 0.2$ and $\xi_{max} = 8$. Parameters are $m = 5$, $\rho = 1$, $d = 0.6$, $\lambda_\mu = 2.5$, $\sigma_\mu = 0.1$, $\alpha = 0$, $\gamma = -0.32$, $\sigma_v = \sigma_\rho = \sigma_{\mu_\rho} = 0.5$. Four different values of interaction strength β are considered: $\beta = 0.3$ (dotted lines), $\beta = 0$ (solid lines), $\beta = -0.2$ (dashed lines) and $\beta = -0.3$ (dashed-dotted lines). (a) First spatial moment $Z_1(t)$ over time t; (b) Second spatial moment, expressed as a PCF $C(r)$ (with bin width $\delta r = 0.2$), computed at $t = 60$.

increasingly strong cluster spatial patterns.

On the other hand, for $\beta > 0$ (e.g. $\beta = 0.3$ in Fig. 7.5) the directional bias acts as a repulsive force which counteracts the cluster-generating effects of short-range dispersal. At very early times, the spatial pattern is clustered because the overall weak effect of neighbourhood interactions enables short-range dispersal to rapidly generate structure (i.e. birth rate is high) while biased movement is insufficiently strong to completely break down this structure (results not shown). However, as the average density increases, neighbour-dependent biased movement overcomes short-range dispersal as the predominant structure-generating mechanism and eventually leads to a regular pattern at steady state (Fig. 7.5(b)). This regular spatial pattern reduces the proximity between neighbours so that the overall effect of interactions on an individual's birth rate is lower than would be expected in the absence of spatial structure. As a result, neighbour-dependent birth permits growth of the population to a higher steady state average cell density than that of the mean-field dynamics (Fig. 7.5(a)).

The results presented here share similar traits to those of the spatial logistic model [65,80]. In both cases, the steady state average cell density is lower than that expected from the non-spatial dynamics for a cluster spatial pattern and higher than expected for a regular pattern.

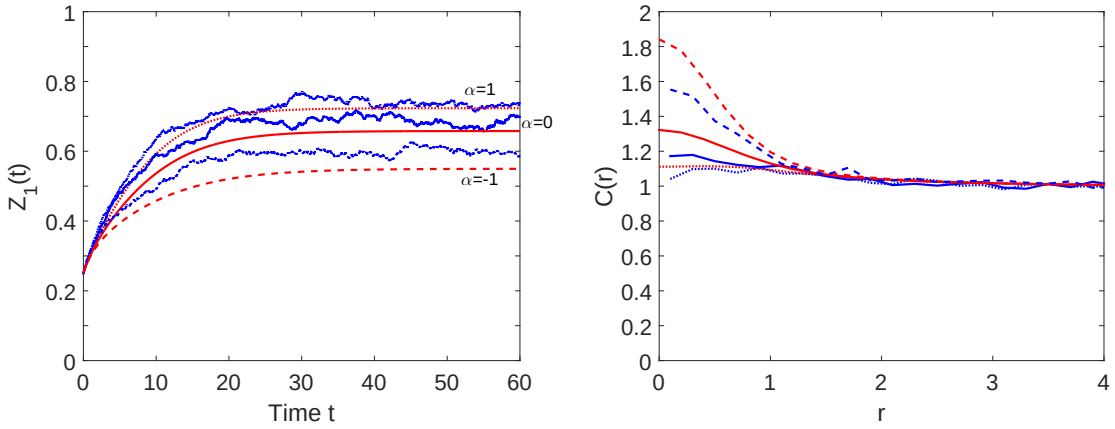


Figure 7.6: Effect of neighbour-dependent motility for different interaction strengths α. Collective dynamics for a cell population, of initial size $N(0) = 100$, undergoing birth, death and movement in a 20 x 20 domain. The IBM is simulated until $t = 60$ and results are averaged over 20 realisations. First and second spatial moments predicted by the IBM (blue lines) are compared to those approximated by the spatial moment model (red lines), solved using $\Delta = 0.2$ and $\xi_{max} = 8$. Parameters are $m = 5$, $\rho = 1$, $d = 0.6$, $\lambda_\mu = 2.5$, $\sigma_\mu = 0.1$, $\beta = 0$, $\gamma = -0.32$, $\sigma_w = \sigma_\rho = \sigma_{\mu_\rho} = 0.5$. Three different values of interaction strength α are considered: $\alpha = 1$ (dotted lines), $\alpha = 0$ (solid lines) and $\alpha = -1$ (dashed lines). (a) First spatial moment $Z_1(t)$ over time t; (b) Second spatial moment, expressed as a PCF $C(r)$ (with bin width $\delta r = 0.2$), computed at $t = 60$.

A similar decline in the population was also observed with the spatial logistic model for particularly strong clustering generated by short-range interactions/dispersal [65].

7.3.7 Neighbour-dependent Birth and Motility

Finally, we consider the effect of neighbour-dependent motility on the dynamics of a cell population undergoing neighbour-dependent birth and death (but in the absence of directional bias). Previously, we showed that for $\alpha > 0$ neighbourhood interactions give rise to an increase in movement rate which promotes a regular spatial pattern. For $\alpha < 0$, movement rate is reduced by interactions and thereby promotes clustering. Figure 7.6 shows the first and second spatial moments for a value of $\alpha > 0$ and $\alpha < 0$. Results for $\alpha = 0$ are also shown for comparison.

As was the case for movement with directional bias, the neighbour-dependent motility mechanism can work cooperatively with short-range dispersal to promote clustering when $\alpha < 0$. Offspring dispersed within close proximity of other individuals have their motility rates reduced thereby making it harder for them to escape the spatial structure. This results in a stronger cluster spatial pattern at steady state (Fig. 7.6(b)) and thus a lower steady state

average cell density than occurs with only intrinsic movement (Figure 7.6(a)). In contrast, for $\alpha > 0$, neighbour-dependent motility counteracts short-range dispersal enabling cells within clustered regions to move at a faster rate which diffuses the spatial structure. For the value of $\alpha < 0$ employed in Figure 7.6, neighbour-dependent motility is sufficiently strong to partially counteract the clustering generated by short-range dispersal but too weak to generate a regular spatial pattern. Therefore, at steady state the spatial structure is very weakly clustered (Fig. 7.6(b)) and average cell density is lower than that expected in the absence of neighbour-dependent movement (Fig. 7.6(a)).

7.4 Discussion

Incorporating birth and death into our model for cell movement provides insights into how these key processes generate spatial structure and how, in turn, this structure impacts the growth of a cell population. We are particularly interested in the role that directional bias plays in these dynamics. We previously found that both neighbour-dependent movement rate and directional bias generated spatial structure in the system. In the presence of cell birth and death, spatial structure is also affected by the dispersal of offspring and a neighbour-dependent birth rate. Our numerical results show that while dispersal generates clustering, neighbourhood interactions that reduce an individual's birth rate act to break down the structure to promote a Poisson spatial pattern. Varying the range over which dispersal/interactions occur (i.e. by varying kernel widths) or varying the strength of interactions affects the extent to which spatial structure is generated. Our results show that spatial structure can cause both the rate of population growth and the average cell density at steady state to differ from that of the non-spatial dynamics. In general, the presence of clustering causes slower growth to a lower than expected steady state average cell density. A regular spatial pattern allows the population to grow rapidly to average cell densities above those observed in the non-spatial dynamics.

