Models of Cranial Neural Crest Cell Migration

Louise Dyson

Balliol College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Hilary 2013

Acknowledgements

I would like to express my deep gratitude to my supervisors, Prof. Philip Maini and Dr. Ruth Baker for introducing me to the world of Mathematical Biology, sharing the highs and lows of research and their continued help and encouragement. I am also grateful to Prof. Paul Kulesa for his kind hospitality at the Stowers Institute for Medical Research in Kansas City and to Dr. Rebecca McLennan for taking the time to explain experimental biology to a mathematician. Many thanks also to all the scientists I met at the Stowers institute for sharing many interesting conversations and exciting science.

Thank you to everyone in the Centre for Mathematical Biology and those outside it that have provided support and copious amounts of tea and food. Particularly to Suruchi, Gabs, Aaron and Jess for forming near-continual tea-trains, and to Becky, Owen, Katy, Lyndsey and Freya for their unfailing support and encouragement. Also to Michael, for always being there with a friendly smile. Finally, to my family, who made so much possible with their love and support.

This project was funded by the Engineering and Physical Sciences Research Council through the Systems Biology Doctoral Training Centre.

Abstract

From the developing embryo to the evacuation of football stadiums, the migration and movement of populations of individuals is a vital part of human life. Such movement often occurs in crowded conditions, where the space occupied by each individual impacts on the freedom of others. This thesis aims to analyse and understand the effects of occupied volume (volume exclusion) on the movement of the individual and the population.

We consider, as a motivating system, the rearrangement of individuals required to turn a clump of cells into a functioning embryo. Specifically, we consider the migration of cranial neural crest cells in the developing chick embryo. Working closely with experimental collaborators we construct a hybrid model of the system, consisting of a continuum chemoattractant and individual-based cell description and find that multiple cell phenotypes are required for successful migration. In the crowded environment of the migratory system, volume exclusion is highly important and significantly enhances the speed of cell migration in our model, whilst reducing the numbers of individuals that can enter the domain. The developed model is used to make experimental predictions, that are tested *in vivo*, using cycles of modelling and experimental work to give greater insight into the biological system.

Our formulated model is computational, and is thus difficult to analyse whilst considering different parameter regimes. The second part of the thesis is driven by the wish to systematically analyse our model. As such, it concentrates on developing new techniques to derive continuum equations from diffusive and chemotactic individual-based and hybrid models in one and two spatial dimensions with the incorporation of volume exclusion. We demonstrate the accuracy of our techniques under different parameter regimes and using different mechanisms of movement. In particular, we show that our derived continuum equations almost always compare better to data averaged over multiple simulations than the equivalent equations without volume exclusion. Thus we establish that volume exclusion has a substantial effect on the evolution of a migrating population.

Contents

Chapter 1

Introduction

The migration of cell populations is crucial in the development and homeostasis of adult organisms, allowing the formation of highly structured organs and protecting from infection and disease through a targeted immune system and inflammatory response (Luster, 2001; Alt and Lauffenburger, 1987). Consequently, failures of migration or inappropriate cell movement can result in severe developmental disorders, chronic inflammatory syndromes, such as asthma and rheumatoid arthritis, and the metastasis of cancerous growths (Roussos *et al.*, 2011; Trainor, 2005; Johnston and Bronsky, 1995; Tallquist and Soriano, 2003; Rizzoti and Lovell-Badge, 2007; Theveneau and Mayor, 2011). One system that is widely studied is the directed migration of the neural crest (Johnston and Bronsky, 1995; Kulesa *et al.*, 2010). Since cellular migration is well conserved across and within organisms (Kulesa *et al.*, 2004), this provides a model system to investigate the migration of cell populations. We will take the migration of the cranial neural crest as an example system to study cell migration in detail by undertaking mathematical modelling in close collaboration with experimental biologists.

In this chapter we give a brief introduction to the biological system and review the current state of the field in terms of experimental understanding, previous mathematical models and the techniques currently used to approximate individual-based models (IBMs) using continuum equations.

1.1 Biological background

Just as any journey must begin with a single step, so the development of any organism must begin with a single cell. The first step towards a fully mature, multicellular organism is the division of

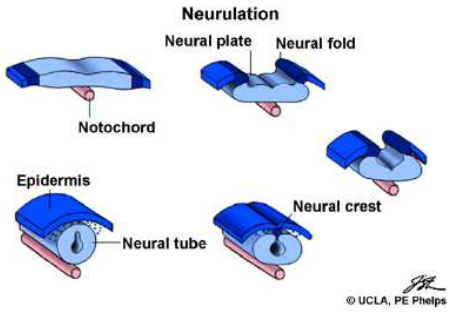


Figure 1.1: A schematic representation of neurulation, starting from the top left, and proceeding clockwise. Reproduced from Spicher and Michel (2007), with permission from Elsevier.

this initial zygote into a ball of cells (blastula). This division is so rapid that cell growth is almost completely neglected during this phase. The blastula then undergoes a period of slower division, during which the cells grow and become motile, enabling a highly coordinated rearrangement of cells (gastrulation) to take place. This rearrangement of cells leads to the establishment of three germ layers; the ectoderm (gives rise to the nervous system and epidermis), the endoderm (produces the lining of the digestive tube and respiratory tube) and the mesoderm (forms bones, connective tissue, circulatory system and reproductive system). A thickening of the ectoderm, called the neural plate, then undergoes neurulation, folding inwards to make the neural tube (Figure 1.1). This will become the central nervous system for the organism, including the brain, spinal cord and retina. Other cells within the ectoderm become neural crest cells (NCCs), subsequently migrating to specific areas within the embryo and differentiating into a prodigious number of different types of cells. NCCs are divided by their spatial position in the neural crest into four main groups (which may overlap): cranial; trunk; vagal and cardiac neural crests. NCCs from different groups will then migrate to various regions in the embryo and differentiate into distinct cell types (Gilbert, 2006).

There has previously been a particular focus of research on enteric neural crest cells (ENCCs) (Landman *et al.*, 2003; Simpson *et al.*, 2006a; Zhang *et al.*, 2010), which derive from the trunk neural crest, but cranial neural crest cells (CNCCs) are now becoming more well-studied (McLennan and Kulesa, 2007; Carmona-Fontaine *et al.*, 2008; Kulesa *et al.*, 2010). The head is largely the product of CNCCs, and many craniofacial abnormalities, such as DiGeorge syndrome (related to maternal alcohol use; Trainor, 2005; Johnston and Bronsky, 1995; Tallquist and Soriano, 2003; Rizzoti and Lovell-Badge, 2007; Theveneau and Mayor, 2011) and the development of a cleft palate (Tallquist and Soriano, 2003; Eberhart *et al.*, 2008; Trainor, 2005; Johnston and Bronsky, 1995; Latta and Golding, 2012; Theveneau and Mayor, 2011, 2012; Olesnicky Killian *et al.*, 2009), may be the result

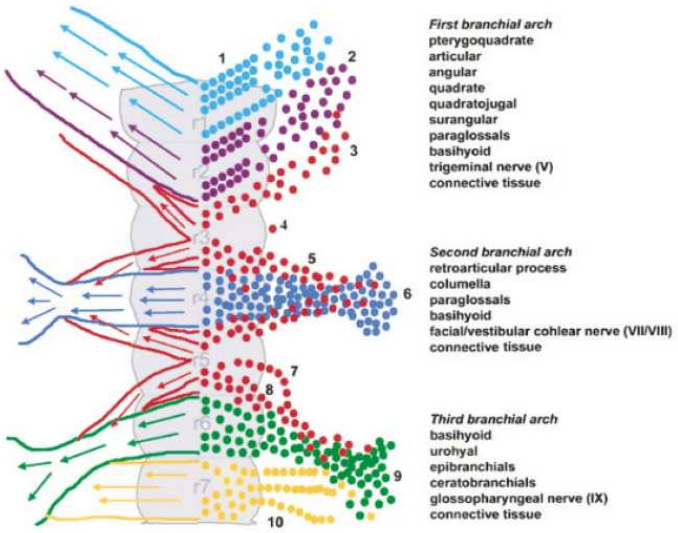


Figure 1.2: A schematic showing the migration routes of CNCCs from different rhombomeres (labelled r1-r7). Each circle represents a neural crest cell, and colours represent the rhombomere of origin; r1 (turquoise), r2 (purple), r4 (blue), r6 (green) and r7 (yellow). NCCs originating from r3 or r5 are shown in red. Different types of behaviour are represented and numbered here; NCCs from r1 and r2 migrate diagonally (1,2), often forming chain-like arrays which then disassemble leaving the cells free to migrate as individuals. Cells from r3 will either join the r2 stream (3) or the r4 stream (5,6), or stop and retract filopodia (4). Similarly, cells from r6 and r7 migrate along specific routes (9), with NCCs from r7 forming chains (10) and cells from r6 being joined by cells from r5 (7,8). Reproduced from Kulesa *et al.* (2004), with permission from Wiley Periodicals Inc.

of a failure of CNCC migration. The cranial neural crest is therefore highly important for normal embryo development. This thesis will focus on the mechanisms underlying CNCC migration, a subject that is further complicated by the continual growth of the embryo, expanding the domain of migration.

The neural tube in the region of the developing head is divided into three regions; the forebrain, the midbrain and the hindbrain. CNCCs originate in the hindbrain (rhombencephalon), which has a characteristic patterning of periodic swellings called rhombomeres (r1-r7). The developmental fate and migration route of the CNCCs is dictated by the rhombomere of origin. Streams of cells are found laterally to the even-numbered rhombomeres, whereas CNCCs originating from odd-numbered rhombomeres will typically join an adjacent stream, leaving an exclusion zone next to the odd-numbered rhombomeres (Figure 1.2). CNCCs will typically migrate to and colonise one of the pharyngeal arches (also known as branchial arches), giving rise to bones, cartilage and other connective tissue.

1.2 Experimental background

We wish to investigate the r4 migratory stream (see Figure 1.2) bound for the second branchial arch (ba2) in the developing chick embryo. This system is studied experimentally by our collaborators at the Stowers Institute for Medical Research (Kansas City), allowing experimental work to be undertaken in order to test modelling predictions. However, it should also be noted that NCC patterning is highly conserved among vertebrates (Pratt *et al.*, 1975; Johnston and Bronsky, 1995; Kulesa *et al.*, 2004), so that the information gathered in other organisms may be relevant to chick embryos and this, in turn, has implications for human development.

There are several experimental review papers in the field of NCC migration, including two that focus on the cranial crest (Noden, 1993; Kulesa *et al.*, 2010). The earlier of these hypothesised sequential, hierarchical interactions between the different populations of cells (*e.g.* between the mesenchymal cells, the CNCCs and the underlying mesodermal cells). However, the much more recent paper by Kulesa *et al.* (2010), whilst accepting the importance of the general microenvironment, asserts that cell-cell interactions within the CNCC population is also a contributing factor to the successful migration and spatial segmentation of cells. Three main phases of migration are identified by Kulesa *et al.* (2010): the initial emergence of cells from the neural tube and acquisition of directed migration towards the branchial arches; the successful migration of the majority of cells to the destination; and the entry into, and colonisation of, the branchial arch. This thesis will be primarily concerned with the second of these phases.

1.2.1 Anatomy of a migratory stream

CNCCs in the r4 migratory stream acquire directionality soon after leaving the neural tube (delamination), and migrate in a narrowing band towards ba2. When cells are close to their destination the stream widens before the final invasion and population of the arch. In total, cells migrate a distance of around 1100μm taking approximately 24 hours to reach their destination.

The cranial neural crest migratory stream is a heterogeneous population, with the front 30% of cells displaying different characteristics to the remaining 70% of the stream. In particular, cells leading the migratory stream have distributed filopodia (Kulesa *et al.*, 2004) and wandering trajectories (Kulesa *et al.*, 2008). In contrast, later emerging cells are bipolar, with filopodia extending along the direction of migration (Teddy and Kulesa, 2004). Trailing cells, particularly those from the r7

migratory stream (see Figure 1.2), are observed to move in a 'follow-the-leader' manner, forming long chains of cells with filopodia aligned between the cells. Moreover, cells that exit the neural tube early on in migration remain at the wavefront and colonise the distal portion of ba2, so that the order of exiting the neural tube is preserved throughout migration (Kulesa *et al.*, 2008). Hence the cells retain similar morphologies throughout migration. The differences between the front 30% and the back 70% of cells led to the hypothesis that there are two sub-populations: leading and trailing cells (Kulesa *et al.*, 2008).

Kulesa *et al.* (2008) also found that leading cells in the migratory stream have a higher proliferation rate than trailing cells. However, this result was obtained by assuming a cell doubling growth curve ($N(t) = N(0)2^{kt}$) and fitting to two data points. In contrast, when proliferation is examined throughout migration, it is found that cells do not divide frequently *en route*, but instead multiply quickly after migration is complete (Kulesa, pers. comm.).

The migratory system has been shown to be robust to environmental perturbations. In particular, if a barrier is inserted in front of the migratory stream then the lead cells become trapped, but cells behind the front may then migrate around the barrier to become the new leaders of the stream. Indeed, cells may even move through the repulsive regions lateral to the migratory pathway in order to move past the barrier (Kulesa *et al.*, 2005).

1.2.2 Cell interactions

During migration, cell-cell interactions are facilitated via lamellipodia or short, thin filopodia for local cell contacts, while long range filopodia enable contact with more distant cells. Cell contacts may range from between neighbouring individuals to cells up to 70μm away (Teddy and Kulesa, 2004). It has been suggested that sequential interactions between different types of cells (*e.g.* between the mesenchymal cells, the CNCCs and the underlying mesodermal cells) may be used to achieve the distinctive characteristics of the migratory stream (Noden, 1993). However, recent evidence implies that cell-cell contact within the CNCC population may be more important than these inter-population interactions (Kulesa *et al.*, 2010). β-catenin is a possible mediator of cell-cell adhesion and interactions, and overstimulation of this protein completely inhibits neural crest cell migration (de Melker *et al.*, 2004).

Cells may also interact indirectly by degrading the extracellular matrix (ECM), thus allowing subsequent cells to pass more freely through the environment. In particular, inhibiting matrix metal-

loproteinases production by CNCCs dramatically reduced cell delamination and migration, whilst injecting excess metalloproteinases into embryos resulted in premature and accelerated migration (Monsonego-Ornan *et al.*, 2012). The importance of cell-cell interactions and ECM degradation has also been indicated in other species (Nie *et al.*, 2009; Carmona-Fontaine *et al.*, 2008).

1.2.3 Chemotaxis

In addition to information exchange through cell-cell interactions, there is also increasing evidence to support a chemotactic mechanism for invasion. Vascular endothelial growth factor (VEGF) is known to be a chemoattractant both *in vitro* and *in vivo*, since cells will migrate towards VEGF-soaked beads or VEGF-producing cells that are inserted laterally to the migratory route (McLennan and Kulesa, 2007). It has been found that VEGF is produced by the ectoderm overlying the CNCC migratory pathway (see Figure 1.3), and that this production is fairly constant throughout migration (McLennan *et al.*, 2010). In addition, if endogenous VEGF is removed from the system, CNCCs fail to migrate to invade ba2 successfully (McLennan *et al.*, 2010). Alternative chemoattractants have been discovered in other species (Olesnicky Killian *et al.*, 2009; Kubota and Ito, 2000; Eberhart *et al.*, 2008) and other migratory streams in chick embryos (Rupp and Kulesa, 2007). Interestingly, whilst Ephs and ephrins maintain the excluded zones between migratory streams in most studied vertebrates, the roles of ephrin-B1 and ephrin-B2 are reversed in chick embryos (Mellott and Burke, 2008), indicating that not all mechanisms are perfectly preserved between species.

In chicks, the pathway used for chemotactic sensing of VEGF begins with the membrane-bound coreceptor for vascular endothelial growth factor (VEGF), neuropilin 1. When neuropilin 1 expression is reduced in cells or endogenous VEGF is removed, cells fail to fully invade ba2. Cells with reduced neuropilin 1 also display a morphology that is consistent with a loss of polarity and mobility (McLennan and Kulesa, 2010). However, when cells with reduced neuropilin 1 are transplanted into a younger host embryo, the transplanted cells are able to recover their normal migratory ability, indicating that the effect is reversible (McLennan and Kulesa, 2007, 2010).