Movement, in particular directional bias, also plays an important role in the population dynamics. Intrinsic movement diffuses the spatial structure because offspring can move out of the clusters generated by short-range dispersal. Neighbour-dependent directional bias away from regions of crowding is more effective at breaking down this structure because cells undergo directed movement out of clusters. If the bias is sufficiently strong, movement can

completely counteract the effects of dispersal to generate a regular spatial pattern. On the other hand, bias towards crowded regions strengthens the clustering generated by dispersal. A neighbour-dependent movement rate gives rise to similar effects: either clustering is broken down because cells with close neighbours move with increased rate, or the pattern is strengthened due to low motility of clustered cells. However, directional bias is the stronger of these two mechanisms.

Our numerical results display similar trends to those predicted by the spatial logistic model [65, 80], despite some differences in the underlying mechanisms of each model. For instance, we observe similar effects of clustering and regular spatial patterns on the growth of the population. In general, varying the kernel widths resulted in greater departures from the mean-field dynamics with the spatial logistic model than were observed in our results. This is because the spatial logistic model did not incorporate movement, therefore the spatial structure generated by short-range interactions and dispersal was stronger than for our model (where movement partially diffused the structure). In the spatial logistic model, competitive short-range neighbourhood interactions generated a regular spatial pattern by increasing the death rates of individuals with close neighbours. This is not observed with our model because, in the absence of neighbour-dependent movement, neighbourhood interactions can only break down the spatial structure generated by short-range dispersal. However, incorporating directional bias allows a wider range of spatial structures to be explored, including regular patterns. For these results we observe departures from the mean-field dynamics that are qualitatively similar to those of the spatial logistic model, even though the mechanisms generating the spatial structures differ [65].

We explore the effect of neighbourhood interactions on population growth by allowing the birth rate of an individual to comprise both an intrinsic component and a component which depends on the presence of other cells in its neighbourhood. Implementing this mechanism for neighbour-dependent birth alongside intrinsic death provides insights into how crowding effects, such as contact inhibition of proliferation, give rise to spatial structure and influence population growth. With neighbour-dependent birth the growing population eventually reaches a steady state average cell density, which is an observed feature of population growth studied *in vitro* [1]. However, other approaches which incorporate neighbour-dependent effects into birth and death processes are also discussed in the literature. For instance, models which implement intrinsic birth with neighbour-dependent death are popular in the ecology

literature as they allow the effects of competition or predation on population growth to be explored [65, 70, 80, 86]. One advantage of this approach is that an individual's death rate is increased by competitive neighbourhood interactions. Therefore, spatial logistic-type growth can still be observed but without the risk of an individual's birth or death rate ever becoming negative, even when interactions are strong. With neighbour-dependent birth, the range of interaction strengths we can consider is restricted by our need to avoid negative expected birth rates, thus implementing neighbour-dependent death provides greater scope for exploring parameter space.

Another important difference concerns the effect of each mechanism on the spatial structure. Our results show that a Poisson spatial pattern is promoted by short-range neighbourhood interactions that reduce birth rate. However, this ability to break down spatial structure is completely reliant on the dispersal of offspring, making the independent effect of interactions difficult to determine. A more rigorous analysis of parameter space, in particular kernel width and interaction strength, may provide more insights into this process. With short-range dispersal, which generates clustering, the overall structure is only ever a Poisson or a cluster spatial pattern. In terms of population growth this means that the resulting steady state average cell density can lie below that expected from the mean-field dynamics, but never above. In contrast, neighbour-dependent death promotes a regular spatial pattern because the death rates of individuals in close proximity are increased, resulting in the over-dispersal of surviving individuals [65]. This force is separated from dispersal because it does not affect birth rate so the independent effects of these two mechanisms can be extricated from the overall dynamics. In addition, neighbour-dependent death permits population growth to average cell densities above that of the mean-field dynamics, which is an effect we do not observe except in the presence of neighbour-dependent directional bias or motility.

By comparison, neighbour-dependent birth has received less attention in the literature. This could perhaps partly be explained by a prevalence of ecological examples that involve competition and predation compared to competitive interactions affecting birth rate. It is also plausible that the advantages of neighbour-dependent death mentioned above make it a more appealing approach. Bolker and Pacala (1999) derived a birth-death model for plant competition in which neighbourhood interactions reduced an individual's birth rate and rates that would otherwise be negative were set to zero. They also discussed how this nonlinearity affected the accuracy of the moment equations for some parameter ranges [105]. A predator-

prey model featuring neighbour-dependent birth was presented by Murrell (2005) [86]. However, in this case the predator's birth rate was fully dependent on neighbourhood interactions with prey (birth coincided with successful attacks on prey), i.e. there was no intrinsic birth. Therefore, the strength of interaction was not limited by the positivity constraint. Models describing birth rate as a sum of an intrinsic and neighbour-dependent component are also discussed by Dieckmann and Law (2000), and Plank and Law (2015) [61, 85].

Despite these differences, we see similar behaviour in the population dynamics predicted by our model implementing neighbour-dependent birth as were observed in the spatial logistic model with neighbour-dependent death [65, 80]. Incorporating neighbour-dependent death would also be biologically relevant for cell movement. For instance, the rate of cell death can be increased in overcrowded regions due to hypoxia (a lack of extracellular oxygen) or the local depletion of nutrients [106]. Therefore, there is scope for future work to explore the effects of neighbour-dependent death on the dynamics of a cell population, particularly in the presence of directional bias.

Our work has some important implications from a biological perspective, in particular for the *in vitro* study of cell movement. The results from our model would suggest that directional bias has the potential to drastically alter the behaviour of a growing population compared to if movement were unbiased. It would be useful to test this prediction in an *in vitro* setting. For instance, the growth of a population of cells that secrete a strong chemoattractant or -repellant could be compared to the growth in a population where the release of chemotactic factors was inhibited. If the dynamics do indeed significantly differ, as our model suggests is possible, then it is important to take this into account when making predictions about the outcomes of experiments with cells that undergo biased movement. In addition, if the aim of an experiment is to manipulate collective movement in some way, for example to enhance or retard movement, then the mechanisms that generate directional bias may be appropriate targets.

Chapter 8

Concluding Remarks

In this work we have developed a model for collective cell behaviour at two biological scales. Our model provides insights into how the short-range interactions experienced by cells lead to self-generated spatial structure in a population and what consequences this has for the collective dynamics. In particular, we focused on the effect of neighbour-dependent directional bias on cell movement. To explore the stochastic cell behaviour that arises at a microscopic scale, we first derived an individual-based model for movement in one-dimensional space. The IBM was constructed in a lattice-free framework to avoid the drawbacks associated with restricting movement along an artificially imposed grid. We incorporated mechanisms into our model that allowed an individual's direction and rate of movement to depend on the degree of crowing in its neighbourhood.