1.3 Modelling of the neural crest

To date there have been only two mathematical models focussing specifically on the migration of the cranial neural crest (Kulesa *et al.*, 2008; Wynn *et al.*, 2012; see Section 1.4.1). In addition to

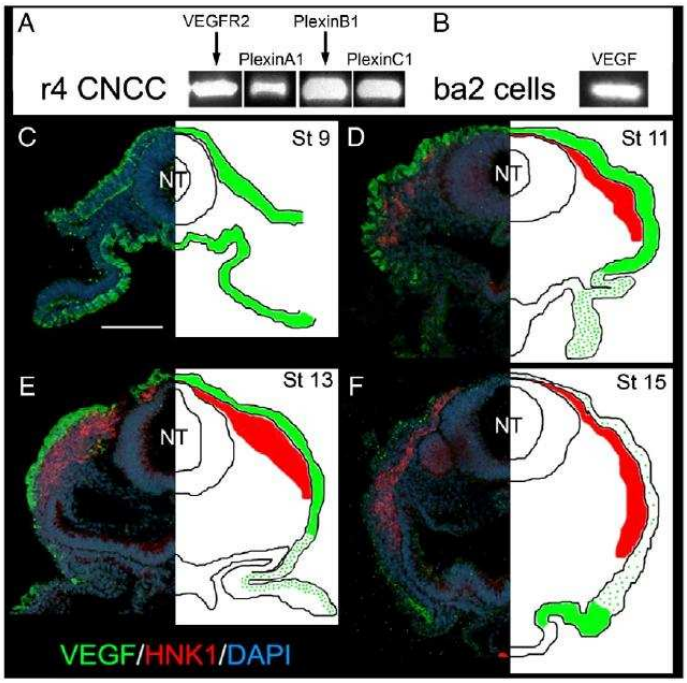


Figure 1.3: Expression of VEGF at different developmental stages of a chick embryo (shown in green) is shown to be constant throughout the migratory pathway and is produced by the ectoderm overlying the CNCC migratory pathway. Reproduced from McLennan *et al.* (2010), with permission from Elsevier.

these models, there has been a large amount of theoretical modelling work undertaken examining the migration of ENCCs (Landman *et al.*, 2003, 2007; Simpson *et al.*, 2006a,b, 2007b; Trewenack and Landman, 2009; Zhang *et al.*, 2010). Here cells colonise the developing intestine during embryogenesis, subsequently proliferating and differentiating to form the gut nervous system. The system is similar to CNCC migration in that it involves the movement of cells through a growing domain. However the migratory domain is different, since ENCCs migrate along the walls of the intestine, a lengthening tube. In addition, cell proliferation is thought to be very important in the ENCC colonisation of the intestine, in contrast to the cranial neural crest, where proliferation occurs after migration. Bearing this in mind, it is still useful to consider models focusing on ENCC migration since similar mathematical tools and techniques may apply in both situations.

1.3.1 Continuum modelling

Three main partial differential equation (PDE) models of ENCC migration have been developed and expanded upon by Landman *et al.* (2003); Simpson *et al.* (2006a) and Trewenack and Landman (2009).

1.3.1.1 Landman *et al.* (2003)

In this one-dimensional model migrating cells (n) and chemoattractant (g) diffuse and grow logistically on a uniformly growing domain with velocity, v. Cells are assumed to move in response to a chemoattractant and are removed by apoptosis or conversion to a cell type that is not considered in the model. The chemoattractant is internalised by the cells, thus creating a gradient that leads to the movement of cells into regions not previously populated. This leads to the non-dimensional system:

$$\frac{\partial n}{\partial t} = \frac{\partial}{\partial x}\left(D_n \frac{\partial n}{\partial x}\right) - \chi \frac{\partial}{\partial x}\left(n \frac{\partial g}{\partial x}\right) - \frac{\partial}{\partial x}(vn) + n(1-n) - \lambda_2 n, \tag{1.1}$$

$$\frac{\partial g}{\partial t} = \frac{\partial}{\partial x}\left(D_g \frac{\partial g}{\partial x}\right) - \frac{\partial}{\partial x}(vg) + \lambda_3 g(1-g) - \lambda_4 n g, \tag{1.2}$$

where D_n and D_g are the diffusivities of n and g, respectively, χ is a parameter relating to chemotaxis and λ_2, λ_3 and λ_4 are parameters relating to apoptosis, proliferation and internalisation of chemoattractant, respectively. These equations are considered on the domain $0 < x < L(t)$, with zero flux boundary conditions. Different types of uniform domain growth are considered, including linear, exponential and logistic growth and solutions are found numerically (the numerical method used in Landman *et al.*, 2003 is discussed in Simpson *et al.*, 2006b). The presence and type of domain growth was found to have a large influence on the success and form of cell colonisation of the domain, strongly suggesting that domain growth should be included when modelling migration in a growing embryo. However, whilst chemotaxis is included in the model, cell proliferation is also very important, since the cell profiles found are at carrying capacity behind the wavefront. Cell proliferation plays an important role in the enteric neural crest but it is unclear how these mechanisms would work in the less proliferative CNCCs.

1.3.1.2 Simpson *et al.* (2006a)

Before the widespread use of green fluorescent protein (GFP) to label cells during migration, chick-quail transplants were used to create a population of distinct cells which could then be followed as migration progressed. By transplanting the donor cells into different regions of the host domain the behaviour of different parts of the stream may be examined (Simpson *et al.*, 2007b). This system has been modelled mathematically with models consisting of two populations of cells: donor cells (D) and host cells (H), with the option of also including a glial derived neurotrophic factor

(GDNF), which acts as a chemoattractant for both the host and donor cells (Simpson *et al.*, 2006a, 2007b; Landman *et al.*, 2007). The cells are assumed to move according to some flux term and proliferate logistically with a combined carrying capacity. Different forms for the flux term are investigated, including chemotaxis and chemokinesis towards the GDNF, and linear and nonlinear diffusion. However, whilst the form of the wavefront of invasion changes depending on the flux used, the qualitative results for the different forms are very similar. Thus the system is highly robust to different forms of migration. The model is represented by

$$\frac{\partial D}{\partial t} + \frac{\partial J_D}{\partial x} = \lambda_D D\left(1 - \frac{D+H}{K}\right), \tag{1.3}$$

$$\frac{\partial H}{\partial t} + \frac{\partial J_H}{\partial x} = \lambda_H H\left(1 - \frac{D+H}{K}\right), \tag{1.4}$$

$$\frac{\partial G}{\partial t} = F(D, H, G), \tag{1.5}$$

where J_D and J_H are flux terms and λ_D and λ_H give the rate of proliferation of the donor and host cells, respectively. $F(D, H, G)$ gives the net rate of increase of GDNF as time progresses. Simulating this model with different initial distributions of cells and zero flux boundary conditions gives an insight into different areas of the migratory stream. The model is essentially a modified Fisher equation and predicts that invasion occurs due to the proliferation of cells at the wavefront and movement into the unoccupied area ahead of the wave (Simpson *et al.*, 2006a). Since this is precisely the mechanism by which Fisher's equation generates an invasive wave, this is perhaps unsurprising. Domain growth may also be considered within this modelling framework (Landman *et al.*, 2007).

Whilst this system is driven by proliferation, which is less important in the cranial neural crest, the use of experimental and modelling techniques to explore the behaviour of subsections of the migratory population may be useful in investigating our biological system. In particular, with the use of green fluorescent protein (GFP) it is now possible to follow different subsections of the migratory stream and determine their proliferative and migratory characteristics (Kulesa *et al.*, 2008). Explicitly considering separate subsections of the stream as different cell populations may therefore be an important modelling technique.

1.3.1.3 Trewenack and Landman (2009)

A third way to consider ENCC migration is as an initial wave of migrating and proliferating cells (precursor cells, c) which subsequently differentiate into relatively non-migratory and non-proliferating cells (differentiated cells, n). This system is modelled as two distinct populations, where the precursor cells diffuse, proliferate and may also become differentiated cells, giving the non-dimensional equations:

$$\frac{\partial c}{\partial t} = \frac{\partial^2 c}{\partial x^2} + c(1 - c - \gamma n) - \lambda c(1 - Kn), \tag{1.6}$$

$$\frac{\partial n}{\partial t} = \lambda c(1 - Kn). \tag{1.7}$$

Here, K is the ratio of the carrying capacities, λ is the ratio of the differentiation and proliferation rates and γ represents the reduced contribution that the differentiated cells make to the carrying capacity of the domain (Trewenack and Landman, 2009). Much analysis may be performed on this system, including predicting the wavespeed through stability analysis, examining the dependence of the wavespeed on initial conditions and performing perturbation analysis for the situation in which λ is small. Since perturbing system parameters in this model is found to disrupt travelling wave solutions or reduce their wavespeed, it is concluded that invasive waves *in vivo* may be interrupted or stopped completely by a change in parameters.

This model again uses a Fisher-type migratory mechanism, which generates an invasive wavefront by assuming that cells proliferate up to the tissue carrying capacity. Whilst this mechanism may be justified in the enteric neural crest, it is known that CNCCs proliferate very little during migration (Kulesa, pers. comm.).

1.4 Individual-based modelling

IBMs of cellular migration may be characterised as either on- or off-lattice. On-lattice models consist of a grid of points (the lattice), and each individual may occupy one or more of the points. Most models of this kind use a square-based lattice, with each lattice site containing many individuals. In contrast, individuals in an off-lattice model are defined by a continuous coordinate, which may be anywhere in the domain. An on-lattice model may be thought of as a coarse-grained equivalent of an off-lattice model, where each individual occupies the lattice site closest to it. Thus finer spatial

details are neglected but, particularly for larger numbers of individuals, simulations are faster, since multiple individuals at a given lattice site are equivalent and smaller movements that remain close to a lattice site are not explicitly simulated.

1.4.1 Individual-based modelling of neural crest migration

There has been one off-lattice model of CNCC migration to date (Kulesa *et al.*, 2008). This IBM imposed a 'gradient' function for the amount of an unspecified chemoattractant at each point in the considered domain. Cells survey the local environment in a stochastic manner and choose which direction to go in based on the 'gradient' function. Two subpopulations were considered: leading and trailing cells. It was found that if the leading cells were more efficient (*i.e.* if they sampled the environment more frequently) then the two distinct subpopulations were maintained. However, if noise was added to the 'gradient' function then the cells migrate less far into the domain and comprise less distinct subpopulations (Kulesa *et al.*, 2008). However, this model gives no insight into how such a gradient may arise *in vivo* or how the dramatic growth of the migratory domain may affect the behaviour of the migratory stream.

1.4.1.1 Wynn *et al.* (2012)

A subset of IBMs explicitly take into account the volume occupied by an individual (volume exclusion), so that if a lattice site or area of the domain is occupied by an individual then no other agents may also occupy that site. Volume exclusion enables direct investigation into the effects of population crowding and cell-cell interactions. Wynn *et al.* (2012) investigated CNCC migration using on-lattice modelling with volume exclusion to study the leader-follower chains seen in some migratory streams. At each time step each individual was given the chance to move in a direction picked at random from a set of possibilities. Two subpopulations of cells, leaders and followers, were used in the model, differing only in the number of directions from which they choose each movement. In addition, degradation of the ECM was modelled by defining lattice sites as either 'open' or 'closed', with cell movement biased towards 'open' sites. To investigate the circumstances under which chains of migrating cells are seen experimentally, the authors found the mean persistence time of chains of cells. It was found that chains persisted better when leading and trailing cells had a high degree of directionality, and that a prelaid path of degraded ECM was not enough to drive chain persistence but that it could enhance it if other mechanisms (such as chemotaxis or biased

movement) were also present. An extension to the model included cell filopodia, with each cell in contact with some other lattice sites through their filopodia and trailing cells assumed to have fewer filopodia than leading cells. The presence of individuals at other lattice sites then modified the probabilities of attempted movements in those directions. The inclusion of cell contact mechanisms allowed much greater persistence of cell chains, even when the directional bias of trailing cells was reduced. Thus Wynn *et al.* (2012) demonstrated that cell-cell contact can enhance the migration of trailing cells in chains and showed that ECM degradation was not sufficient to maintain chain migration without cell-cell contact or high directionality.

Thus Wynn *et al.* (2012) present evidence that the chains of cells observed in some migratory crest cell populations require cell-cell communication, and not simply an indirect transfer of information through the migratory environment (specifically, the ECM). However, the mechanisms by which cells may establish directional bias are not discussed and the choice of an on-lattice modelling regime may affect their results. In particular, constraining cells to move only horizontally or vertically may increase the likelihood that individuals will migrate directly behind each other, and thus be considered to be in a chain of cells.

1.4.1.2 Simpson *et al.* (2007a)

The migration of the enteric neural crest has been studied using an on-lattice exclusion process, where cells move and proliferate on a two-dimensional grid. The effects of cell crowding are explicitly included by ensuring that each occupied lattice site may only have a set number of neighbouring cells (the carrying capacity). If movement or proliferation would leave the cell above carrying capacity then the operation is aborted. Initially all the cells to the left of the origin are occupied, so that an invasive wave is generated to the right. The wavespeed of this invasion is then investigated by defining the position of the wavefront to be the first position, counting from in front of the wave, at which there were three or more cells in a column. It is then possible to compare the predicted wavespeed from the corresponding Fisher-Kolmogorov equation to the speed in the simulated lattice-based model, finding that the speed obtained from simulations is more sensitive than the predicted wavespeed to changes in the proliferation parameter. However, the wavespeed was found to be insensitive to the carrying capacity used in the simulations, which is consistent with the known properties of the corresponding continuum Fisher-Kolmogorov equation (Simpson *et al.*, 2007a). The continuum equations used in this study were not derived directly from the underlying individual-based mechanisms, but instead simply assumed the Fisher-Kolmogorov equation. Thus the inconsistencies found indicate the need

for further work in deriving suitable continuum equations to model biological systems. These are discussed further in the following section.

A similar method may be used in wound healing, using an on-lattice exclusion process, where cells may proliferate or move on a one- or two-dimensional grid. Initially all the cells to the left of the origin are occupied, so that an invasive wave is generated to the right. The wavespeed of this invasion is then predicted by defining the position of the wavefront to be approximately equal to the total mass of created particles. It is then possible to compare the predicted wavespeed from the Fisher-Kolmogorov equation to the simulated wavespeed (Callaghan *et al.*, 2006).

1.5 Deriving continuum equations from IBMs

Continuum models, such as those in Section 1.3.1, have historically been widely used to model migration and diffusive processes. The equations used are rarely derived from considering the behaviour of individuals in detail, however, so that it is therefore difficult to study cell-cell interactions such as volume exclusion in this framework. It is also difficult to parameterise continuum models from experimental data, since it is unclear how to connect experimental data gathered at the level of individual cells with model parameters, as several aspects of migration may be integrated into a single parameter. To address these issues, recent mathematical analysis has begun to focus on methods for systematically deriving continuum PDEs from IBMs. In particular, Simpson and coworkers have focussed on analysing on-lattice models with volume exclusion (Simpson *et al.*, 2009c; Baker and Simpson, 2010).

1.5.1 Simpson *et al.* (2009c)

Simpson *et al.* (2009c) begin with a simple exclusion process, examining a biased random walk on a square lattice. At each time step each agent at (x, y) moves to $(x, y \pm \Delta)$ with probability $P/4$ and to $(x \pm \Delta, y)$ with probability $P(1 \pm \rho)/4$. Here, P is the probability of moving during a given time step and ρ is a parameter giving the bias of the random walk. If $N_n^m(i, j)$ is the occupancy of site (i, j) after n steps of the mth simulation of the model then the mean occupancy,

$$\langle N_n(i,j) \rangle = \frac{1}{M} \sum_{m=1}^{M} N_n^m(i,j), \tag{1.8}$$

satisfies the following approximate equation

$$\begin{aligned}
&\langle N_{n+1}(i,j)\rangle - \langle N_n(i,j)\rangle \\
&= -\frac{P}{4}\langle N_n(i,j)\rangle[1-\langle N_n(i,j+1)\rangle] - \frac{P}{4}\langle N_n(i,j)\rangle[1-\langle N_n(i,j-1)\rangle] \\
&\quad -\frac{P}{4}(1+\rho)\langle N_n(i,j)\rangle[1-\langle N_n(i+1,j)\rangle] - \frac{P}{4}(1-\rho)\langle N_n(i,j)\rangle[1-\langle N_n(i-1,j)\rangle] \\
&\quad +\frac{P}{4}\langle N_n(i,j+1)\rangle[1-\langle N_n(i,j)\rangle] + \frac{P}{4}\langle N_n(i,j-1)\rangle[1-\langle N_n(i,j)\rangle] \\
&\quad +\frac{P}{4}(1-\rho)\langle N_n(i+1,j)\rangle[1-\langle N_n(i,j)\rangle] + \frac{P}{4}(1+\rho)\langle N_n(i-1,j)\rangle[1-\langle N_n(i,j)\rangle]. \qquad (1.9)
\end{aligned}$$

The positive terms in equation (1.9) are due to individuals moving into site (i,j), whilst the negative terms are due to transitions out of site (i,j). For example, the first term, $-P/4\langle N_n(i,j)\rangle[1-\langle N_n(i,j+1)\rangle]$ represents the movement of individuals from site (i,j) to $(i,j+1)$, where $P/4$ is the probability that an individual attempts to move, $\langle N_n(i,j)\rangle$ is the average number of such individuals at site (i,j) and $[1-\langle N_n(i,j+1)\rangle]$ is the probability that the site $(i,j+1)$ is free for individuals to move to. The occupancy of adjacent sites is assumed to be independent and hence the terms are multiplied together to give the first term, $-P/4\langle N_n(i,j)\rangle[1-\langle N_n(i,j+1)\rangle]$.

Equation (1.9) may be rearranged to give a PDE; by dividing through by the time step, τ, and letting Δ and τ tend to zero, whilst holding Δ^2/τ constant we arrive at

$$\frac{\partial N}{\partial t} = D\nabla^2 N - \nu\frac{\partial}{\partial x}[N(1-N)], \qquad (1.10)$$

where

$$D = \lim_{\Delta,\tau\to 0}\frac{P\Delta^2}{4\tau}, \quad \nu = \lim_{\Delta,\tau\to 0}\frac{\rho P\Delta}{2\tau}, \qquad (1.11)$$

and N is a continuous representation of $\langle N_n(i,j)\rangle$.