To scale up our model and consider behaviour at a macroscopic scale, we used our IBM to derive a population-level description in terms of the dynamics of spatial moments. Unlike the majority of population-level models, our moment approximation accounts for the presence of spatial structure in a population by including the dynamics of the second spatial moment (an average density of cell pairs).

By carrying out simulations and computing numerical solutions for a homogeneous space, we were able to show that the spatial moment model provided a good approximation to the underlying IBM, except in the presence of strong clustering. Analysis of our results showed that neighbourhood interactions affecting either the direction or rate of an individual's movement can give rise to spatial structure in a cell population. Both mechanisms were capable of behaving either as a repulsive force, to generate a regular spatial pattern, or as an attractive force to promote clustering. Qualitatively similar spatially homogeneous patterns are observed in cell populations cultured *in vitro* [31, 37, 74].

As well as offering insights into the role of directional bias for determining collective behaviour, the one-dimensional model provided a useful stepping stone for an extension to two-dimensional space. Our results confirmed that the moment approximation also agreed

well with the IBM's prediction in two spatial dimensions. Using imaging data generated from *in vitro* cell experiments, we were able to parameterise the model and validate that it was capable of predicting behaviour in a moving population of 3T3 fibroblast cells.

Finally, we carried out a further extension of the two-dimensional model to account for cell birth and death. This allowed us to consider how the growth of a population is affected by self-generated spatial structure, in the presence of directional bias. Our results showed that the short-range dispersal of daughter cells and a neighbour-dependent birth rate also affect spatial structure in a moving population. This spatial structure is diffused by intrinsic cell movement but can still cause the growth of the population to differ from the expected non-spatial growth. A neighbour-dependent directional bias causes even greater departures from the expected non-spatial growth because it is capable of generating especially strong spatial structure in the growing population.

Overall, our work highlights the importance of accounting for spatial structure in models of collective cell movement, particularly for populations affected by neighbour-dependent directional bias. Deriving a spatial moment approximation from an IBM with directional bias, allows us to relate the underlying stochastic cell behaviour to the emergent collective dynamics of the population and provides scope for further analysis. Movement with neighbour-dependent directional bias has received comparatively little attention in the moment dynamics literature to date and has mostly been considered in only one spatial dimension [71, 89]. Our work expands the knowledge in this area by exploring biased movement both in two-dimensional space and in the presence of cell birth and death. From a biological perspective, our model could prove useful for quantifying the behaviour of cell populations studied *in vitro*, especially for cell types that undergo biased movement. A model that can predict the outcomes of experiments could be particularly beneficial, for instance, in the context of drug testing.

Solving the equations for the dynamics of spatial moments required a closure assumption to truncate the system at second order. Employing different closures can lead to quite different model predictions and it is not immediately obvious which closure should be chosen in order to obtain the best performance. We tested a small number of closures in our model and found that the power-3 closure provided good performance overall for cell movement, both with neighbour-dependent motility and directional bias. However, in the presence of cell birth and death, this closure was no longer the most suitable choice. Instead, for neighbour-

dependent birth and intrinsic death the asymmetric power-2 closure with weighting $\{4, 1, 1\}$ provided better performance. This closure has also been applied successfully in models with neighbour-dependent death [65]. A more rigorous analysis of possible closures may reveal other suitable choices, but these two closures provide reasonable approximations for the neighbour-dependent mechanisms employed in our model.

There are several ways in which the work here could be extended. Applying our model in a spatially homogeneous setting is an important first step towards understanding how spatial structure is generated in moving cell populations and how this spatial structure affects collective behaviour. However, extending this study to explore collective movement through a non-homogeneous space would widen the scope for application of our model in a variety of contexts. For example, this could provide key insights into the formation and behaviour of invasion waves of individuals, such as those arising in spreading populations of malignant tumour cells [7] or in wound healing [8]. From an experimental perspective, models for movement in a non-homogeneous space could prove incredibly useful for understanding and quantifying behaviour in cell invasion assays, such as scratch or circular barrier assays [45].

The tools for exploring a non-homogeneous space are already in place as we have derived the equations for the dynamics of spatial moments in terms of physical locations. However, solving the equations numerically may require some careful thought. Expressing the equations in terms of pair separation displacements for a homogeneous space reduced the number of variables, therefore the numerical solutions were reasonably straightforward to compute. Solving numerically for a non-homogeneous space will require heavier computations, particularly in two spatial dimensions, as the spatial moments need to be calculated at each physical location over time. The implementation of boundary conditions for the spatial moment dynamics in a non-homogeneous setting will also require careful consideration. A homogeneous space allowed us to make the assumption that conditions at the boundary were approximately mean-field such that there was no spatial structure at large pair displacements. However this boundary condition could not be applied to a non-homogeneous space where the average density of a pair is dependent on the physical location of each agent in space.

Many *in vitro* studies of cell populations generate experimental data in two spatial dimensions, such as time-lapse images, and our two-dimensional model is amenable for direct application to this kind of data. There is also scope for further extension of the model to describe cell movement through three-dimensional space. This would allow the model to be

used in conjunction with three-dimensional experimental techniques, such as a Transwell migration assay [45,107]. However, numerically solving the equations for the dynamics of spatial moments in three-dimensional space would be considerably more computationally intensive. For example, an integration over a location in three-dimensional space would actually involve an integration over each of the three dimensions. Therefore, it would be worthwhile to consider how the efficiency of the numerical computations could be improved.

Our model is constructed so as to allow some flexibility in the choice of kernels and rate functions. Therefore, the model can be readily adapted to suit a variety of different experimental situations. Furthermore, the model could be easily extended to describe collective behaviour in a population consisting of different types of cells. General frameworks for the dynamics of spatial moments for different types of interacting species are already discussed in the literature [61].

By closing the dynamical system of spatial moments at second order, we neglect the presence of spatial structure at higher orders. This provides spatial information at the most basic level yet still provides valuable insight into systems where short-range interactions play a predominant role. However, extending the moment approximation to account for the dynamics of higher spatial moments may reveal further information and provide a more accurate approximation for systems where higher order structure has a greater effect.

Collective movement is a phenomena observed throughout nature, from migrating cells to swarming animals and pedestrian flow in traffic. Despite their diverse contexts, the underlying mechanisms governing this collective behaviour are often similar. In particular, neighbour-dependent directional bias could also be an important driving factor of movement in systems outside of the field of cell biology. Perhaps the closest ties lie in ecology, where the movements of animals can also generate, for example, clustered or regular patterns of individuals.

Gaining a deeper understanding of the underlying processes that drive collective cell behaviour is a crucial step towards being able to make accurate predictions about cell populations studied *in vitro*. Mathematical models can shed light on these complex systems by ascertaining which mechanisms are important for determining collective dynamics and predicting the effect that manipulating these mechanisms has on the population. Therefore, there is considerable scope for the application of modelling to *in vitro* studies that aim to either enhance or perturb collective cell movement.