If there is no volume exclusion in the model, then equation (1.9) is simplified to give

$$\begin{aligned}
\langle N_{n+1}(i,j)\rangle - \langle N_n(i,j)\rangle = &-\frac{P}{4}\langle N_n(i,j)\rangle - \frac{P}{4}\langle N_n(i,j)\rangle - \frac{P}{4}(1+\rho)\langle N_n(i,j)\rangle \\
&-\frac{P}{4}(1-\rho)\langle N_n(i,j)\rangle + \frac{P}{4}\langle N_n(i,j+1)\rangle + \frac{P}{4}\langle N_n(i,j-1)\rangle \\
&+\frac{P}{4}(1-\rho)\langle N_n(i+1,j)\rangle + \frac{P}{4}(1+\rho)\langle N_n(i-1,j)\rangle, \qquad (1.12)
\end{aligned}$$

and as Δ and τ tend to zero, whilst holding Δ^2/τ and $\rho\Delta/\tau$ constant, as before, we obtain the

linear drift equation

$$\frac{\partial N}{\partial t} = D\nabla^2 N - \nu\frac{\partial N}{\partial x}. \tag{1.13}$$

Note that, returning to the case with volume exclusion, if $\rho = 0$, so that the random walk is unbiased, we also recover the diffusion equation (since $\nu = 0$), and volume exclusion is predicted to have no effect on the population movement.

The same model is then studied with multiple subpopulations, with continuum limits being derived as before. The solutions of the continuum equations agree well with the simulation data, and an intriguing result is noted, namely that the density profile of each subpopulation does not always obey a maximum principle (where the maximum value of the solution over the domain decreases monotonically with time) even though the total population density profile does obey this principle. In addition, when movement may be biased or unbiased, the population-level results cannot distinguish between the two types of movement. This has implications for the analysis of experimental results, which often only produce population-level data (Simpson *et al.*, 2009b). It is also possible to introduce multiple species into the model to simulate transplant experiments undertaken and it is interesting to note that the continuous approximations for the subpopulations are then found to be different to that for the total population (Simpson *et al.*, 2009c). This indicates that the behaviours of subpopulations of the stream may not be easily deduced from the overall behaviour of the stream. A similar technique may be used to study on-lattice models with different movement mechanisms (Simpson *et al.*, 2009a, 2010b; Fernando *et al.*, 2010).

1.5.2 Baker and Simpson (2010; 2011)

The assumption used by Simpson *et al.* (2009c) that the occupancy of adjacent lattice sites is approximately independent over repeated simulation is often used in derivations of this type (Simpson *et al.*, 2010a; Deroulers *et al.*, 2009). This assumption is explored by Baker and Simpson (2010; 2011) using k-point distribution functions, $\rho^{(k)}$, the multivariate probability distribution functions for the occupancy probability of k-tuplets of lattice sites. To derive a continuum equation for the distribution of cells we therefore wish to find the one-point distribution function, $\rho^{(1)}$. The authors

study a birth-death process with migration, so that

$$\frac{\partial \rho^{(1)}(A_l)}{\partial t} = P_m \sum_n \frac{\alpha_{n,I}}{z} \left[\rho^{(2)}(0_l, A_n) - \rho^{(2)}(A_l, 0_n)\right] + P_p \sum_n \frac{\alpha_{n,I}}{z} \rho^{(2)}(0_l, A_n) - P_d \rho^{(1)}(A_l), \tag{1.14}$$

where P_m, P_p and P_d are the rates of movement, proliferation and death, respectively, and A_i and 0_i denote the occupation or not of site i by an agent. For example, $\rho^{(2)}(0_l, A_n)$ is the probability that site l is unoccupied and site n is occupied. Thus it is clear that to find the one-point distribution function we must know the two-point distribution functions. Similarly, to find the two-point functions, knowledge of the three-point distribution functions is required, and so on. This is a long chain of equations, dependent on the lattice grid size, and cannot be closed to obtain a tractable system of differential equations without making an approximation, such as the one described earlier, that the occupancy of adjacent sites is independent. The authors explore the effects of taking different closure assumptions by considering the correlation functions

$$F_{\lambda,\mu}(|\mathbf{l} - \mathbf{m}|) = \frac{\rho^{(2)}(\sigma_l, \sigma_m)}{\rho^{(1)}(\sigma_l)\rho^{(1)}(\sigma_m)}, \tag{1.15}$$

where $\sigma_i = \{A, 0\}$ describes the occupation of site i for $i = l, m$. $F_{\lambda,\mu}$ depends only on the distance, $|\mathbf{l} - \mathbf{m}|$, between the two lattice sites for the homogeneous initial conditions considered in the 2010 paper. Hence $F_{\lambda,\mu}(\Delta) \equiv 1$ represents the usual assumption of independence for lattice spacing of size Δ. Using a random initial seeding of agents in the domain, it is possible to find the evolution of the correlation functions under different closure assumptions. The authors investigate the additional accuracy gained by taking the Kirkwood superposition approximation (KSA) at the three-point level (Singer, 2004), rather than making the usual closure assumption of independence at the two-point level.

Baker and Simpson (2010) studied an initially homogeneous agent distribution in two and three spatial dimensions and determine that the higher order closure approximation given by the KSA gives greater accuracy when compared to model simulations, particularly in cases with higher proliferation rates. The KSA is found to be a better assumption in higher dimensions, although the usual assumed independence of neighbouring lattice sites is also violated to a lesser extent so that both of the considered approximations perform better in higher spatial dimensions.

Simpson and Baker (2011) extended the previous study to include initially inhomogeneous distributions of individuals in one spatial dimension by considering the correlation functions to be

$$F_{\lambda,\mu}(l,m) = \frac{\rho^{(2)}(\sigma_l, \sigma_m)}{\rho^{(1)}(\sigma_l)\rho^{(1)}(\sigma_m)}, \tag{1.16}$$

and once again obtain an improvement to the equations derived by using the KSA rather than the usual independence assumption. The authors assume that $F_{\lambda,\mu}(l, l+M) \to 0$ as $M \to \infty$, and thus reduce the equations solved to a computationally tractable system by only calculating $F_{\lambda,\mu}(l, l+M)$ for $M = 1, \ldots, 5$.

Both in higher spatial dimensions, with homogeneous initial conditions and in initially inhomogeneous models in one dimension, greater proliferation rates decrease the accuracy of the independence assumption compared to the higher order closure approximation. This is to be expected, since we then expect the positions of agents to be more highly correlated, so that assuming independence of adjacent lattice sites is less valid.

This extensive literature on deriving continuum equations from on-lattice volume exclusion processes provides techniques for systematically analysing how different mechanisms of movement may impact on the derived continuum equations. It is clear that volume exclusion may significantly change the form of the equations, leading to nonlinear terms. However, some systems such as those involving migrating chains of cells, cell-cell interactions or growth of the migratory domain may be more natural to model in an off-lattice setting. Thus the on-lattice techniques, whilst useful in many systems, may not always be applicable, and may result in artefacts that reflect the type of lattice chosen (Flache and Hegselmann, 2001).

1.5.3 Off-lattice models

Attempts at deriving continuum equations from off-lattice IBMs have mostly been restricted to the modelling of inanimate particles using statistical mechanics techniques resulting in the Boltzman equation and modifications to it. These are not necessarily applicable to cellular populations, where more complex movement may occur and the system may gain or lose energy through a variety of non-physical routes. Another group of studied models consider individuals to move at a constant velocity, with probabilities of changing speed and direction at each time point (Codling, 2003; Codling *et al.*, 2008). These 'velocity-jump' models are in contrast to the 'position-jump' models described earlier in which individuals are assumed to move instantaneously to a new position when movement

occurs. Velocity-jump models have not often been used to study volume exclusion processes, but are particularly suited to the movement of bacteria, which are usually characterised to continually be either moving in a particular direction ('running'), or changing direction ('tumbling'). One example system is the case of rod-shaped bacteria, which can reverse their direction of motion. These are modelled by Gejji *et al.* (2012) as a one-dimensional on-lattice velocity-jump process considered with volume exclusion. The authors do not derive continuum equations directly from the underlying microscopic model, but instead posit that cell density may be described by a non-linear diffusion equation

$$\frac{\partial p}{\partial t} = \frac{\partial}{\partial x}\left(D(p)\frac{\partial p}{\partial x}\right), \tag{1.17}$$

and use Boltzmann-Matano analysis to find $D(p)$ numerically from simulated data.

Two papers have been published that consider deriving continuum equations from an off-lattice exclusion process. The first considers a system of hard spheres with interparticle collisions and an optional external force on all particles (Bruna and Chapman, 2012). The authors model collisions between particles as reflective boundary conditions on collision surfaces and take matched asymptotic expansions in the size of a particle to derive a PDE for the volume concentration of particles. Interestingly, while on-lattice volume exclusion diffusive processes result in the unmodified diffusion equation, Bruna and Chapman (2012) derive a non-linear diffusion coefficient that is modified by the local concentration of particles. The hard-sphere model considered is appropriate for studying inanimate particles, but does not easily extend to considering more complex interactions between individuals, such as those seen in cell-based biological systems.

The second paper deriving a continuum equation from an off-lattice model with volume exclusion divides the domain up into columns, assuming that the concentration of agents across the width of a column will be approximately constant. In this way the authors reduce the problem to the equivalent on-lattice model and hence require the addition of proliferation for volume exclusion to have any effect on the derived equation (Plank and Simpson, 2012).

1.5.4 Chemotaxis

Finally we discuss the various attempts that have been made to systematically derive a continuum approximation to describe an individual-based chemotactic model. On-lattice models have been considered without volume exclusion by Painter and Sherratt (2003) and Baker *et al.* (2010). Whilst

Baker *et al.* (2010) consider an on-lattice model with agent-based chemoattractant and cells, the model is presented in the appendix of the paper, and is not analysed in detail. In contrast Painter and Sherratt (2003) derives equations in which cells produce a chemoattractant, giving rise to aggregation. The derived equations are studied using linear stability analysis and solutions are plotted under varying parameter values, but are not compared to simulations of the original on-lattice model. It is not clear, therefore, whether the chemoattractant is intended to be modelled as an individual-based species or a continuum, and how it should be coupled to the individual-based cells in either case.

Chemotactic motion in off-lattice velocity-jump models is studied by Erban and Othmer (2004), who derive continuum equations in the case where the chemoattractant is a given function within the domain and is not affected by the cell distribution. The authors use moment closure techniques to derive continuum equations from the system and analyse the system under different applied signal functions.

1.6 Aims and outline

This chapter began by introducing the biological system under consideration: the migration of CNCCs from the neural tube to ba2. We have discussed previous mathematical models of this system and the related migratory system describing migration of the enteric neural crest. However, experimental understanding of the mechanisms of CNCC migration far outstrips the complexity of the mathematical models used to date. In addition, the outcome of experimental perturbations is observed either as the interaction of individual cells, at a fine time resolution but for a short period of time (Teddy and Kulesa, 2004), or as the migration of the total cell population over long periods of time (Kulesa *et al.*, 2008). It is difficult, therefore, to use these data to infer the mechanisms by which individual cells affect the population-level migration that is observed.

The aim of this thesis is to use a combination of individual-based and hybrid modelling techniques in close collaboration with our experimental colleagues to interrogate the mechanisms of CNCC migration. We wish to understand the method by which cells gain directional information from their environment and use this to move as a directed, cohesive population to colonise the migratory route. To investigate the system we develop a hybrid model that can incorporate different cell-cell interactions which can be direct, through cell-cell contact, or indirect, by explicitly modelling the space occupied by cells and prohibiting cells to overlap each other. Motivated by experimental

evidence and hypothesised migratory mechanisms, we also examine the way in which cells may use a spatially homogeneous production of chemoattractant to create a chemotactic gradient. To further investigate the effect that the space occupied by cells has on the migration of individual cells, and the population as a whole, we must develop new systematic techniques for deriving continuum equations from off-lattice IBMs with volume exclusion. Finally, we formulate new methods to understand how different model formulations of chemotaxis are reflected in the corresponding continuum equations and use these equations to analyse hybrid models of CNCC migration. The modelling methodologies developed here have significant application in many different areas, from chemotaxis in cellular migration during embryo growth and in the immune system, to the effects of volume exclusion on the movement of crowds of people during emergency evacuations and the modelling of molecular movement in crowded environments.

We will begin in Chapter 2 by developing an IBM for CNCC migration using the current hypothesised mechanisms underlying the biological system, and experimental data from the wildtype system. Using this model, we will then predict, in Chapter 3, the response of the system to perturbations and compare our predictions to experimental results. We will use our predictions to infer information about the biological system by comparing with the experiments and use experimental data to extend and enhance our model.

In the remainder of the thesis we will build systematic techniques to analyse our hybrid IBM from Chapters 2 and 3. We begin in Chapter 4 with a simple one-dimensional off-lattice exclusion process, describing cell motility with volume exclusion. We will compare numerical solutions to our derived equations with averaged simulations of the model and with numerical solutions of the equations derived from a system without volume exclusion. In this way we may investigate the effect that the, often taken, assumption of point-like cells has on the equations derived and the evolution of cell densities. Hence we may discover whether, and under what circumstances, it is important to include volume exclusion when modelling cellular migration. We extend this analysis in Chapter 5 to two spatial dimensions.

In the penultimate chapter (Chapter 6) we consider chemotactic systems and compare the derived equations under different modelling regimes. We discuss ways of coupling on- and off-lattice individual-based species with a continuum species and compare numerical solutions to the derived equations with averaged model simulations.

Finally, in Chapter 7, we use the techniques developed in Chapters 4 to 6 to analyse models from

Chapters 2 and 3 and explore their response to different parameter regimes using the derived equations. We also discuss further work to be undertaken, both in the study of CNCC migration, and in the derivation of continuum equations from IBMs.

Chapter 2

Setting up a hybrid modelling framework

In this chapter we progressively build a mathematical model that can be used to study the biological system described in Section 1.1. The aim is to construct a model that can robustly achieve the migratory behaviour seen *in vivo*. The model will then be used in Chapter 3 to predict the response of the biological system to experimental perturbations and interpret the results of those experiments.

2.1 Simple diffusion

In our system, CNCCs migrate long distances from the neural tube to the second branchial arch. Since the domain is initially empty of CNCCs and after 24 hours there are still only on the order of 100 cells in the migratory domain (Kulesa *et al.*, 2008), it is important to investigate the effects of stochasticity on the system. In addition if we wish to study cell-cell interactions, or to explicitly include mechanisms on an individual cell level, then we need to use an IBM for cells.

We begin with the simplest possible model of cellular movement, diffusion, constrained to a rectangular domain that represents the migratory pathway. The domain measures 300μm from the neural tube at $x = 0$ to the second branchial arch at $x = 300$ and is 120μm high (see Table 2.1 for parameter values used in the models). Two cells enter at the neural tube every 15 minutes (as used in Kulesa *et al.*, 2008) with uniformly distributed y-coordinates.

Cells move with a rate α a fixed distance d in a random direction. This is implemented by incrementally increasing time by Δt so that each cell has a probability $\alpha\Delta t$ of moving during each time step. Using smaller values for Δt and/or d allows for a more fine-grained description of the system, where the position of a cell at any particular time is known more exactly, whilst increasing the computational cost of simulations.

Using the average speed, v, of CNCCs (around 45μm/hr; Kulesa *et al.*, 2008) we find $\alpha = v/d$. Then we require that the probability, $\alpha\Delta t$, of any particular cell moving in time Δt is small enough that we may assume at most one cell moves during each time step.

2.1.1 Simulations

In our simulations, diffusing cells typically migrate at most 25μm into the domain in 24 hours (Figure 2.1), compared to the 575μm seen experimentally. However, we cannot directly compare the distances travelled since the migratory route more than triples in length during migration imparting a significant additional advective flux to the cells which must be taken into account in modelling work.

2.2 Domain growth

To allow direct comparison of simulations and experimental data and to take account of the additional advective flux imparted by tissue growth we next incorporate domain growth into the model.

McLennan *et al.* (2012) measured the length of the migratory route at different stages during embryo development and found growth to be approximately logistic. We therefore fit a logistic curve to the domain length data, of the form

$$L(t) = \left(\frac{L_\infty e^{L_\infty \beta (t - t_s)}}{L_\infty - 1 + e^{L_\infty \beta (t - t_s)}} + 1 - \frac{L_\infty e^{L_\infty \beta (-t_s)}}{L_\infty - 1 + e^{L_\infty \beta (-t_s)}} \right) L_0, \tag{2.1}$$

where L_0 is the initial length of the domain, β and L_∞ give an indication of the rate of growth and the limiting length of the domain, respectively. Using least squares regression we obtain $L_\infty = 960\mu$m, $\beta = 0.0580/\mu$m/hr and $t_s = -16$ hours (see Figure 2.2).

Since growth occurs due to the proliferation of the overlying ectodermal cells on which the CNCCs migrate, we assume that growth is uniform and that the CNCCs have an additional advective flux

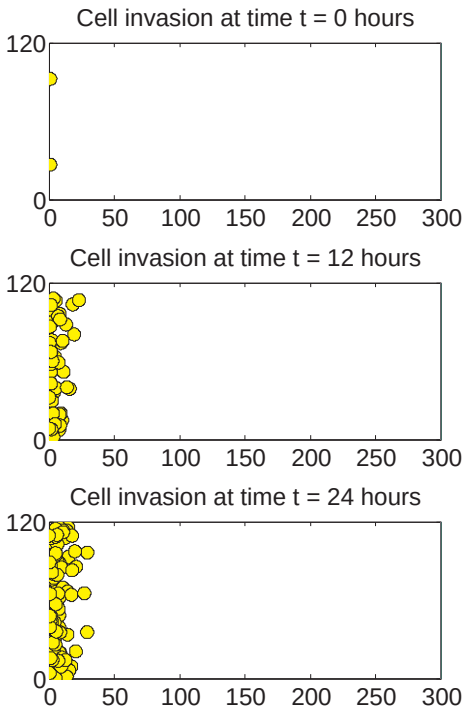


Figure 2.1: Cells migrating by simple diffusion on a fixed domain (see Section 2.1) only invade approximately 25μm into the domain in 24 hours. Note that although cells are represented as yellow discs, there is no excluded volume effect so that individual discs may overlap.