Appendix A

A.1 Model Validation Using *In Vitro* Data

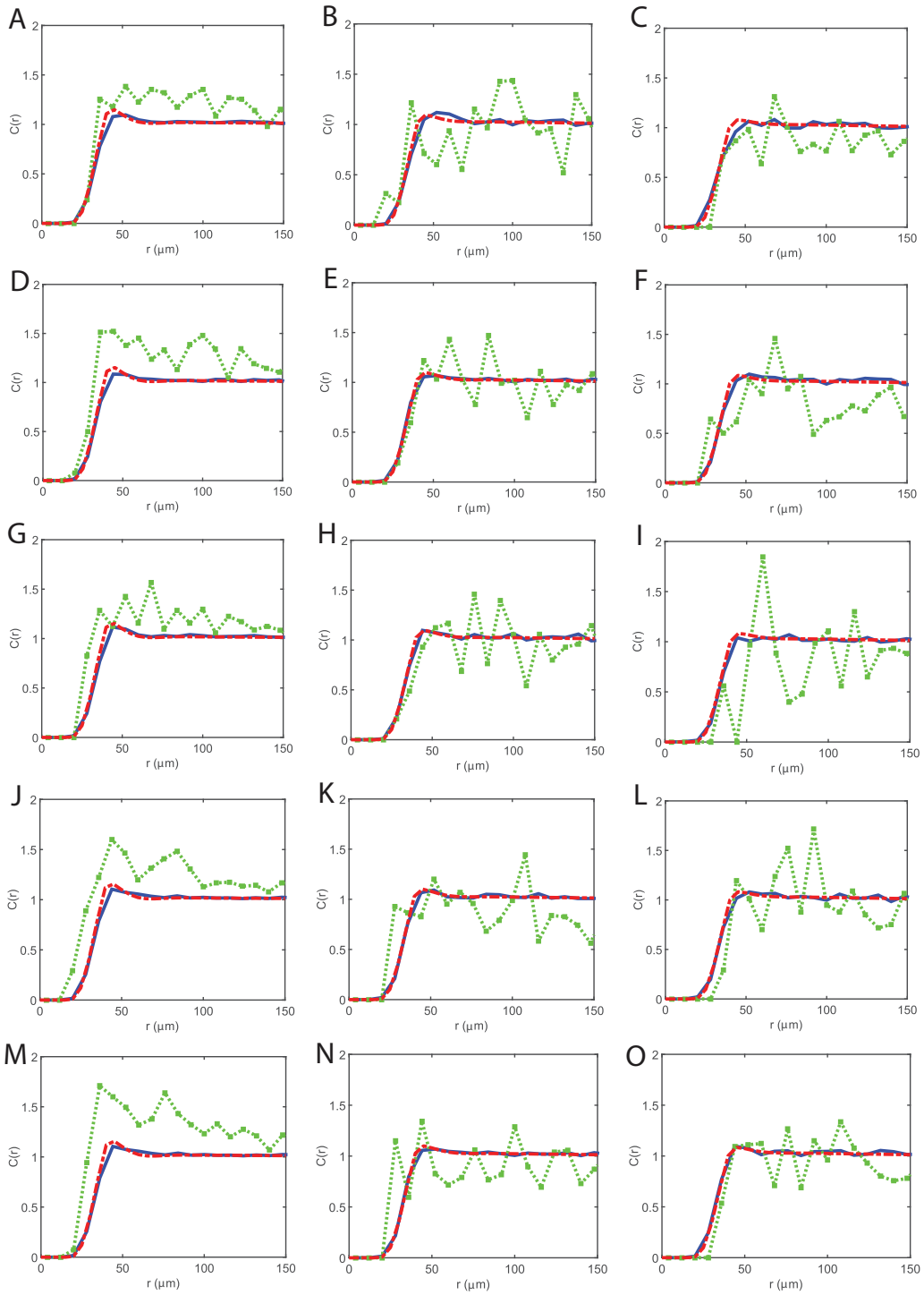


Figure A.1: Spatial structure in populations of 3T3 fibroblast cells in a 4500μm x 450μm region. PCF $C_{exp}(r)$ (green squares-dotted line) generated from an experimental image (obtained from wells containing cell suspension of approximate initial density 20,000 cells/ml), for $\delta r = 8$ μm. PCF $C_{IBM}(r)$ (blue solid line) obtained from averaging results from 200 simulations of the IBM at $t = 15$ hours. PCF $C_{SM}(r)$ (red dashed line) approximated by spatial moment model at $t = 15$ hours, for $\Delta = 5$ μm and $\xi_{max} = 150$ μm. Parameters are $\alpha = 0$ hour^{-1}, $\beta = 1000$ μm, $\sigma_w = \sigma_v = 10$ μm, $m = 5$ hour^{-1}, $\lambda_\mu = 0.1$ μm^{-1}, $\sigma_\mu = 2.5$ μm. Time point 1, wells 1-3: A) 220 cells; B) 114 cells; C) 80 cells. Time point 2, wells 1-3: D) 228 cells; E) 123 cells; F) 95 cells. Time point 3, wells 1-3: G) 229 cells; H) 118 cells; I) 90 cells. Time point 4, wells 1-3: J) 235 cells; K) 125 cells; L) 88 cells. Time point 5, wells 1-3: M) 235 cells; N) 123 cells; O) 92 cells.

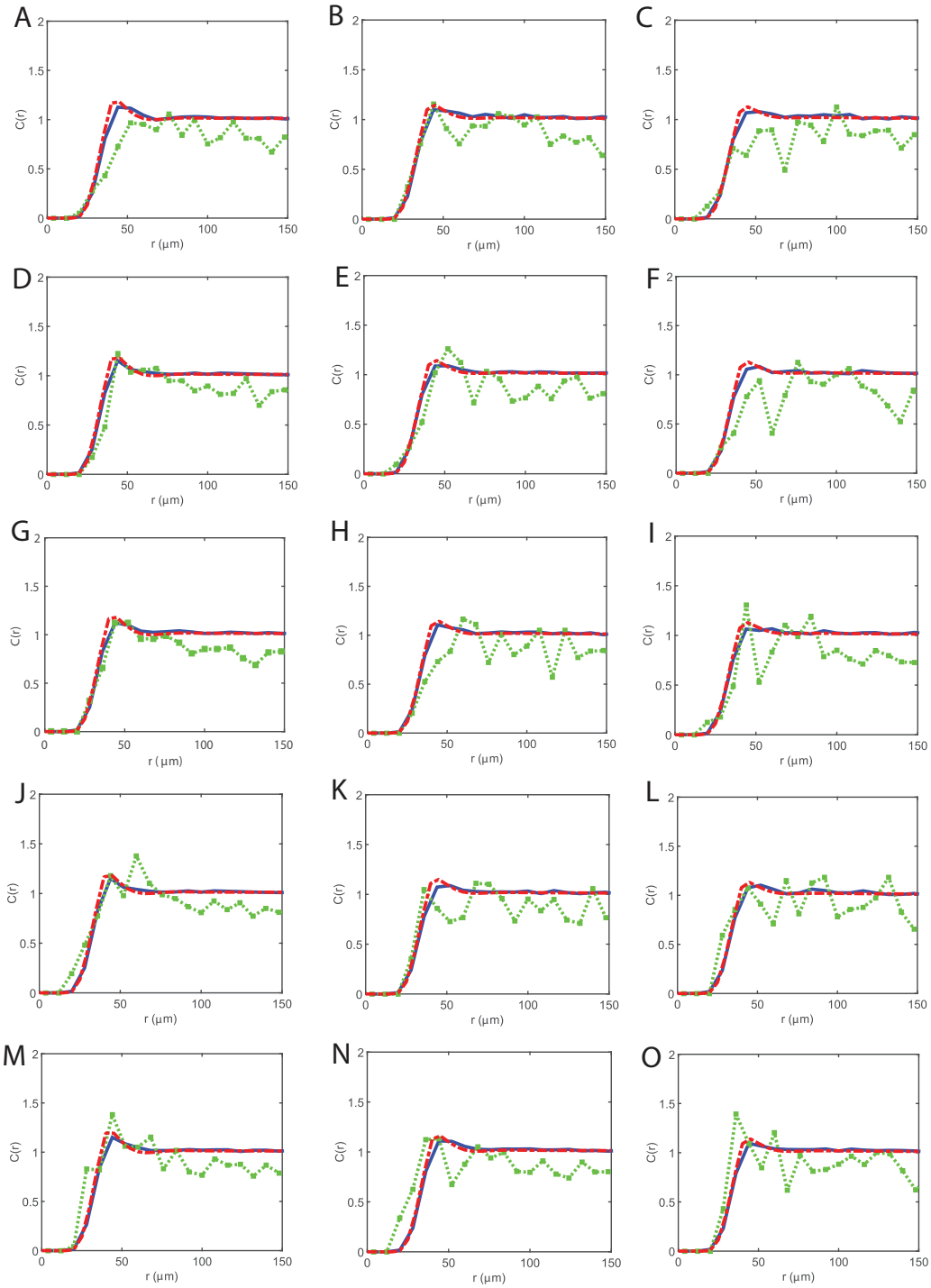


Figure A.2: Spatial structure in populations of 3T3 fibroblast cells in a 4500μm x 450μm region. PCF $C_{exp}(r)$ (green squares-dotted line) generated from an experimental image (obtained from wells containing cell suspension of approximate initial density 30,000 cells/ml), for $\delta r = 8$ μm. PCF $C_{IBM}(r)$ (blue solid line) obtained from averaging results from 200 simulations of the IBM at $t = 15$ hours. PCF $C_{SM}(r)$ (red dashed line) approximated by spatial moment model at $t = 15$ hours, for $\Delta = 5$ μm and $\xi_{max} = 150$ μm. Parameters are $\alpha = 0$ hour^{-1}, $\beta = 1000$ μm, $\sigma_w = \sigma_v = 10$ μm, $m = 5$ hour^{-1}, $\lambda_\mu = 0.1$ μm^{-1}, $\sigma_\mu = 2.5$ μm. Time point 1, wells 1-3: A) 289 cells; B) 211 cells; C) 178 cells. Time point 2, wells 1-3: D) 290 cells; E) 208 cells; F) 182 cells. Time point 3, wells 1-3: G) 286 cells; H) 207 cells; I) 180 cells. Time point 4, wells 1-3: J) 290 cells; K) 222 cells; L) 185 cells. Time point 5, wells 1-3: M) 318 cells; N) 245 cells; O) 201 cells.

References

[1] A. Tremel, A. Cai, N. Tirtaatmadja, B. D. Hughes, G. W. Stevens, K. A. Landman, and A. J. O'Connor. Cell migration and proliferation during monolayer formation and wound healing. *Chemical Engineering Science*, 64(2):247–253, January 2009. 1, 2, 7, 8, 11, 23, 69, 81, 97

[2] T. J. Shaw and P. Martin. Wound repair at a glance. *Journal of Cell Science*, 122(18):3209–13, September 2009. 1

[3] A. R. A. Anderson and M. A. J. Chaplain. Continuous and discrete mathematical models of tumor-induced angiogenesis. *Bulletin of Mathematical Biology*, 60(5):857–99, September 1998. 1, 5, 6, 9, 11

[4] H. M. Byrne, T. Alarcon, M. R. Owen, S. D. Webb, and P. K. Maini. Modelling aspects of cancer dynamics: a review. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, 364(1843):1563–78, June 2006. 1, 9, 11

[5] C. Deroulers, M. Aubert, M. Badoual, and B. Grammaticos. Modeling tumor cell migration: from microscopic to macroscopic models. *Physical Review E*, 79(3):031917, 2009. 1, 8, 9, 11

[6] R. Horwitz and D. Webb. Cell migration. *Current biology : CB*, 13(19):R756–9, September 2003. 1, 5, 6

[7] P. Friedl and K. Wolf. Tumour-cell invasion and migration: diversity and escape mechanisms. *Nature Reviews Cancer*, 3(5):362–74, May 2003. 1, 6, 103

[8] P. K. Maini, D. L. S. McElwain, and D. I. Leavesley. Traveling wave model to interpret a wound-healing cell migration assay for human peritoneal mesothelial cells. *Tissue Engineering*, 10(3-4):475–82, 2004. 1, 6, 7, 11, 103

[9] T. Alarcón, H. M. Byrne, and P. K. Maini. Towards whole-organ modelling of tumour growth. *Progress in Biophysics and Molecular Biology*, 85(2-3):451–72, 2004. 1, 2, 8, 9

[10] C. Le Clainche and M. Carlier. Regulation of actin assembly associated with protrusion and adhesion in cell migration. *Physiological Reviews*, 88:489–513, 2008. 1, 5, 17

[11] P. Rørth. Collective cell migration. *Annual Review of Cell and Developmental Biology*, 25:407–29, January 2009. 1, 5, 6, 8

[12] S. Vedel, S. Tay, D. M. Johnston, H. Bruus, and S. R. Quake. Migration of cells in a social context. *Proceedings of the National Academy of Sciences of the United States of America*, 110(1):129–34, January 2013. 2, 6, 7, 23, 64, 66, 69

[13] A. Q. Cai, K. A. Landman, and B. D. Hughes. Multi-scale modeling of a wound-healing cell migration assay. *Journal of Theoretical Biology*, 245(3):576–94, April 2007. 2, 7, 8, 9, 10, 23

[14] R. N. Binny, M. J. Plank, and A. James. Spatial moment dynamics for collective cell movement incorporating a neighbour-dependent directional bias. *Journal of The Royal Society Interface*, 12(106):20150228, April 2015. 4, 68

[15] R. N. Binny, P. Haridas, A. James, R. Law, M. J. Simpson, and M. J. Plank. Spatial structure arising from neighbour-dependent bias in collective cell movement. *PeerJ*, 2015. 4

[16] S. Kurosaka and A. Kashina. Cell biology of embryonic migration. *Birth Defects Research Part C: Embryo Today*, 84(2):102–122, 2008. 5, 6