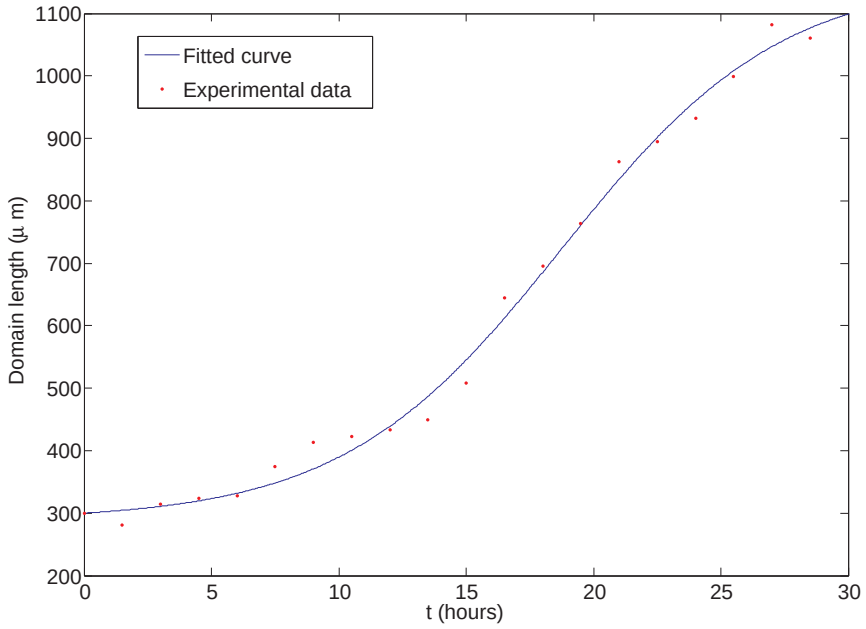


Figure 2.2: Growth profile for the domain of migration, showing experimental data and fitted logistic curve (see Equation (2.1)).

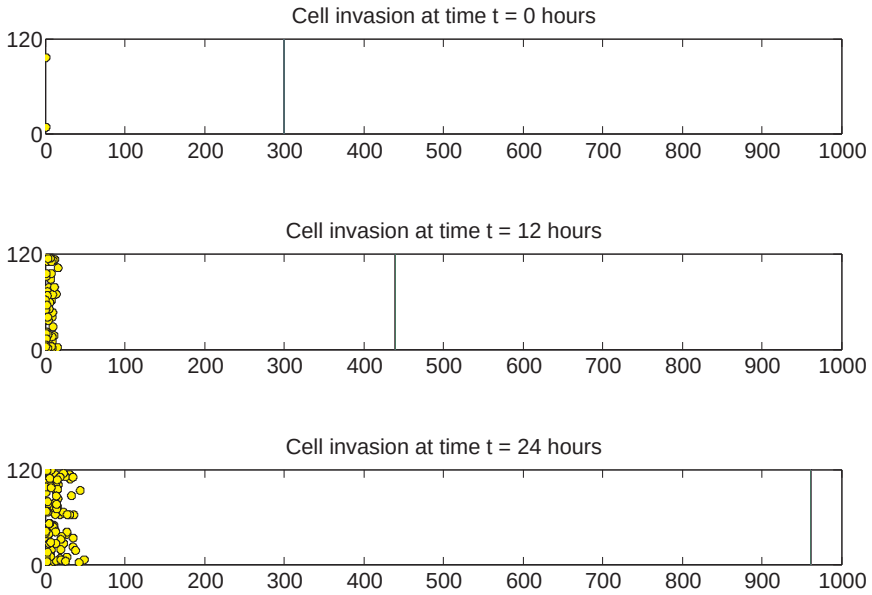


Figure 2.3: Cells migrating by simple diffusion on a growing domain (see Sections 2.1–2.2) only invade approximately 50μm into the domain in 24 hours. The length of the domain is shown at each time step by a vertical line. Note that although cells are represented as yellow discs, there is no excluded volume effect so that individual discs may overlap.

due to domain growth. Computationally this is achieved by multiplying the x-coordinate of each cell by $L(t)/L(t-\Delta t)$ at each time step of length Δt.

2.2.1 Simulations

The additional advective flux imparted by domain growth does allow cells to travel further into the domain than in the case of the fixed domain. However, simulated cells still only migrate around 50μm into the domain (Figure 2.3), much less far than seen experimentally. One possible mechanism for increasing the speed of migration is that the large numbers of cells entering the migratory domain bias the movement of cells towards the positive x-direction. To investigate the effects of this type of population pressure, however, we must consider the space occupied by individuals in the domain.

2.3 Volume exclusion

Cells are often modelled as point particles that are able to move through and around each other. In reality, however, migration can be greatly affected by the presence of other cells, since individuals cannot overlap. Large groups of cells restrict the available space for migration and, as such, can bias

movement away from highly occupied areas. Conversely, cell movement can also be greatly restricted by other cells preventing migration through crowded spaces. In the cranial neural crest, the y-axis of the migratory route is only eight cell diameters wide, leading us to conclude that volume exclusion is an important mechanism to consider.

We implement volume exclusion in our model by preventing any cell movement that would result in a distance of less than a cell diameter between two cell centres. Movements to within a cell radius of the domain boundaries are also rejected. In addition, when new cells enter the domain they must not overlap with any existing cell. If a cell attempts to enter at an occupied point, then other points are tried at random to a maximum of m points. If all points fail then no cell enters the domain, representing the case where cells are prevented from delaminating from the neural tube by the pressure of other cells. For low values of m the population of migrating cells is greatly reduced, since cells next to $x = 0$ have a high probability of preventing new individuals from entering the domain. Once m is high enough, however, further increases do not significantly affect the population since a cell failing to enter at many different places implies that very few (or no) feasible entry points exist. As m gets very large, the computation simulation time increases whilst not significantly affecting the model outcome.

2.3.1 Simulations

Simulations of diffusive cells with domain growth and volume exclusion migrate further into the domain than in the point-particle model, although fewer cells can successfully enter the domain. In particular, the initial cells to enter the domain are typically found at approx. $x = 50\mu$m after 24 hours (Figure 2.4) compared to $x = 25\mu$m in the previous model (Figure 2.1). We note that cells are inserted further into the domain when volume exclusion is present, since they cannot be within a cell radius of the boundary. The increase in migration distance, however, is much greater than a cell radius and we conclude that the main increase in distance is from the bias in motion introduced by volume exclusion since movements away from crowded areas are more likely to be successful.

The distance migrated is still far short of that seen experimentally, however, where leading cells can reach up to 575μm. Thus we conclude that passive diffusion does not provide successful migration on the timescales required; instead cells need an active migratory mechanism.

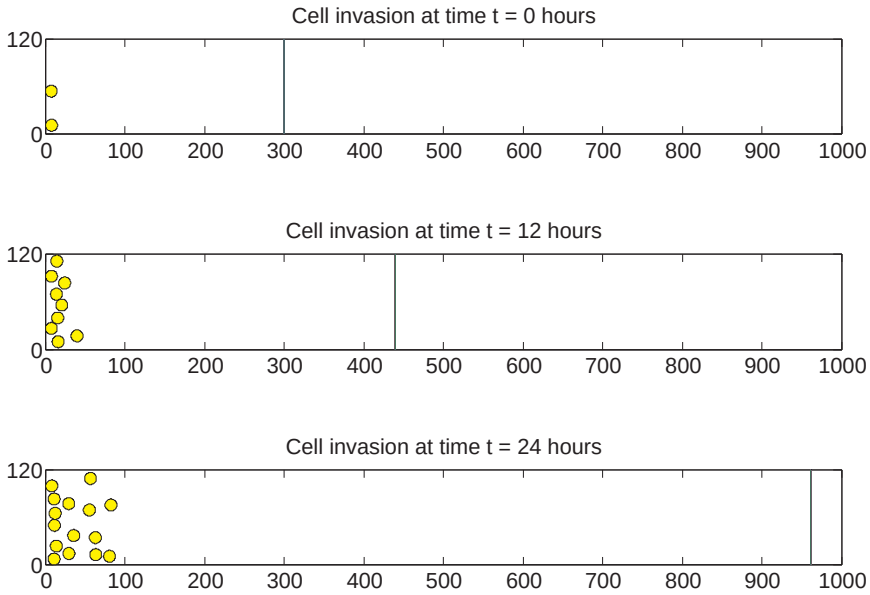


Figure 2.4: Cells migrating by simple diffusion with domain growth and volume exclusion (see Sections 2.1 to 2.3) only invade approximately 100μm into the domain in 24 hours. Note that cells are inserted further into the domain than when volume exclusion is not incorporated, to prevent overlapping with the domain boundary.

2.4 Chemotaxis

We will investigate chemotaxis as a form of active migration, motivated by experimental evidence for the presence of a chemoattractant *in vivo* (McLennan *et al.*, 2010). There are multiple different ways that we could implement dependence on chemoattractant levels *in silico*, which can involve cells moving at different rates depending on the level of chemoattractant at the cell position and/or the level at another point away from the cell position. Different methods produce similar results, however, in which cell movement is biased in directions with positive chemoattractant gradients.

As with diffusion, at each time point all cells sample in a random direction, but whether or not the cell moves is now determined by the chemoattractant gradient. We assume that the chemoattractant can be determined away from the cell position by extending a filopodium. The chemoattractant at the end of the filopodium is then compared with the levels at the cell body and the cell moves if the gradient is favourable. We note that cells *in vivo* may have many filopodia, with slowly changing positions relative to the cell. Our single fast-moving filopodium, therefore, is a modelling simplification to the biological system.

The question remains of how to determine the level of chemoattractant at a point. We would expect that individuals measure the chemoattractant concentrations close to their cell body and filopodia

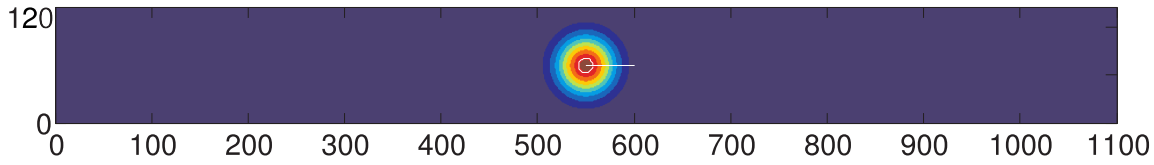


Figure 2.5: The weighting function for chemoattractant sensing for $\delta = 30$ (see Section 2.4).

but have less information about concentrations further away. The level of chemoattractant, c, sensed at a point (either the cell body or the end of a filopodium), (x_c, y_c), is therefore found by multiplying the chemoattractant by a function that gives greater importance to chemoattractant closer to the considered point (Figure 2.5). We use

$$I = \int_0^{x_L} \int_0^{y_L} \frac{c(x,y)}{\delta\sqrt{\pi}} \exp\left\{-\frac{\left((x-x_c)^2 + (y-y_c)^2\right)}{\delta^2}\right\} \mathrm{d}x\ \mathrm{d}y, \tag{2.2}$$

where δ is a measure of the distance from (x_c, y_c) at which chemoattractant may still be measured and x_L and y_L are the length and height of the domain, respectively.

2.4.1 Simulations

We simulate with a fixed chemoattractant gradient on a growing domain with no chemoattractant dilution, so that the maximum chemoattractant concentration in the domain remains fixed during domain growth. With fixed chemoattractant cells typically migrate around $600 - 700\mu$m in 24 hours (Figure 2.6) with our chosen set of parameters, which is much closer to the 575μm seen experimentally. The distance migrated does vary with the parameter, δ, which determines how far the cells can sense the chemoattractant concentration. As δ decreases, the sensing distance decreases, and the cells spread out in a wider stream as they migrate, since there is less chance of the sensing area overlapping with the domain boundary. Cells that are more spread out across the width of the domain spend less time moving directly towards the end of the domain, and hence move less far in a given time. The distance migrated ranges from around 350μm for $\delta = 1$ to 700μm for $\delta = 15$, indicating that our choice of δ is important. We note that whilst δ is not given explicitly by experimental data, a reasonable choice can be made by considering the size of a cell body and assuming that cells do not sense chemoattractant further away than the length of a filopodium.

Chemotaxis offers a cell migration mechanism that is robust to stochastic variations between simulations. However, the question remains as to how the chemoattractant gradient is established and

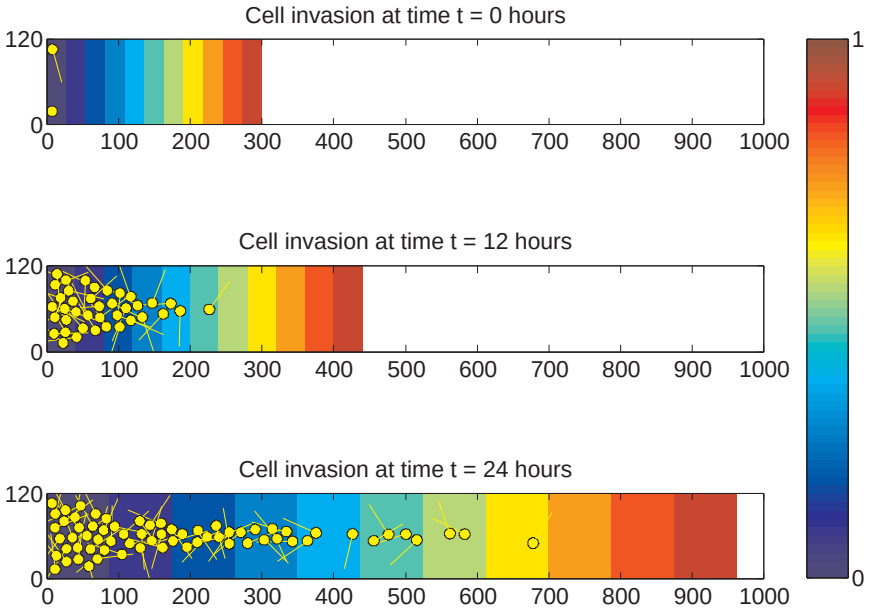


Figure 2.6: Cells migrating by chemotaxis with domain growth and volume exclusion (see Sections 2.1 to 2.4) typically invade approximately $600 - 700\mu$m into the domain in 24 hours. Note that the chemoattractant profile is fixed, so that domain growth does not lead to dilution of the chemoattractant.

maintained against destabilising factors such as dilution due to domain growth.

2.5 Dynamic chemoattractant

McLennan *et al.* (2010) found evidence that VEGF is a chemoattractant *in vivo* with spatially homogeneous production. To investigate the creation of a chemoattractant gradient from spatially homogeneous production, we will now explicitly model the chemoattractant dynamics. Since there are a relatively large number of chemoattractant molecules compared to the number of cells, we will use a continuum PDE for the chemoattractant concentration. We hypothesise that cells may internalise VEGF as they sense their surroundings, thus depleting VEGF in the area around the cell. The levels of VEGF are then lower in areas close to the neural tube, where cells are present for longer periods, providing a positive chemoattractant gradient to the end of the migratory domain.

We assume standard logistic production of chemoattractant throughout the region with linear production rate, χ, but note that VEGF profiles over time are qualitatively similar under different production terms. The only exception to this is if the production rate is so high as to overwhelm the internalisation of VEGF by cells. Chemoattractant consumption is often modelled on-lattice

by depleting the chemoattractant at lattice points that contain cells. In an off-lattice model, we cannot simply have a sink term that is only present at cell locations, since functions that are zero everywhere except for a point cannot be implemented numerically. Instead we use the same approximation to a delta function for chemoattractant consumption terms as we used for the sensing mechanism (equation (2.2)), since this approximates point consumption terms.

With the given assumptions the chemoattractant concentration, $c(x, y)$, neglecting domain growth, is governed by the following equation.

$$\frac{\partial c}{\partial t} = \overbrace{D_c\left(\frac{\partial^2 c}{\partial x^2} + \frac{\partial^2 c}{\partial y^2}\right)}^{\text{diffusion}} \overbrace{-\lambda \sum_{i=1}^{N(t)} \frac{c(x,y)}{\delta\sqrt{\pi}} \exp\left[-\frac{\left((x-x_i)^2 + (y-y_i)^2\right)}{\delta^2}\right]}^{\text{internalization}} + \overbrace{\chi c(1-c)}^{\text{production}}, \tag{2.3}$$

where D_c is the diffusion coefficient for the chemoattractant, δ is a measure of the distance from the ith cell centre at (x_i, y_i) at which chemoattractant is consumed. There are a total of $N(t)$ cells in the domain at time t, and cell consume chemoattractant at a rate λ.