[17] P. Martin. Wound healing–aiming for perfect skin regeneration. *Science*, 276(5309):75–81, April 1997. 5

[18] M. Reffay, L. Petitjean, S. Coscoy, E. Grasland-Mongrain, F. Amblard, A. Buguin, and P. Silberzan. Orientation and polarity in collectively migrating cell structures: statics and dynamics. *Biophysical journal*, 100(11):2566–75, June 2011. 6

[19] Philip Vitorino and Tobias Meyer. Modular control of endothelial sheet migration. *Genes & development*, 22(23):3268–81, December 2008. 6, 8

[20] X. Trepat, M. R. Wasserman, T. E. Angelini, E. Millet, D. A. Weitz, J. P. Butler, and J. J. Fredberg. Physical forces during collective cell migration. *Nature Physics*, 5(6):426–430, May 2009. 6, 55

[21] T. E. Angelini, E. Hannezo, X. Trepat, M. Marquez, J. J. Fredberg, and D. A. Weitz. Glass-like dynamics of collective cell migration. *Proceedings of the National Academy of Sciences of the United States of America*, 108(12):4714–9, March 2011. 6

[22] D. T. Tambe, C. C. Hardin, T. E. Angelini, K. Rajendran, C. Y. Park, X. Serra-Picamal, E. H. Zhou, M. H. Zaman, J. P. Butler, D. A. Weitz, J. J. Fredberg, and X. Trepat. Collective cell guidance by cooperative intercellular forces. *Nature Materials*, 10(6):469–75, June 2011. 6

[23] M. F. Ware, A. Wells, and D. A. Lauffenburger. Epidermal growth factor alters fibroblast migration speed and directional persistence reciprocally and in a matrix-dependent manner. *Journal of Cell Science*, 111(16):2423–32, August 1998. 6, 22, 64, 66

[24] A. Q. Cai, K. A. Landman, and B. D. Hughes. Modelling directional guidance and motility regulation in cell migration. *Bulletin of Mathematical Biology*, 68(1):25–52, January 2006. 6, 7, 8, 25, 55

[25] K. J. Painter and T. Hillen. Volume-filling and quorum-sensing in models for chemosensitive movement. *Canadian Applied Mathematics Quarterly*, 10(4):501–543, 2002. 6, 25, 55

[26] M. B. Miller and B. L. Bassler. Quorum sensing in bacteria. *Annual Reviews in Microbiology*, 55:165–99, 2001. 6

[27] B. Huber, K. Riedel, M. Hentzer, A. Heydorn, A. Gotschlich, M. Givskov, S. Molin, and L. Eberl. The cep quorum-sensing system of Burkholderia cepacia H111 controls biofilm formation and swarming motility. *Microbiology (Reading, England)*, 147(Pt 9):2517–28, September 2001. 6

[28] T. Gregor, K. Fujimoto, N. Masaki, and S. Sawai. The onset of collective behavior in social amoebae. *Science (New York, N.Y.)*, 328(5981):1021–5, May 2010. 6

[29] S. T. Johnston, M. J. Simpson, and M. J. Plank. Lattice-free descriptions of collective motion with crowding and adhesion. *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics*, 88(6):062720, December 2013. 6, 7, 10

[30] E. Raz and H. Mahabaleshwar. Chemokine signaling in embryonic cell migration: a fisheye view. *Development*, 136(8):1223–9, April 2009. 6

[31] J. N. Kay, M. W. Chu, and J. R. Sanes. MEGF10 and MEGF11 mediate homotypic interactions required for mosaic spacing of retinal neurons. *Nature*, 483(7390):465–9, March 2012. 7, 101

[32] M Abercrombie. Contact inhibition and malignancy. *Nature*, 281(5729):259–262, September 1979. 7, 8, 17, 23

[33] D. C. Walker, G. Hill, S. M. Wood, R. H. Smallwood, and J. Southgate. Agent-Based Computational Modeling of Wounded Epithelial Cell Monolayers. *IEEE Transactions on Nanobioscience*, 3(3):153–163, 2004. 7, 8, 9

[34] D. C. Markham, R. E. Baker, and P. K. Maini. Modelling collective cell behaviour. *Discrete and Continuous Dynamical Systems*, 34(12):5123–5133, June 2014. 7, 11

[35] J. E. F. Green, S. L. Waters, J. P. Whiteley, L. Edelstein-Keshet, K. M. Shakesheff, and H. M. Byrne. Non-local models for the formation of hepatocyte-stellate cell aggregates. *Journal of Theoretical Biology*, 267(1):106–20, November 2010. 7, 11, 69

[36] P. J. Diggle, S. J. Eglen, and J. B. Troy. Modelling the bivariate spatial distribution of amacrine cells. In A Baddelely, P Gregori, J Mateu, R Stoica, and D Stoyan, editors, *Case Studies in Spatial Point Process Modeling, Lecture Notes in Statistics*, volume 185, pages 215–233. Springer, New York, October 2006. 7, 11

[37] D. J. G. Agnew, J. E. F. Green, T. M. Brown, M. J. Simpson, and B. J. Binder. Distinguishing between mechanisms of cell aggregation using pair-correlation functions. *Journal of Theoretical Biology*, 352:16–23, July 2014. 7, 11, 13, 101

[38] B. J. Binder and M. J. Simpson. Quantifying spatial structure in experimental observations and agent-based simulations using pair-correlation functions. *Physical Review E*, 88(2):022705, August 2013. 7, 12, 49

[39] P. W. Keeley, C. Zhou, L. Lu, R.W. Williams, S. Melmed, and B. E. Reese. Pituitary tumor-transforming gene 1 regulates the patterning of retinal mosaics. *Proceedings of the National Academy of Sciences of the United States of America*, 111(25):9295–300, June 2014. 7, 69

[40] R. Van Horssen and T. L. M. ten Hagen. Crossing barriers: the new dimension of 2D cell migration assays. *Journal of Cellular Physiology*, 226(1):288–90, January 2011. 7, 8

[41] Y. Kam, C. Guess, L. Estrada, B. Weidow, and V. Quaranta. A novel circular invasion assay mimics in vivo invasive behavior of cancer cell lines and distinguishes single-cell motility in vitro. *BMC Cancer*, 8:198, January 2008. 7, 8

[42] M. J. Simpson, K. A. Landman, and B. D. Hughes. Cell invasion with proliferation mechanisms motivated by time-lapse data. *Physica A: Statistical Mechanics and its Applications*, 389(18):3779–3790, September 2010. 7, 9, 10

[43] M. J. Simpson, K. K. Treloar, B. J. Binder, P. Haridas, K. J. Manton, D. I. Leavesley, D. L. S. McElwain, and R. E. Baker. Quantifying the roles of cell motility and cell proliferation in a circular barrier assay. *Journal of the Royal Society, Interface*, 10(82):20130007, May 2013. 7, 8, 10, 11, 14