2.5.1 Domain growth

Uniform domain growth affects the form of the chemoattractant equation, which can be derived in the following way. If $L = L(t)$ is the length of the domain, then conservation of the total mass of chemoattractant in an arbitrary box $[L_1(t), L_2(t)] \times [y_1, y_2]$ gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{y_1}^{y_2}\int_{L_1(t)}^{L_2(t)} c(x,y,t)\ \mathrm{d}x\ \mathrm{d}y = \int_{y_1}^{y_2}\int_{L_1(t)}^{L_2(t)} \left(\frac{\partial^2 c}{\partial x^2} + \frac{\partial^2 c}{\partial y^2} + f(c,x,y,t)\right)\ \mathrm{d}x\ \mathrm{d}y, \tag{2.4}$$

where $f(c, x, y, t)$ is the net production of chemoattractant at (x, y). Using Reynold's Transport Theorem (Reynolds and Brightmore, 1903) we can evaluate the left-hand side of this equation:

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{L_1(t)}^{L_2(t)} c(x,y,t)\ \mathrm{d}x = \int_{L_1(t)}^{L_2(t)} \left(\frac{\partial c}{\partial t} + \frac{\partial}{\partial x}(uc)\right)\ \mathrm{d}x, \tag{2.5}$$

where $u(x, t)$ is the flux due to the expansion of the domain. Since equation (2.5) holds for all boxes $[L_1(t), L_2(t)] \times [y_1, y_2]$, we have

$$\frac{\partial c}{\partial t} + \frac{\partial}{\partial x}(uc) = D_c\left(\frac{\partial^2 c}{\partial x^2} + \frac{\partial^2 c}{\partial y^2}\right) - \lambda \sum_{i=1}^{N(t)} \frac{c(x,y)}{\delta\sqrt{\pi}} \exp\left[-\frac{\left((x-x_i)^2 + (y-y_i)^2\right)}{\delta}\right] + \chi c(1-c). \tag{2.6}$$

To find $u(x,t)$ in terms of $L(t)$, let the movement of a material point X due to growth alone be given by $\Gamma(X,t)$, so $u(x,t) = d\Gamma/dt$. For isotropic growth $\Gamma(X,t) = Xr(t)$ where $r(0) = 1$. Now

$$u(x,t) = \frac{\mathrm{d}\Gamma}{\mathrm{d}t} = X\dot{r} = \frac{x}{r}\dot{r}, \tag{2.7}$$

so

$$\frac{\partial c}{\partial t} + \frac{\partial}{\partial x}(uc) = \frac{\partial c}{\partial t} + \frac{\dot{r}}{r}\left(x\frac{\partial c}{\partial x} + c\right). \tag{2.8}$$

Rescaling to a stationary domain $\bar{x} = x/L(t)$, $\bar{t} = t$ to solve numerically, and immediately dropping the bars gives

$$\frac{\partial c}{\partial t} = D_c\left(\frac{1}{L(t)^2}\frac{\partial^2 c}{\partial x^2} + \frac{\partial^2 c}{\partial y^2}\right) + f(c, Lx, t) - \frac{\dot{r}}{r}c, \tag{2.9}$$

$$= D_c\left(\frac{1}{L^2}\frac{\partial^2 c}{\partial x^2} + \frac{\partial^2 c}{\partial y^2}\right) - \lambda\sum_{i=1}^{N(t)}\frac{c(x,y)}{\delta\sqrt{\pi}}\exp\left[-\frac{\left(L^2(x-x_i)^2 + (y-y_i)^2\right)}{\delta}\right] + \chi c(1-c) - \frac{\dot{L}}{L}c, \tag{2.10}$$

since $r(t)$ does not depend on X, so $L(t) = L(0)r(t)$. Hence

$$\frac{\dot{r}}{r} = \frac{\dot{L}(t)/L(0)}{L(t)/L(0)} = \frac{\dot{L}}{L}. \tag{2.11}$$

The VEGF concentration is held constant at zero at each of the four boundaries of the region to artificially simulate the exclusion zones between the migrating streams without explicitly modelling inhibitory factors. This prevents the VEGF concentration from becoming artificially high near the boundaries due to the greater consumption by cells in the interior of the domain (see Figure 2.7).

Since our chemotactic mechanism only considers the sign of the chemoattractant gradient, not the magnitude, and we have no data as to the concentrations of VEGF *in vivo*, we will normalise the chemoattractant so that $c(x,y,0) = 1$ initially. However, to allow numerical solutions of the chemoattractant profile with the given boundary conditions, we interpolate between the initial condition within the domain, and the fixed boundary condition so that our initial condition is smooth (see Figure 2.8). The chemoattractant profile is solved at each time step using the NAG solver d03ra (see Appendix A.2). Note that we are assuming the PDE system is parabolic despite the very low diffusivity of VEGF. As $D_c \to 0$ the system becomes hyperbolic and d03ra is not recommended for

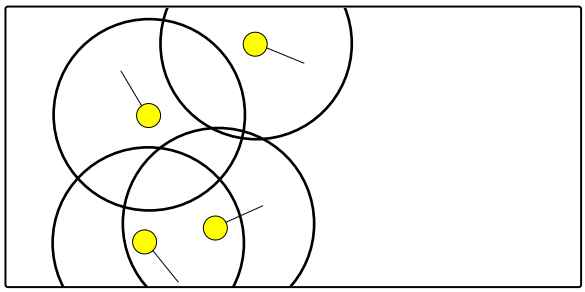


Figure 2.7: The chemoattractant in the interior of the domain is consumed more quickly since it is more likely to be close to multiple cells.

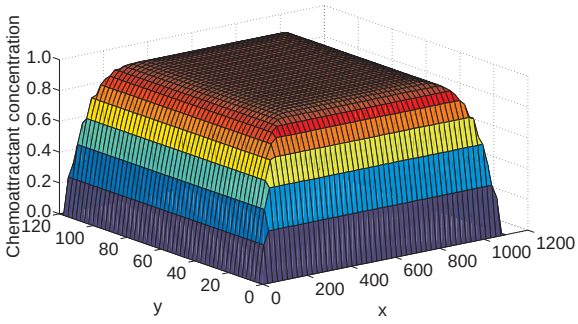


Figure 2.8: Chemoattractant initial condition.

solving such systems. However, numerical solutions are found to change gradually as the diffusivity increases from zero, implying that the numerical solver is successful under these conditions. For zero diffusivity, the system can be solved analytically for any given time step.

2.5.2 Diffusion of VEGF

The diffusivity of VEGF has not been determined *in vivo*, but has been estimated *in vitro* as $7 \times 10^{-7}\text{cm}^2/s = 2.52 \times 10^5 \mu\text{m}^2/\text{hr}$ (Chen *et al.*, 2007). This cannot be simulated with the boundary condition previously described since if the repulsive zones between migratory streams are represented by holding the chemoattractant concentration at the boundaries at zero then there is a high flux of chemoattractant out of the domain. Using a high value of diffusivity results in a huge loss of chemoattractant out of the domain, so that it is necessary to use zero flux boundary conditions instead. Using this estimated diffusivity results in a homogeneous chemoattractant profile as the y-coordinate varies (Figure 2.9). However, it is expected that the diffusivity *in vivo* will be much lower than this, since VEGF is known to bind to cell surfaces and extracellular matrix, thus significantly reducing the diffusivity. As the diffusivity is reduced the profile across the y-axis becomes non-constant, and the clustering of cells next to the entrance to the domain (at $x = 0$), which is already

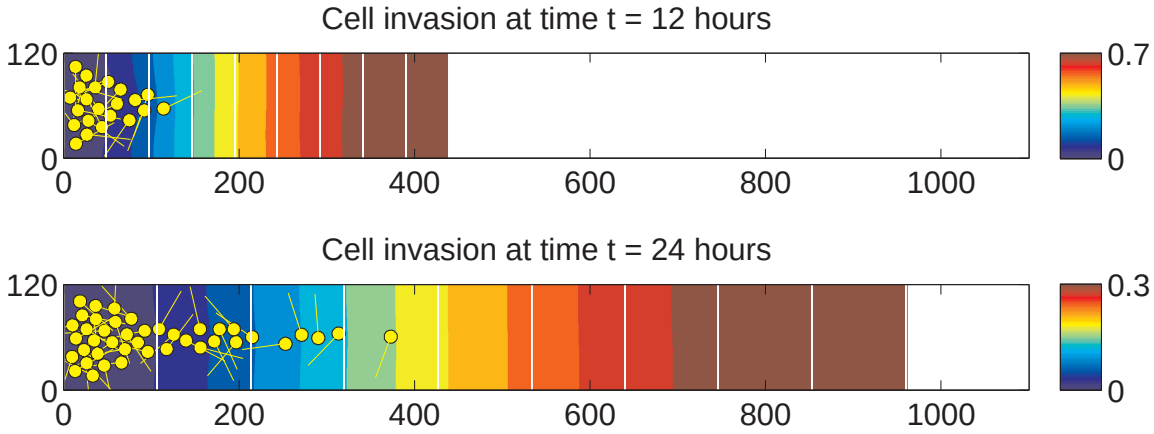


Figure 2.9: Representative simulation of the model described in Section 2.5 with a high diffusion coefficient, $D_c = 2.52 \times 10^5 \mu\text{m}^2/\text{hr}$. Chemoattractant concentration is denoted by the colour of the background. Note that the global levels of chemoattractant are decreasing due to the expansion of the domain. The colour bars adjacent to each diagram show the scale for the chemoattractant concentration. White lines show the movement of evenly spaced x-coordinates due to domain growth. Cells and filopodia are shown in yellow. Other parameters values are given in Table 2.1.

becoming a problem when $D_c = 2.52 \times 10^5 \mu\text{m}^2/\text{hr}$ (Figure 2.9) becomes more pronounced as D_c is reduced (Figure 2.10). Since the rate of diffusion of VEGF is not known *in vivo* and we expect the diffusion to be much reduced by binding to surfaces and matrix compared to the values found *in vitro*, we will err on the side of caution and take a very low diffusivity ($D_c = 0.1\mu\text{m}^2/\text{hr}$).

2.5.3 Simulations

When the chemoattractant gradient is dynamically established through cell consumption, leading chemotactic cells invade to $500 - 600\mu$m, slightly less far than in Section 2.4 (Figures 2.6 and 2.10). The distance migrated is similar for a range of diffusivities (Figures 2.9 and 2.10, other results not shown). The reduction in distance travelled compared to a static gradient can be attributed to a delay in migration due to the time required to establish the chemoattractant gradient. In the case with negligible chemoattractant diffusion, the less homogeneous chemoattractant profile across the domain width may also be a factor, leading to a favourable gradient towards the top and bottom domain boundaries as well as in the positive x-direction (see Figure 2.11). Cells at the front of the stream then have more wandering trajectories across the width of the domain, whilst still maintaining a bias in the positive x-direction.

However, later emerging cells, particularly in the reduced-diffusion case, congregate next to the neural tube (at $x = 0$), since the chemoattractant is depleted to such an extent that the gradient

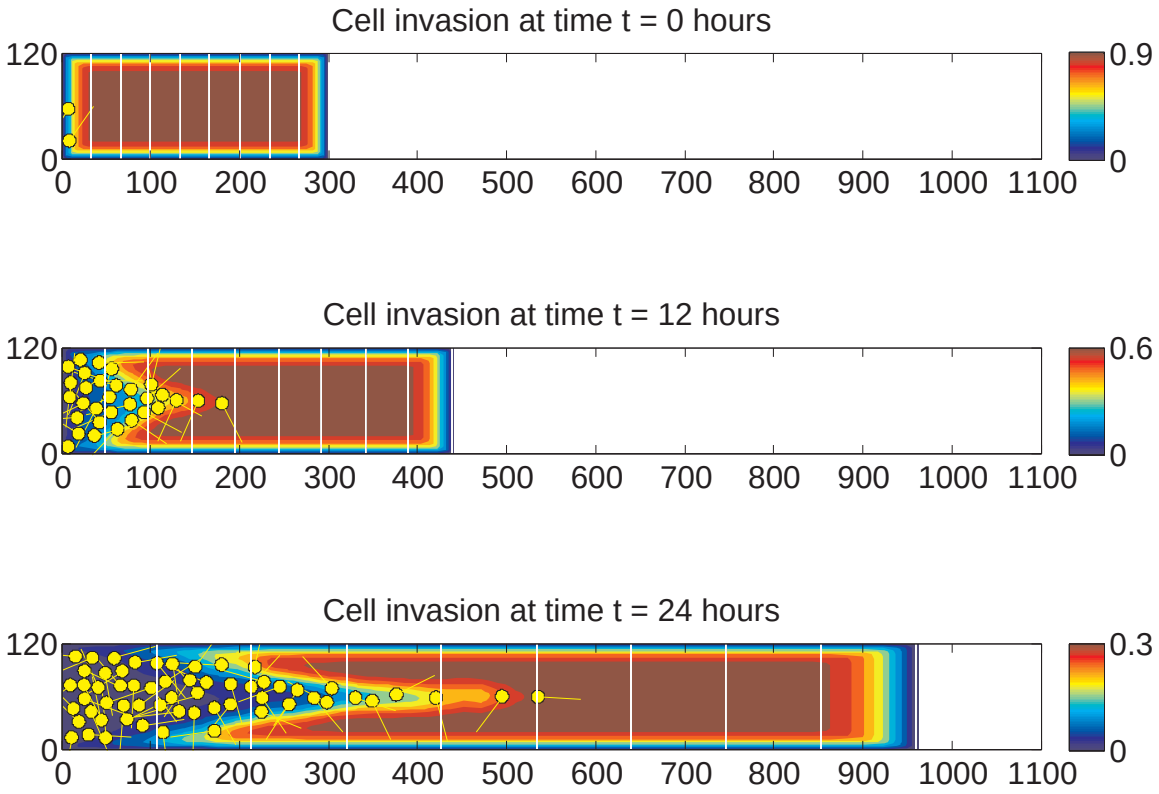


Figure 2.10: Representative simulation of the model described in Section 2.5. Chemoattractant concentration is denoted by the colour of the background. Note that the global levels of chemoattractant are decreasing due to the expansion of the domain. The colour bars adjacent to each diagram show the scale for the chemoattractant concentration. White lines show the movement of evenly spaced x-coordinates due to domain growth. Cells and filopodia are shown in yellow. Parameters can be found in Table 2.1. The full video for this simulation may be found at http://tinyurl.com/cx966lh with the file name without_followers.avi.

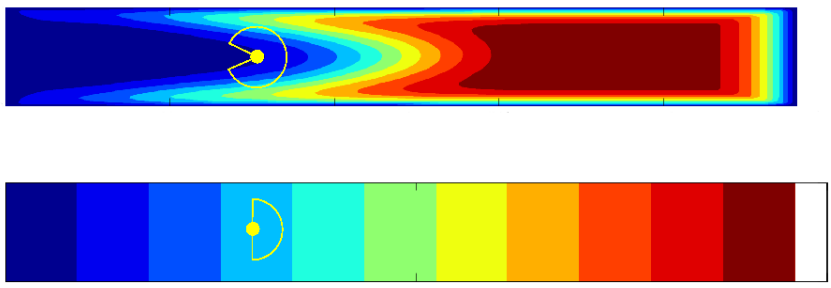


Figure 2.11: A cell (represented by a yellow disc) at the front of the migratory stream with dynamically determined chemoattractant gradient may move backwards towards $x = 0$, whilst in the fixed chemoattractant model, cells must always move towards the end of the domain. The outlined circular portions show the region of the domain in which the chemoattractant gradient is positive from the cell position.

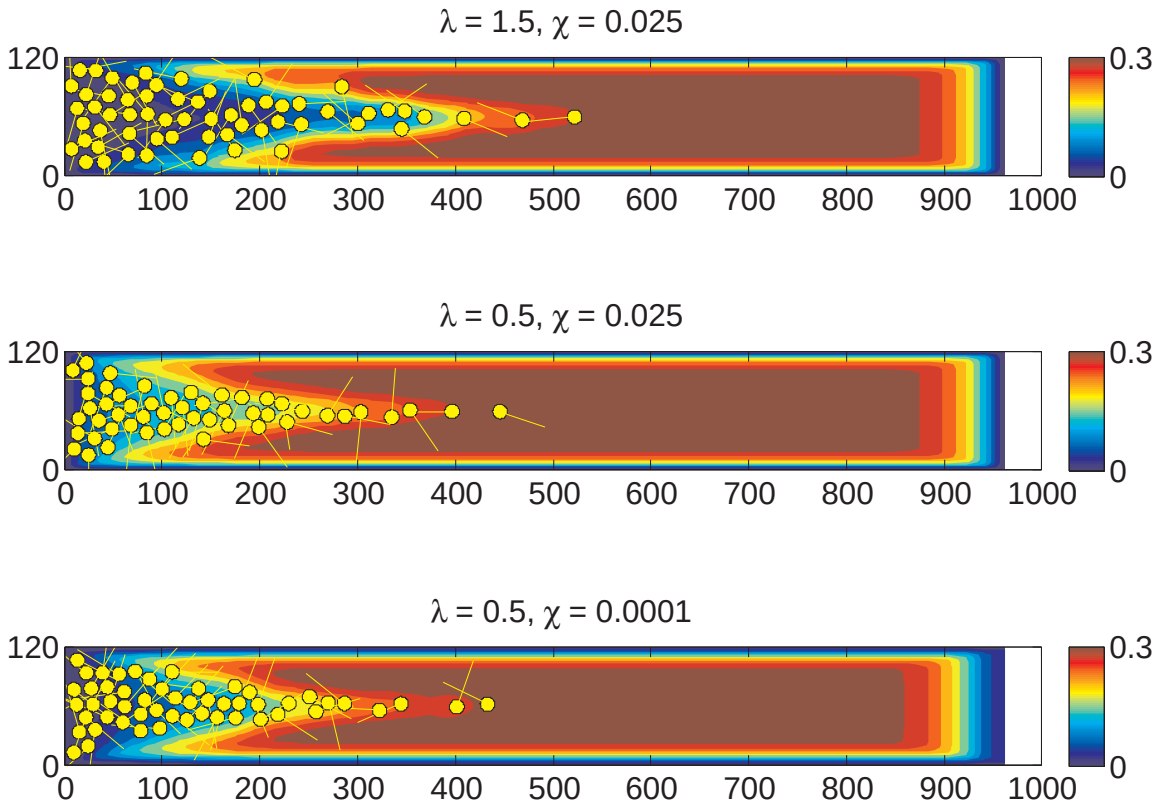


Figure 2.12: Increasing chemoattractant production (χ) or decreasing chemoattractant consumption (λ) does not prevent cells clustering at the entrance to the domain ($x = 0$).

can no longer give directional cues. This cell clustering is exacerbated by the general dilution of chemoattractant resulting from domain growth. In particular, we note that 24 hours into migration the maximum chemoattractant in the domain is approximately a third of the maximum when migration begins. Hence the cells split into two subpopulations: those at the front, where a gradient of chemoattractant allows directed movement; and those behind the front, where the levels of chemoattractant are uniformly low. Thus later emerging cells become trapped next to the neural tube and do not migrate any further into the domain.