[44] M. J. Plank and M. J. Simpson. Models of collective cell behaviour with crowding effects: comparing lattice-based and lattice-free approaches. *Journal of the Royal Society, Interface*, 9(76):2983–96, November 2012. 8, 10, 73

[45] A. Valster, N. L. Tran, M. Nakada, M. E. Berens, A. Y. Chan, and M. Symons. Cell migration and invasion assays. *Methods (San Diego, Calif.)*, 37(2):208–15, October 2005. 8, 103, 104

[46] C. L. Stokes, D. A. Lauffenburger, and S. K. Williams. Migration of individual microvessel endothelial cells: stochastic model and parameter measurement. *Journal of Cell Science*, 99:419–30, June 1991. 8, 9

[47] C. M. Topaz, M. R. D'Orsogna, L. Edelstein-Keshet, and A. J. Bernoff. Locust dynamics: behavioral phase change and swarming. *PLoS Computational Biology*, 8(8):e1002642, January 2012. 8

[48] N. Shiwakoti, M. Sarvi, G. Rose, and M. Burd. Animal dynamics based approach for modeling pedestrian crowd egress under panic conditions. *Transportation Research Part B: Methodological*, 45(9):1433–1449, November 2011. 8

[49] Ruth E. Baker and Matthew J. Simpson. Models of collective cell motion for cell populations with different aspect ratio: Diffusion, proliferation and travelling waves. *Physica A: Statistical Mechanics and its Applications*, 391(14):3729–3750, July 2012. 9, 10

[50] Daniel W. Stroock. Some stochastic processes which arise from a model of the motion of a bacterium. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 28(4):305–315, 1974. 9

[51] R. L. Hall. Amoeboid movement as a correlated walk. *Journal of Mathematical Biology*, 4:327–335, 1977. 9

[52] E. A. Codling, M. J. Plank, and S. Benhamou. Random walk models in biology. *Journal of the Royal Society Interface*, 5(25):813–34, August 2008. 9

[53] S. C. Peterson and P. B. Noble. A two-dimensional random-walk analysis of human granulocyte movement. *Biophysical journal*, 12(8):1048–55, August 1972. 9

[54] W. Alt. Biased random walk models for chemotaxis and related diffusion approximations. *Journal of mathematical biology*, 9:147–177, 1980. 9

[55] G. A. Dunn and A. F. Brown. A unified approach to analysing cell motility. *Journal of Cell Science. Supplement.*, 8:81–102, 1987. 9

[56] V. Grimm, U. Berger, F. Bastiansen, S. Eliassen, V. Ginot, J. Giske, J. Goss-Custard, T. Grand, S. K. Heinz, G. Huse, A. Huth, J. U. Jepsen, C. Jø rgensen, W. M. Mooij, B. Müller, G. Peer, C. Piou, S. F. Railsback, A. M. Robbins, M. M. Robbins, E. Rossmanith, N. Rüger, E. Strand, S. Souissi, R. A. Stillman, R. Vabø, U. Visser, and D. L. DeAngelis. A standard protocol for describing individual-based and agent-based models. *Ecological Modelling*, 198(1-2):115–126, September 2006. 9

[57] M. J. Plank and M. J. Simpson. Lattice-free models of cell invasion: discrete simulations and travelling waves. *Bulletin of Mathematical Biology*, 75(11):2150–66, November 2013. 10

[58] L. Dyson, P. K. Maini, and R. E. Baker. Macroscopic limits of individual-based models for motile cell populations with volume exclusion. *Physical Review E*, 86(3):031903, September 2012. 10

[59] L. Dyson and R. E. Baker. The importance of volume exclusion in modelling cellular migration. *Journal of Mathematical Biology*, 71(3):691–711, September 2015. 10

[60] M. Bruna and S. J. Chapman. Excluded-volume effects in the diffusion of hard spheres. *Physical Review E*, 85(1):011103, January 2012. 10, 15

[61] M. J. Plank and R. Law. Spatial point processes and moment dynamics in the life sciences: a parsimonious derivation and some extensions. *Bulletin of Mathematical Biology*, 77(4):586–613, April 2015. 11, 14, 15, 27, 28, 29, 49, 79, 99, 104

[62] R. A. Fisher. The wave of advance of advantageous genes. *Annals of Eugenics*, 7(4):355–369, June 1937. 11

[63] A. N. Kolmogorov, I. G. Petrovsky, and N. S. Piskunov. Étude de léquation de la diffusion avec croissance de la quantité de matière et son application à un problème biologique. *Moscow University Mathematics Bulletin*, 1:1–25, 1937. 11

[64] E. F. Keller and L. A. Segel. Model for chemotaxis. *Journal of Theoretical Biology*, 30(2):225–34, February 1971. 11

[65] R. Law, D. J. Murrell, and U. Dieckmann. Population growth in space and time: spatial logistic equations. *Ecology*, 84(1):252–262, 2003. 11, 13, 14, 43, 50, 73, 87, 89, 90, 94, 95, 97, 98, 99, 103

[66] N. N. Bogoliubov. Kinetic Equations. *Acad Sci USSR Journal of Physics*, 10(3):265–274, 1946. 11

[67] M. Born and H. S. Green. A General Kinetic Theory of Liquids. I. The Molecular Distribution Functions. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 188(1012):10–18, December 1946. 11

[68] J. G. Kirkwood. The Statistical Mechanical Theory of Transport Processes I. General Theory. *The Journal of Chemical Physics*, 14(3):180, 1946. 11

[69] J. Yvon. *La théorie statistique des fluides et l'équation d'état, Actual. Sci. & Indust. No. 203.* Hermann, Paris, 1935. 11

[70] D. J. Murrell and R. Law. Beetles in fragmented woodlands: a formal framework for dynamics in ecological landscapes of movement. *Journal of Animal Ecology*, 69(3):471–483, 2000. 11, 13, 14, 15, 43, 50, 98

[71] A. M. Middleton, C. Fleck, and R. Grima. A continuum approximation to an off-lattice individual-cell based model of cell migration and adhesion. *Journal of Theoretical Biology*, 359:220–232, October 2014. 11, 13, 14, 50, 102

[72] J. Illian, A. Penttinen, H. Stoyan, and D. Stoyan. *Statistical analysis and modelling of spatial point patterns.* Wiley, Chichester, 2008. 12, 14, 26, 27

[73] K. K. Treloar, M. J. Simpson, B. J. Binder, D. L. S. McElwain, and R. E. Baker. Assessing the role of spatial correlations during collective cell spreading. *Scientific reports*, 4:5713, January 2014. 12

[74] B. J. Binder and M. J. Simpson. Spectral analysis of pair-correlation bandwidth: application to cell biology images. *Royal Society Open Science*, 2:140494, February 2015. 12, 26, 101

[75] S. T. Johnston, M. J. Simpson, and D. L. S. McElwain. How much information can be obtained from tracking the position of the leading edge in a scratch assay? *Journal of the Royal Society, Interface*, 11(97):20140325, August 2014. 12

[76] J. G. Kirkwood. Statistical mechanics of fluid mixtures. *The Journal of Chemical Physics*, 3(5):300–313, 1935. 13