The observation that a chemotactic system in which cells create the gradient they follow does not give a coherent migratory stream holds under a wide range of parameter regimes. In particular, reducing the cell consumption of chemoattractant (λ) or increasing the rate of production of chemoattractant (χ) does not stabilise the chemoattractant gradient for long enough to prevent cells clustering next to the neural tube (Figure 2.12).

2.6 Subpopulations

To enable later emerging cells to maintain directional information throughout the migration, we now include two subpopulations of cells. We hypothesise that those emerging early from the neural tube will continue to follow the chemoattractant gradient, as before. Later emerging cells, however, will obtain directional guidance from other migrating cells.

The choice of using subpopulations of chemotactic and communicative cells is supported by experimental evidence *in vivo* that the leading portion of the migratory stream is morphologically distinct to that further behind (Kulesa *et al.*, 2008; Teddy and Kulesa, 2004). Kulesa *et al.* (2008) observe cells moving in chains, with filopodia connecting the migrating cells. The aim of the extended model, therefore, is to set up chains of trailing cells that follow each other, with a leading cell at the head of each chain providing directional guidance from the underlying chemoattractant gradient. Since it is difficult to distinguish experimentally between other potential mechanisms for communication between the front and back of the migratory stream, our aim is not to determine the precise mechanism but to demonstrate a more realistic alternative to a purely chemotactic stream.

2.6.0.1 Trailing cell mechanism

The trailing population is defined as those cells that enter the domain after a set period of simulation time (t_f), representing the hypothesis that later emerging cells from the neural tube may be phenotypically different to earlier cells. Trailing cells continue to internalise the chemoattractant in the same way as is described in Section 2.5 and extend a filopodium in a random direction at each time step, as before. Instead of measuring the VEGF concentration, however, trailing cells respond to the presence of another cell body at any point along their filopodium. In particular, the shortest distance (d) from the line representing a cell filopodium ($\mathbf{x}$) to the centre of a leading cell or a chained cell ($\mathbf{y}$) is found by using $\mathbf{x} \cdot \mathbf{y} = |\mathbf{x}|\,|\mathbf{y}| \cos\theta$ and noting that if λ is the distance to the point where the shortest line and $\mathbf{x}$ meet (see Figure 2.13), then

$$\lambda = |\mathbf{y}| \cos\theta = \frac{\mathbf{x} \cdot \mathbf{y}}{|\mathbf{x}|}. \tag{2.12}$$

Hence

$$d = \left| \frac{\lambda}{|\mathbf{x}|}\mathbf{x} - \mathbf{y} \right|, \tag{2.13}$$

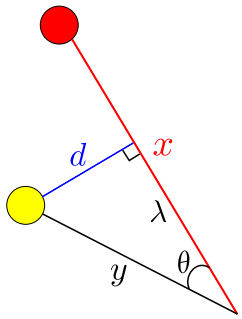


Figure 2.13: Finding the shortest distance (d) from any part of a cell filopodium ($\mathbf{x}$, shown as a red line) to any other cell centre ($\mathbf{y}$, shown as a yellow circle) by finding the distance (λ) from the origin to the point where the shortest line and $\mathbf{x}$ meet.

and if d is less than a cell radius then the trailing cell attaches to the cell at $\mathbf{y}$ and continues to move at each subsequent time step in the direction of that cell.

If a cell chain is found by the trailing cell filopodium, then the trailing cell attaches and continues to move at each subsequent time step in the direction of the attached cell. A cell that is following another is represented graphically by attaching its filopodium to the leading cell and is coloured white. If the two cells become separated during migration by more than the length of a filopodium then the following cell will detach and resume seeking other cells.

2.6.1 Simulations

When trailing cells move by seeking cell chains, early-emerging cells migrate a distance that is consistent with experimental observations (Figure 2.14). In the majority of simulations, later emerging cells successfully find a chain of cells and migrate into the domain. The chemoattractant profile is robust to changes in parameter values since the VEGF gradient is created by the cells. Hence if the cells migrate later in development, or not as quickly, or they consume the chemoattractant less quickly, then the chemoattractant gradient is still established at the correct time for migration. However, since non-chained trailing cells do not move (except for passive movement due to domain growth), cells that do not join a chain at the beginning of migration can occasionally disrupt stream migration. Sensitivity to stochastic fluctuations and the rate of cell sensing may therefore be rectifiable by allowing cells to move between the subpopulations of cells or by giving follower cells some random diffusive motion in addition to directed chain migration. Since the vast majority of wild-type simulations lead to successful migration, however, we will not pursue even greater robustness to stochastic fluctuations unless it should become necessary.

	Description	Reference	Notes	Value
R	cell radius	McLennan and Kulesa (2010)		7.5μm
v	cell speed	Kulesa *et al.* (2008)	lead cells	41.6±10 μm/h
			trailing cells	49.9±17 μm/h
D_c	VEGF diffusivity	unpublished data by Kulesa *et al.*		Low (taken to be 0.1μm/hr)
L_x	domain length	unpublished data by Kulesa *et al.*	beginning of migration time	300μm
			end of migration time	1100μm
L_y	height of the migratory domain		taken from maximum height of migratory stream	120μm
	height of the migratory stream	unpublished data by Kulesa *et al.*	beginning of migratory stream	100μm
			middle of migratory stream	40-50μm
			end of migratory stream	120μm
l_f	filopodium length	Teddy and Kulesa (2004)	longer filopodium at the leading edge	20μm - 70μm (taken as 50μm)
P_c	rate of cell exiting the neural tube	Kulesa *et al.* (2008)		2 cells every 15 mins
m	maximum number of initialisation attempts for each cell inserted into the domain	assumed		50
δ	related to the width of sensing and eating chemoattractant	McLennan and Kulesa (2010)	given as twice a cell radius	15μm
λ	rate of consumption of chemoattractant	assumed		1.5/hr
χ	rate of production of chemattractant	assumed		0.0001/hr
t_f	time after which cells entering the domain are 'followers' in the two-population model	Kulesa *et al.* (2008); Teddy and Kulesa (2004)	70% of cells are assumed to be 'followers'	7.25 hours
dt	time step length	assumed		0.05 hours

Table 2.1: Summary of model parameters.

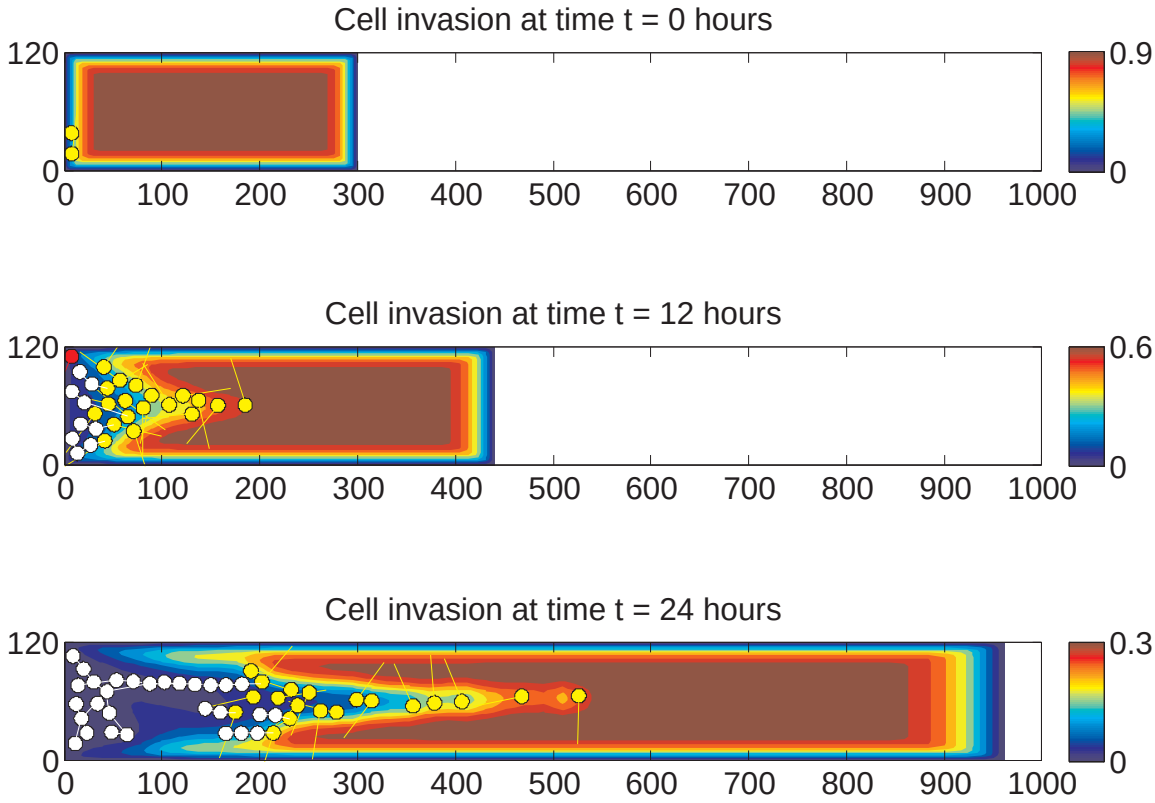


Figure 2.14: Representative simulation of the model described in Section 2.6. Chemoattractant concentration is denoted by the colour of the background. Note that the global levels of chemoattractant are decreasing due to the expansion of the domain. The colour bars adjacent to each diagram show the scale for the chemoattractant concentration. Leading cells (responding to the chemoattractant) and their filopodia are shown in yellow whereas following cells (responding to other cells) and their filopodia are shown in white. Parameters can be found in Table 2.1. The full video for this simulation may be found at http://tinyurl.com/cx966lh with the file name with_followers.avi.

2.6.2 Experimental response to model predictions

In response to our prediction that subpopulations of cells are required for successful migration and experimental observations of differing morphologies within the migratory stream, our collaborators undertook genetic profiling of the leading and trailing populations. When comparing the expression of 84 genes of interest, 19 were found to have significant differences using two different experimental methods of isolating cells (McLennan *et al.*, 2012). Leading cells have upregulated guidance and navigation genes whilst trailing cells express greater levels of cadherin 7 (McLennan *et al.*, 2012), which is associated with cell adhesion. Thus genetic profiling gives additional evidence for the existence of subpopulations within the migratory stream with differing migratory mechanisms.

2.7 Discussion

During this chapter we have developed a hybrid IBM for cellular migration in a growing domain. We have demonstrated that a chemoattractant gradient created dynamically by the cells is not stable enough to allow later emerging cells to successfully migrate into the domain, particularly for slower-diffusing chemoattractants. When trailing cells instead use a different mechanism of invasion, such as following other cells in static chains, we have shown they can form a successful migratory stream. These conclusions are also applicable to other systems with migrating populations and a spatially invariant production of chemoattractant, with or without domain growth (Puré and Cuff, 2001; Boldajipour *et al.*, 2008; Roberts and Brenchley, 2000; Dormann and Weijer, 2006; Lee *et al.*, 2008).

We justify our initial assumptions of precisely two separate populations and a fairly rudimentary cell following mechanism by noting that there is currently insufficient evidence to distinguish between other possible mechanisms. Our aim was not to discover the precise anatomy of the migratory stream, but instead to formulate a possible alternative to the currently hypothesised mechanisms of migration. We have demonstrated that a solely chemotactic population with only one chemoattractant cannot lead to a successful complete migratory stream. This motivated our experimental collaborators to test our prediction of at least two subpopulations of cells, and their genetic profile analysis revealed that the model prediction was correct.

In the next chapter we test our model by predicting the qualitative behaviour of cells under experimental perturbations to the wildtype system. The exact form of successes or failures in our model

predictions will enable us to elucidate the most important aspects of the model and ways in which it should be improved. In so doing, we will also gain more insight into the biology of cranial neural crest cell migration.

Chapter 3

Testing our modelling framework experimentally

In the previous chapter we developed a hybrid IBM that demonstrates successful cellular migration in a growing domain, corresponding to CNCC migration under wildtype conditions. The model consists of a continuum chemoattractant and two populations of discrete cells: leading cells, that move by chemotaxis; and trailing cells, that gain directional information from the leading cells. We will now use that model to predict the outcome of *in vivo* experiments and compare and discuss the results. Where our predictions do not correspond to the experimental results we will discuss possible drawbacks to our model and ways in which these can be overcome. The aim is to interrogate the biological system in greater depth through a combination of experimental and theoretical modelling work. We wish to gain greater insight into the mechanisms of migration and their relative importance. In particular, we will examine the interaction between cellular communication and guidance information gained from the chemoattractant gradient.

We emphasize that the experiments and analysis of experimental results were carried out in isolation from the modelling prediction work to eliminate any possible bias in result interpretation. The work forming Chapters 2 and 3 is published in *Development* (McLennan *et al.*, 2012) on which I am joint first author.

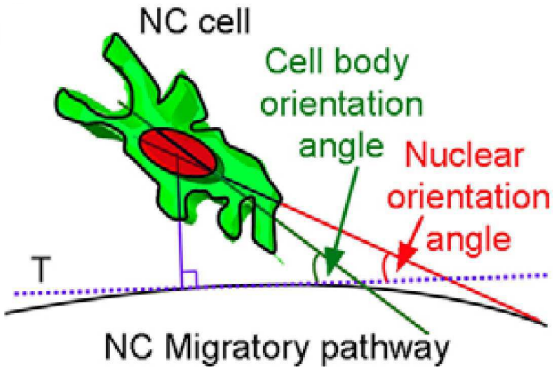


Figure 3.1: Cell orientations are determined by smallest angle between the long axis of the nucleus of the cell and the tangent, T, to the migratory route at the cell's position. Figure taken from McLennan *et al.* (2012).

3.1 Presentation of experimental data

Chicken embryos are used for all the experiments described here. GFP is introduced into CNCCs by electroporation (EP) to allow imaging of the cells at 8, 16 and 24 hours after EP. Average cell orientation angles are given by the smallest angle between the long axis of thenucleus of the cell and the tangent to the migratory route at the cell's position (Figure 3.1) and give a measure of how aligned the cells are to the migratory route. Smaller angles are indicative of cells that are more aligned to the migratory route, whilst a uniformly distributed population would have an average orientation angle of 45°. Data and model predictions are given in Figures 3.2 to 3.7. In particular, representative model simulations are shown in Figures 3.2, 3.4 and 3.6 at 12 and 24 hours into migration. Experimental data are given as representative images (Figures 3.3a, 3.5a, 3.7a) and average cell orientation angles profiled at varying distances along the migratory pathway (Figures 3.3b, 3.5b, 3.7b). The orientation angle profile of a wildtype stream after 24 hours of migration is shown in blue in Figure 3.3. The stream displays a characteristic profile, with higher average angles near the neural tube, becoming more aligned to the pathway towards the middle of the migratory stream. Cells at the leading edge of the migratory population at 16 and 24 hours also have high average angles, indicating reduced alignment to the migratory route.

We note that the embryos are electroporated five hours before the migration of CNCCs begins. Hence all experimental timings must be adjusted accordingly when simulating *in silico*.

We will now investigate a series of experiments, using the model developed in Chapter 2 to predict outcomes and interpret results.

3.2 Ablation experiment

3.2.1 Experimental description

To test whether the migration of lead cells is reliant on communication with the trailing cell population the neural tube is removed from the embryo at ten hours after EP, thus preventing the entry of the trailing cell population into the migratory domain.

If the leading cells require communication with trailing cells then we would expect their migration to be disrupted and unsuccessful. Conversely, if the cells respond to microenvironmental signals and have no direct feedback from the rest of the cellular population then we would expect the cells to migrate exactly as in the wildtype embryo so that a small subsection of the migratory stream is represented. Our model presents a third option: that there is no direct communication of information from the trailing cells to the leading cells, but that the presence of the trailing cells can indirectly affect the leading cells through the chemoattractant profile.

3.2.2 Model prediction

The ablation experiment is simulated *in silico* using our model (see Chapter 2) by ceasing to insert new cells into the domain five hours into the simulation. All the cells in the simulation are therefore of the 'leading' cell profile and move by chemotaxis. Our model predicts that the lower numbers of cells will lead to an uneven distribution of chemoattractant across the width of the domain (Figure 3.2). The higher levels of chemoattractant at the sides of the domain lead to greater exploration of the width of the domain so that each cell moves less quickly in the positive x-direction and the cells spread out along the length of the domain. Thus the migratory stream qualitatively resembles the wildtype case, in that the cells are spread throughout the domain rather than recapitulating the small leading subsection of the stream.

3.2.3 Experimental result

Experimental data are shown in Figure 3.3 with wildtype data shown in blue for reference (Figure 3.3b). Cells in ablated embryos spread out along the migratory route, with trailing cells remaining further back than the equivalent cells in a wildtype embryo. Despite only around a third of the cells entering the domain, therefore, the stream is found to inhabit more than half of the usual

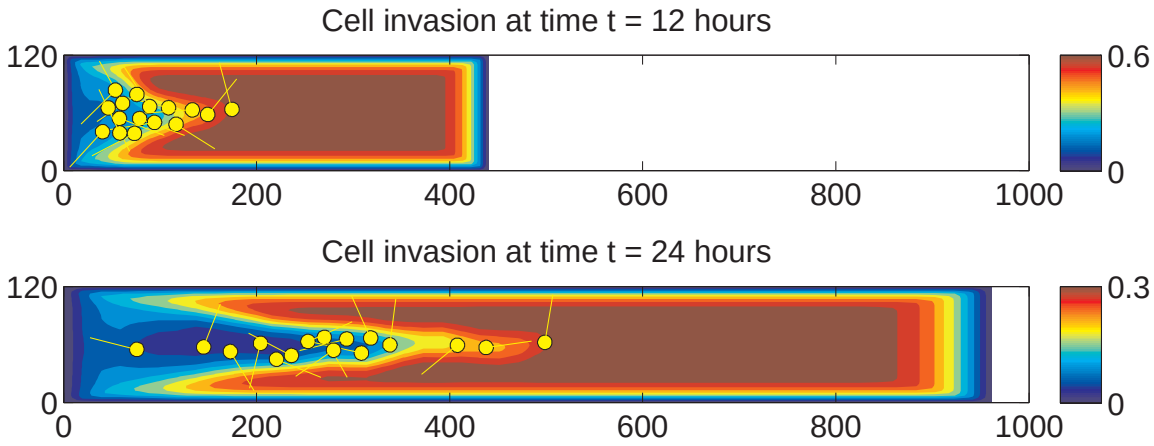


Figure 3.2: The model predicts that ablation of the neural tube five hours after the start of migration will lead to cells spreading out along the domain due to a slower depletion of chemoattractant (see Sections 3.2.1, 3.2.2 and 3.2.3). The full video for this simulation may be found at http://tinyurl.com/cx966lh with the file name ablation.avi.

route after 24 hours (Figure 3.3b). Thus our model prediction of a spread of cells due to slower consumption of the chemoattractant is consistent with experimental data. Although we assume no direct communication from the trailing cells to the leading cells, there is an indirect effect of removing the trailing population, which is mediated through a lower consumption of chemoattractant. Since there are also fewer cells, a smaller population pressure with less occupied space leading to less biasing of movement towards the positive x-direction also contributes towards a slower migration rate *in silico*.

We note that whilst under wildtype conditions there is little cell proliferation *en route* to the destination, when the population is reduced by ablation the cells multiply much more rapidly even whilst migrating. However, there is very little experimental data as to the mechanisms and profile of cell proliferation such as how often cells proliferate, whether some types of cells divide more quickly and the effect of cell division on the speed and direction of cell motion. Whilst we could test particular hypotheses, it would be difficult to make meaningful predictions within our model without further experimental data, since so much of the proliferation mechanism would need to be postulated without supporting data. We therefore leave cell proliferation to future work.

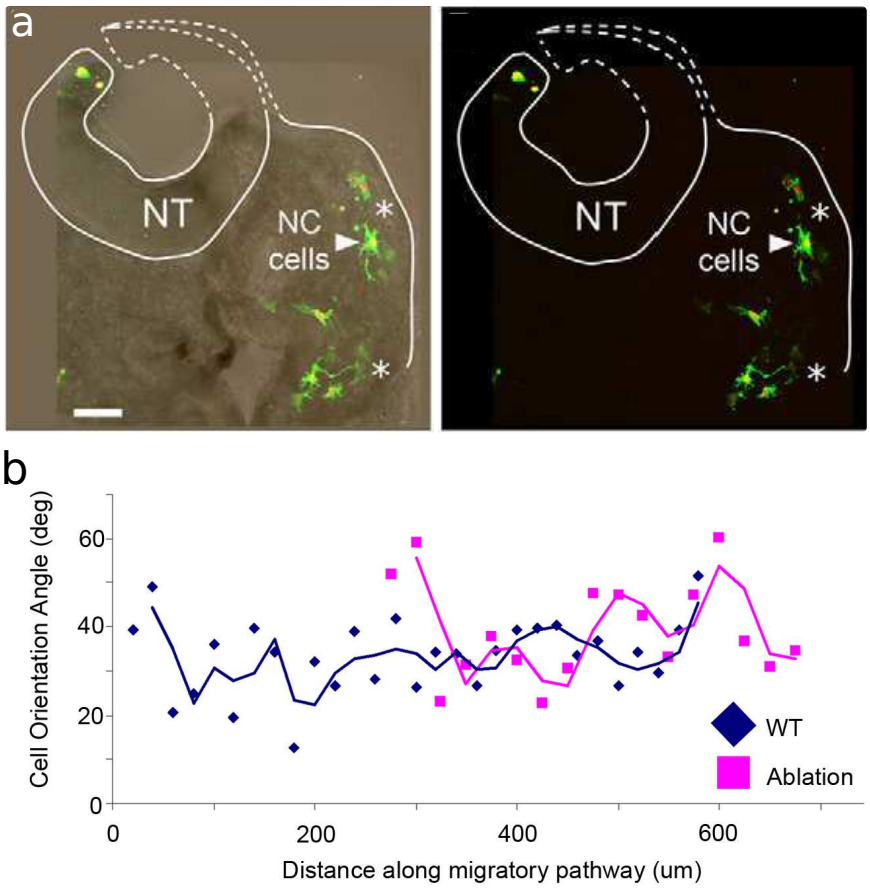


Figure 3.3: Experimental results for an ablated migratory stream (see Section 3.2.3): (a) Representative images 24 hours into migration; (b) Cell orientation angles profiled at varying distances along the migratory pathway.

3.3 Transplant experiment: front-to-back

3.3.1 Experimental description

We have seen that the leading cells do not require communication with the trailing portion of the stream to migrate into the domain. Next we wish to test the response of leading cells to a new environment and altered positions within the stream. Cells are given new positions within the stream by transplanting the front region of the migratory route from a donor embryo into the neural tube of a host embryo at the same level of development. The host embryo is electroporated with GFP as before, whilst the donor embryo cells are marked with red fluorescent protein (RFP) so that the cells may be tracked.

The neural tube forms the entrance to the migratory domain, in that there is a large population of CNCCs within the neural tube that subsequently migrate out into the tissue. Implanting donor tissue into the neural tube also prevents new host cells migrating into the domain, so that the host population is reduced. Without considering the mathematical model, we would expect that transplanted leading cells may maintain their initial ordering with the host population so that they migrate behind the trailing population of host cells. Another possibility is that the transplanted cells may somehow be able to signal to the host cells that they are leading cells and overtake the host cells to lead the migratory stream.

3.3.2 Simulation method

To make a prediction for this experiment using our model, we must consider how to represent a tissue transplantation within our model framework. Since transplantations involve removing a section of tissue, complete with overlying ectoderm, which produces VEGF, we assume that the chemoattractant is moved with the cells. To make a model prediction for the front-to-back experiment, therefore, we simulate the donor embryo for ten hours and record the positions of cells and the levels of chemoattractant in a region 100μm along the x-axis across the width of the domain. The location of this region is user defined and may vary to ensure that the leading cells are taken. To capture the chemoattractant profile at the interface of the transplant, we linearly interpolate to a finer mesh near the region boundaries.

Another embryo is then simulated for ten hours and the donor region with cells is input into the region $x < 0$ to mimic the experimental transplantation into the neural tube. The host system with

transplanted tissue is simulated for a further 14 hours to give a total migration time of 24 hours. No new cells are input into the domain after the transplantation has occurred, since host cells in the neural tube can no longer reach the migratory domain. To enable numerical solution of the resulting chemoattractant profile over the whole domain we smooth the boundary between donor and host tissues. Smoothing is achieved by interpolating between the first grid point into the donor tissue and the first point into the host tissue to create a finer mesh at the interface. We assume that the transplanted tissue does not grow during the time simulated, but we note that if growth is simulated in the transplanted tissue then this does not qualitatively change our results.

3.3.3 Model prediction

Simulations of the front-to-back transplant experiment predict that donor cells will usually not migrate out of the transplant tissue (Figure 3.4). The higher levels of chemoattractant present at the front of the donor migratory stream, that are subsequently transplanted with the cells, ensure that the environment in the transplant is more favourable than that at the back of the host stream. The only situation in which cells may migrate out of the transplant, therefore, is if the host chemoattractant is unusually high at the back of the stream. Higher chemoattractant could occur if the host is at an earlier stage of migration than the donor, leading to less total chemoattractant consumption. Alternatively, if the host embryo has grown less than the donor then chemoattractant dilution would be reduced and maximum chemoattractant concentrations may be higher. In either case chemoattractant concentration in the transplant may be marginally lower than that in the host, and thus allow a small amount of migration. In most cases, however, we predict that the relative chemoattractant concentrations will not allow cellular migration out of the transplanted tissue.

3.3.4 Experimental result

Our model prediction, that no transplanted cells will migrate, is completely different to our initial "naïve" predictions: that the transplanted cells will migrate at the back or at the front of the host stream. We test our predictions, by considering the experimental data.

In ten out of 13 embryos the cells transplanted to the neural tube did not migrate out of the implanted tissue, in agreement with our model predictions. The data presented in Figure 3.5 are therefore taken only from the three embryos in which cellular migration occurred. When donor cells did migrate away from the transplanted tissue they integrated into the trailing region of the host

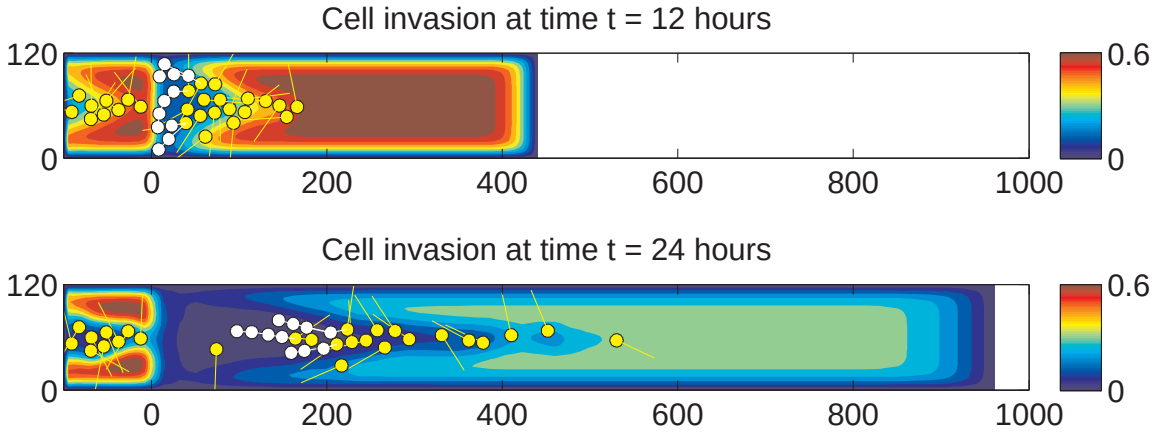


Figure 3.4: The model predicts that donor cells transplanted from the front of a migratory stream to the neural tube of a host embryo will usually not migrate out of the transplanted tissue (see Sections 3.3.1, 3.3.3 and 3.3.4). The full video for this simulation may be found at http://tinyurl.com/cx966lh with the file name front_to_back.avi.

migratory stream and did not regain their previous positions at the head of the stream. The front-to-back experiments thus support our hypothesised migration mechanisms and show that when the cells do successfully leave the transplant they can migrate alongside the host trailing cells.

Since only three of the 13 embryos considered had any cellular migration at all, our collaborators were initially hesitant to publish the data as they regarded the experiments as failed. It was only after in-depth discussions about the model that they were convinced that the results were not inconsistent with our original assumption and in fact supported our hypotheses. Hence without modelling work the data would not have been published and could not have been used to gain greater understanding of the biological system.

3.4 Transplant experiment: back-to-front

3.4.1 Experimental description

In previous sections we have tested the migratory properties of leading cells, and the communication between leading and trailing cells. We now extend this to consider the response of trailing cells to environmental and positional changes, by transplanting the trailing portion of a donor migratory stream ahead of the migrating cells in a host embryo. We aim to discover whether the transplanted trailing cells will migrate in the lead of the host stream, or regain their previous positions behind the host migratory cells.

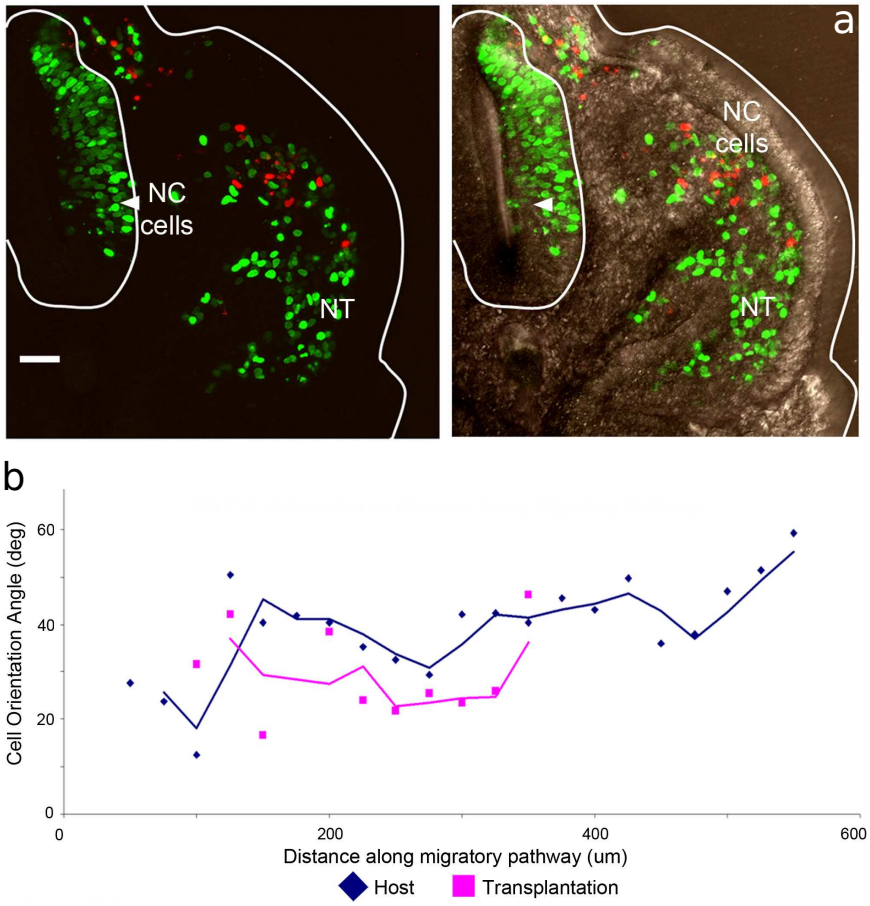


Figure 3.5: Experimental results for cells and tissue transplanted from the front of a donor migratory population to the neural tube of the host embryo (see Section 3.3.4): (a) Representative images 24 hours into migration; (b) Cell orientation angles profiled at varying distances along the migratory pathway.

The back-to-front transplant experiment requires that the cells are taken from the trailing subpopulation, and that there is enough space ahead of the migratory stream in the host embryo. In contrast with the previous transplant experiment, therefore, the donor and host embryos must be at different developmental stages. In particular, the donor embryo is allowed to develop to a later stage than the host embryo, and this has implications for the extent of chemoattractant dilution.

3.4.2 Simulation method

The back-to-front transplant experiment is simulated in a similar way to previous transplantation experiments (Section 3.3.2). In particular, we assume, as before, that the chemoattractant is transplanted along with the cells and tissue.

The donor embryo is simulated for 14 hours before the chemoattractant and cells in a region 70μm along the x-axis and across the width of the domain are recorded. A host embryo is then simulated for ten hours and the region is implanted into a user-defined place ahead of the migrating cells. It is assumed that the region is implanted over the host tissue in that area, so that the host chemoattractant is overwritten with the donor chemoattractant in that region. As in Section 3.3.2, the chemoattractant concentration at the interface between host and donor tissue is smoothed so that the chemoattractant profile may be solved numerically. The host system with transplanted tissue is simulated for a further 14 hours to give a total migration time of 24 hours.

3.4.3 Model prediction

Simulations of back-to-front transplants predict that no cells will move forwards out of the transplanted region (Figure 3.6). The trailing cells taken from the donor embryo will only actively migrate if they are in a chain of cells with a chemotactic leader at the front. Since there are no leading cells ahead of the transplanted cells, the donor cells cannot form chains and will not migrate further along the host migratory route. If the leading host cells reach the transplant site, then the transplanted cells may even move backwards as they follow host leading cells behind them.