[77] D. C. Markham, M. J. Simpson, and R. E. Baker. Simplified method for including spatial correlations in mean-field approximations. *Physical Review E*, 87(6):062702, June 2013. 13

[78] R. E. Baker and M. J. Simpson. Correcting mean-field approximations for birth-death-movement processes. *Physical Review E*, 82(4):041905, October 2010. 13, 15

[79] D. J. Murrell, U. Dieckmann, and R. Law. On moment closures for population dynamics in continuous space. *Journal of Theoretical Biology*, 229(3):421–32, August 2004. 13, 29, 35, 42, 43, 50, 59, 60, 70, 85

[80] B. Bolker and S. W. Pacala. Using moment equations to understand stochastically driven spatial pattern formation in ecological systems. *Theoretical Population Biology*, 52(3):179–97, December 1997. 13, 73, 86, 89, 90, 94, 97, 98, 99

[81] R. Law and U. Dieckmann. Moment Approximations of Individual-based Models. In U. Dieckmann, R. Law, and J. A. J. Metz, editors, *The Geometry of Ecological Interactions: Simplifying Spatial Complexity*, chapter 14, pages 252–270. Cambridge University Press, Cambridge, 2000. 13

[82] M. A. Lewis. Spread rate for a nonlinear stochastic invasion. *Journal of Mathematical Biology*, 41(5):430–454, November 2000. 13, 15, 73

[83] M. A. Lewis and S. Pacala. Modeling and analysis of stochastic invasion processes. *Journal of Mathematical Biology*, 41(5):387–429, November 2000. 13, 14, 15

[84] T. P. Adams, E. P. Holland, R. Law, M. J. Plank, and M. Raghib. On the growth of locally interacting plants: differential equations for the dynamics of spatial moments. *Ecology*, 94(12):2732–2743, December 2013. 14

[85] U. Dieckmann and R. Law. Relaxation projections and the method of moments. In U. Dieckmann, R. Law, and J.A.J Metz, editors, *The Geometry of Ecological Interactions: Simplifying Spatial Complexity*, chapter 21, pages 412–455. Cambridge University Press, Cambridge, 2000. 14, 60, 70, 73, 99

[86] D. J. Murrell. Local spatial structure and predator-prey dynamics: counterintuitive effects of prey enrichment. *The American Naturalist*, 166(3):354–67, September 2005. 14, 73, 98, 99

[87] R. Law, J. Illian, D. F. R. P. Burslem, G. Gratzer, C. V. S. Gunatilleke, and I. A. U. N. Gunatilleke. Ecological information from spatial patterns of plants: insights from point process theory. *Journal of Ecology*, 97(4):616–628, July 2009. 14, 69

[88] M. J. Simpson and R. E. Baker. Corrected mean-field models for spatially dependent advection-diffusion-reaction phenomena. *Physical Review E*, 83(5):051922, May 2011. 14, 15

[89] T. J. Newman and R. Grima. Many-body theory of chemotactic cell-cell interactions. *Physical Review E*, 70(5):051916, November 2004. 15, 102

[90] F. Cooper, G. Ghoshal, and J. Pérez-Mercader. Composite bound states and broken U(1) symmetry in the chemical-master-equation derivation of the Gray-Scott model. *Physical Review E*, 88(4):042926, October 2013. 15

[91] J. Blath, A. Etheridge, and M. Meredith. Coexistence in locally regulated competing populations and survival of branching annihilating random walk. *The Annals of Applied Probability*, 17(5/6):1474–1507, October 2007. 15

[92] O. Ovaskainen, D. Finkelshtein, O. Kutoviy, S. Cornell, B. Bolker, and Y. Kondratiev. A general mathematical framework for the analysis of spatiotemporal point processes. *Theoretical Ecology*, 7(1):101–113, October 2014. 15

[93] Howard M. Taylor and Samuel Karlin. *An introduction to stochastic modeling*. Academic Press, Orlando, third edition, 1998. 19

[94] D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, 81(25):2340–2361, December 1977. 20

[95] I. Barbaric, V. Biga, P. J. Gokhale, M. Jones, D. Stavish, A. Glen, D. Coca, and P. W. Andrews. Time-Lapse Analysis of Human Embryonic Stem Cells Reveals Multiple Bottlenecks Restricting Colony Formation and Their Relief upon Culture Adaptation. *Stem Cell Reports*, 3(1):142–55, July 2014. 22

[96] M. Raghib, N. A. Hill, and U. Dieckmann. A multiscale maximum entropy moment closure for locally regulated space-time point process models of population dynamics. *Journal of Mathematical Biology*, 62(5):605–53, May 2011. 42, 60, 70, 86

[97] J. R. Dormand and P. J. Prince. A family of embedded Runge-Kutta formulae. *Journal of Computational and Applied Mathematics*, 6(1):19–26, March 1980. 46

[98] P. Bogacki and L. F. Shampine. A 3(2) pair of Runge - Kutta formulas. *Applied Mathematics Letters*, 2(4):321–325, 1989. 58

[99] M. J. Simpson, B. J. Binder, P. Haridas, B. K. Wood, K. K. Treloar, D. L. S. McElwain, and R. E. Baker. Experimental and modelling investigation of monolayer development with clustering. *Bulletin of Mathematical Biology*, 75(5):871–89, May 2013. 64

[100] Y. Lee, L. V. McIntire, and K. Zygourakis. Analysis of endothelial cell locomotion: Differential effects of motility and contact inhibition. *Biotechnology and Bioengineering*, 43(7):622–34, March 1994. 69

[101] P. Haase. Spatial pattern analysis in ecology based on Ripley ' s K-function : Introduction and methods of edge correction. *Journal of Vegetation Science*, 6(4):575–582, 1995. 69, 70

[102] E. M. Levine, Y. Becker, C. W. Boone, and H. Eagle. Contact inhibition, macromolecular synthesis, and polyribosomes in cultured human diploid fibroblasts. *Proceedings of the National Academy of Sciences USA*, 53(2):350–356, February 1965. 73

[103] F. Fagotto and B. M. Gumbiner. Cell Contact-Dependent Signaling. *Developmental Biology*, 180(2):445–454, 1996. 73

[104] N.-G. Kim, E. Koh, X. Chen, and B. M. Gumbiner. E-cadherin mediates contact inhibition of proliferation through Hippo signaling-pathway components. *Proceedings of the National Academy of Sciences of the United States of America*, 108(29):11930–11935, 2011. 73

[105] B. M. Bolker and S. W. Pacala. Spatial moment equations for plant competition: understanding spatial strategies and the advantages of short dispersal. *The American Naturalist*, 153(6):575–602, June 1999. 98

[106] T. Alarcón, H. M. Byrne, and P. K. Maini. A mathematical model of the effects of hypoxia on the cell-cycle of normal and cancer cells. *Journal of Theoretical Biology*, 229(3):395–411, 2004. 99

[107] M. J. Simpson, P. Haridas, and D. L. S. McElwain. Do pioneer cells exist? *PloS One*, 9(1):e85488, January 2014. 104