In addition, the chemoattractant levels in the transplant are much more dilute, since the donor embryo was at a later stage of migration than the host. Indeed, since the transplant is taken from the back of the migratory stream where there has been higher cell consumption, the chemoattractant is much lower than the front of the host migratory stream. Hence the transplant forms a barrier

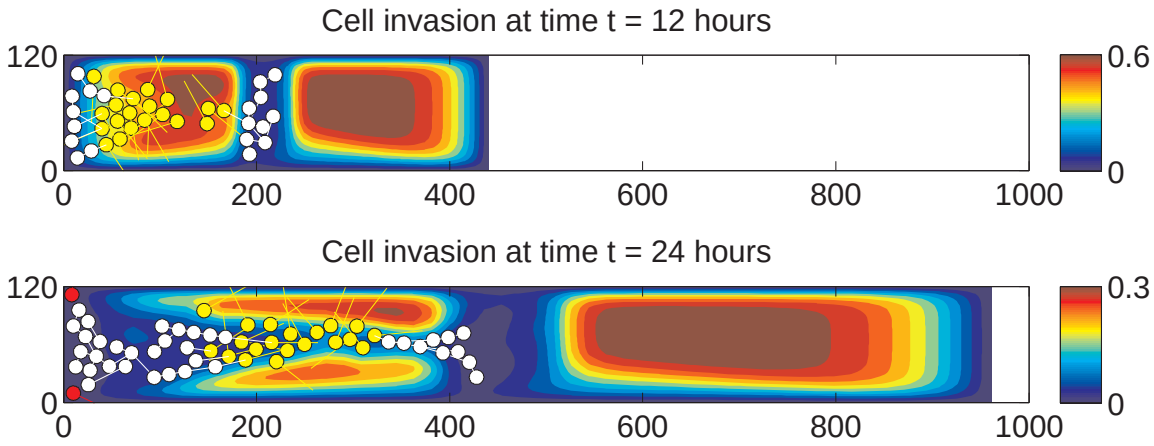


Figure 3.6: The model predicts that cells transplanted from the back of a donor migratory stream in front of host migrating cells will not migrate out of the transplanted tissue, and will prevent host cells from migrating to the end of the domain (see Sections 3.4.1, 3.4.3 and 3.4.4). The full video for this simulation may be found at http://tinyurl.com/cx966lh with the file name back_to_front.avi.

across the domain, trapping leading host cells and preventing further migration.

Our model prediction, therefore, is that the transplanted cells will not migrate out into 'unexplored' host tissue, although they may migrate backwards into the leading host population. The leading host population itself will not be able to migrate through the transplanted tissue and will become trapped behind the transplant.

3.4.4 Experimental result

In contrast to our model prediction, trailing CNCCs transplanted in front of the migratory stream successfully migrate out of the transplanted tissue and into the branchial arch (Figure 3.7). Host cells successfully migrate past the transplant site whilst many donor cells move out into the host tissue. Moreover, the cell orientation profile of the donor cells mimics that of a wildtype migratory stream, with the latter-most donor cells displaying reduced alignment to the migratory route. We note that in the three-dimensional biological system it is possible for cells to migrate around the transplant whilst remaining entirely within host tissue. Thus the transplant does not 'block' the migratory route as it does in our two-dimensional model

In addition, gene expression data reveal that the donor cells after transplantation are similar to wildtype leading cells leading us to conclude that trailing donor cells may alter their mechanisms of movement to those of leading cells. Accordingly, in Section 3.5 we will include conversion between cell types and explore the effects of this on our modelling predictions.

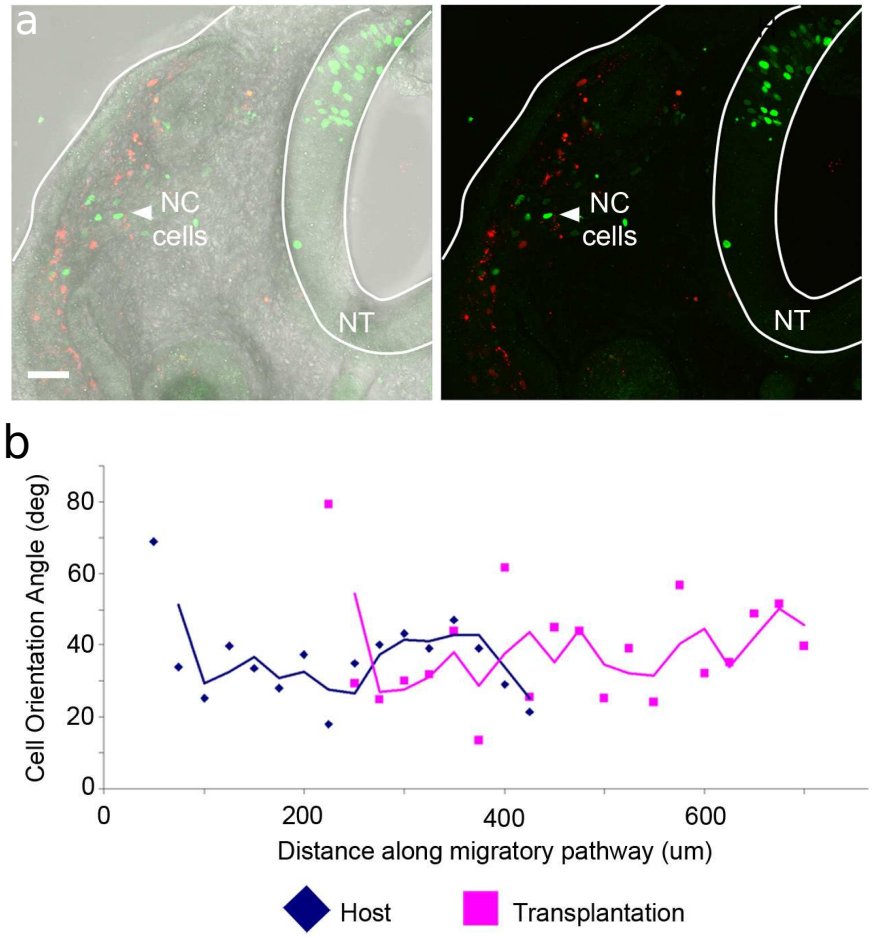


Figure 3.7: Experimental results for cells and tissue transplanted from the back of a donor migratory population to the front of the host migratory stream (see Section 3.4.4): (a) Representative images 24 hours into migration; (b) Cell orientation angles profiled at varying distances along the migratory pathway.

3.5 Population conversion

We previously found that cells may change their genetic profile in response to environmental changes. In addition, there is evidence that cells may modify their behaviours in response to environmental perturbations. For example, when a physical barrier is introduced across the migratory pathway leading cells become trapped whilst those behind change their route to move around the barrier and lead the new migratory stream (Kulesa *et al.*, 2005).

In this section we will include potential cell conversion mechanisms into our model with the aim of explaining the discrepancy between the modelling predictions and experimental results for the back-to-front transplant experiment. The circumstances under which cells may convert between the subpopulations are not well established and so we will present two possible ways of modelling cell conversion: time-based conversion; and gradient-based conversion. For each mechanism we will first check that wildtype migration, without experimental perturbations, still produces successful migration, before running the back-to-front experiment simulations. We will also change the back-to-front experiment simulations by reducing the size of the transplanted tissue, so that host cells may migrate around the area. Another way to achieve this would be to extend the model into three dimensions, but it is unlikely that this would lend greater insight into the system and would significantly increase the time required to simulate the system.

3.5.1 Time-based cell conversion

The first conversion mechanism that we will consider is time-based cell conversion; where cells change populations when their current mechanism is unsuccessful at producing active movement. If a trailing cell does not find a chain to follow or a leading cell does not find a favourable chemoattractant gradient after some time then that cell converts to the other subpopulation. Under wildtype conditions this results in a small number of conversions between types and hence a similar output to the original model with a slightly reduced migratory distance, so long as the time required to convert is not too short (Figure 3.8). For shorter conversion times, most cells convert to the leading subtype and migration breaks down (result not shown). In our simulations we take 10 time steps (0.5 hours) as the time required to convert.

However, the population using time-based conversion is less spatially segregated into subpopulations, particularly at earlier time points. When we simulate the back-to-front experiment, therefore, it is

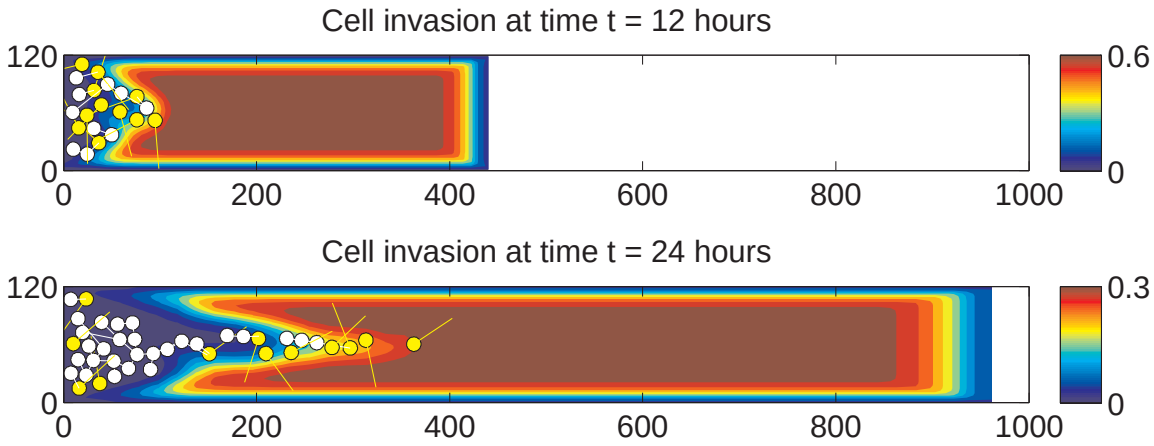


Figure 3.8: Representative simulation of the model described in Chapter 2 with cell conversion between subpopulations dependent on time spent inactive (see Section 3.5.1). Chemoattractant concentration is denoted by the colour of the background. Note that the global levels of chemoattractant are decreasing due to the expansion of the domain. The colour bars adjacent to each diagram show the scale for the chemoattractant concentration. Leading cells (responding to the chemoattractant) and their filopodia are shown in yellow whereas following cells (responding to other cells) and their filopodia are shown in white. Parameters can be found in Table 2.1. The full video for this simulation may be found at http://tinyurl.com/cx966lh with the file name with_followers_convert_1.avi.

difficult to transplant a population of cells that are solely of the trailing subtype. If a mixture of leading and trailing cells are transplanted, then the leading-type cells migrate out of the transplant in both the positive and negative x-directions (Figure 3.9). Trailing donor cells form chains behind the migrating leading donor cells so that all donor cells leave the transplant. However, host cells do not migrate past the transplant as is seen experimentally and the simulated migratory stream lacks cohesion, instead forming clumps of migrating cells.

3.5.2 Gradient-based conversion

Another possible approach to cell conversion is to find an environmental characteristic that localises to the different areas of the migratory stream. One example would be the local chemoattractant profile, which is more homogeneous towards the back of the stream, with greater local variation at the migratory front (see Figure 3.10). To find the local chemoattractant profile in our model, cells sample the chemoattractant gradient in multiple directions away from the cell body. If the gradient is favourable in a high proportion of directions, then the cell will convert towards a leader subtype. Conversely if the gradient is only favourable in a few directions then the cell will convert to a trailing subtype. To prevent cells continually changing back and forth between types, the proportion needed to convert from leading to following must be lower than that needed to convert from following to

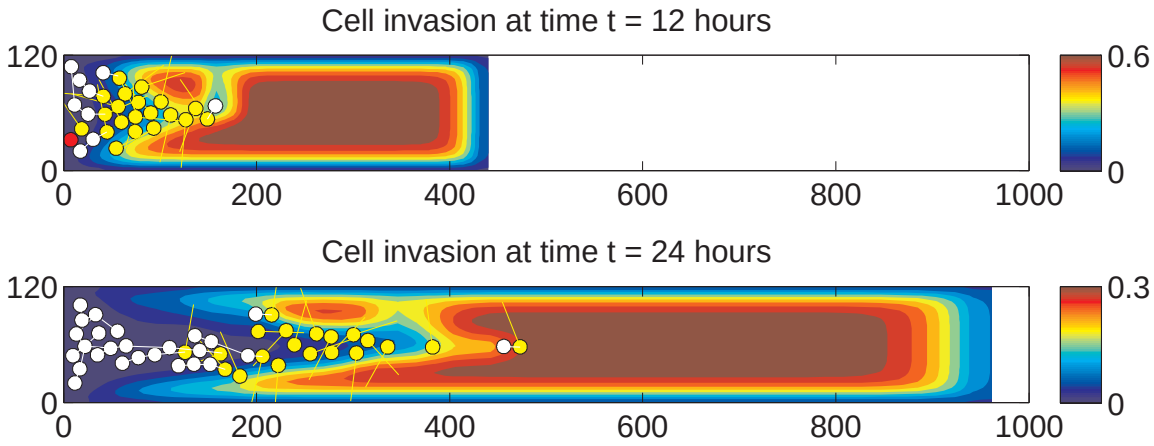


Figure 3.9: When time-dependent cell conversion is included into the model donor cells in the back-to-front transplantation experiment migrate out of the transplant in both directions along the migratory route. Leading cells are yellow and chained trailing cells are white. Trailing cells that are yet to find a chain are shown in red. The full video for this simulation may be found at http://tinyurl.com/cx966lh with the file name back_to_front_convert_1.avi.

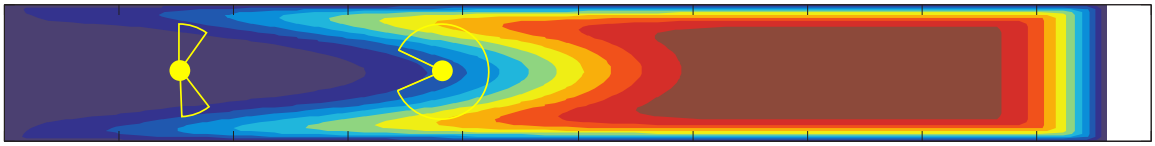


Figure 3.10: A cell (represented by a yellow disc) at the front of the migratory stream is in a local valley of chemoattractant, with a positive chemoattractant gradient in high proportion of directions. Conversely a cell towards the rear of the stream is in a more homogeneous profile of chemoattractant, with positive gradients in only a few directions. Cells may use this difference in profile as an indication of position within the migratory stream (see Section 3.5.2).

leading. In our simulations, cells check 16 directions and will convert from a leader to a follower if the gradient is favourable in less than 10% of directions. Conversely, if the gradient is favourable in more than 50% of directions a follower will convert to a leader.

As with time-based conversion (Section 3.5.1), there is little cell conversion under wildtype conditions and thus the simulations are similar to the original model (Figure 3.11). However, for successful migration, careful tuning of the parameters is needed, indicating that gradient-based conversion is less robust to stochastic variation in cellular behaviour than time-based conversion. Since biological systems in general, and developmental systems in particular, usually need to display robustness to stochastic variations, we take this as evidence against gradient-based cell conversion.

When a reduced transplant, taking only a small fraction of the domain width, is combined with gradient-based cell conversion the transplanted donor cells convert to the leading cell type. The newly formed leading donor cells then migrate out of the transplant, and spread out along the

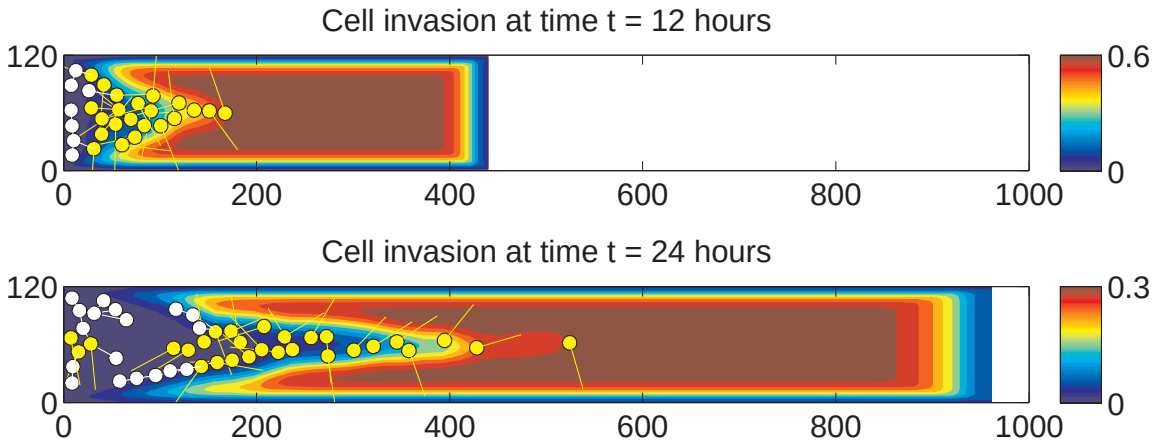


Figure 3.11: Representative simulation of the model described in Chapter 2 with chemoattractant gradient-dependent cell conversion between subpopulations (see Section 3.5.2). Chemoattractant concentration is denoted by the colour of the background. Note that the global levels of chemoattractant are decreasing due to the expansion of the domain. The colour bars adjacent to each diagram show the scale for the chemoattractant concentration. Leading cells (responding to the chemoattractant) and their filopodia are shown in yellow whereas following cells (responding to other cells) and their filopodia are shown in white. Parameters can be found in Table 2.1. The full video for this simulation may be found at http://tinyurl.com/cx966lh with the file name with_followers_convert_2.avi.

latter part of the migratory domain. The host cells migrate alongside the transplanted tissue and together with donor cells form a successful migratory stream, in accordance with experimental results (Figure 3.12).

3.5.3 Ablation experiment

Examining the ablation experimental data in more detail (Section 3.2.3 and Figure 3.3) reveals that the ablated migratory population displays a strikingly similar cell orientation profile to the full wildtype stream, with the back portion of the stream showing reduced alignment to the migratory route. Molecular profiling of the back portion of the stream demonstrates that the profile is closest to that of wildtype trailing cells whereas the profile of the front portion of the ablated stream remains similar to wildtype leading cells. We hypothesise, therefore, that cell conversion could enable the ablated migratory stream to compensate for the removed trailing cells by converting later emerging leading cells into a trailing phenotype. Indeed, genetic profiling data reveal that trailing cells in the ablated migratory stream are more genetically similar to leading cells in wildtype embryos.

With time-dependent conversion, approximately half of the population convert to the trailing subtype after 24 hours of migration (Figure 3.13). The migratory stream is still spread out along the domain, but there are fewer cells near the entrance to the domain, since cells that remain near the

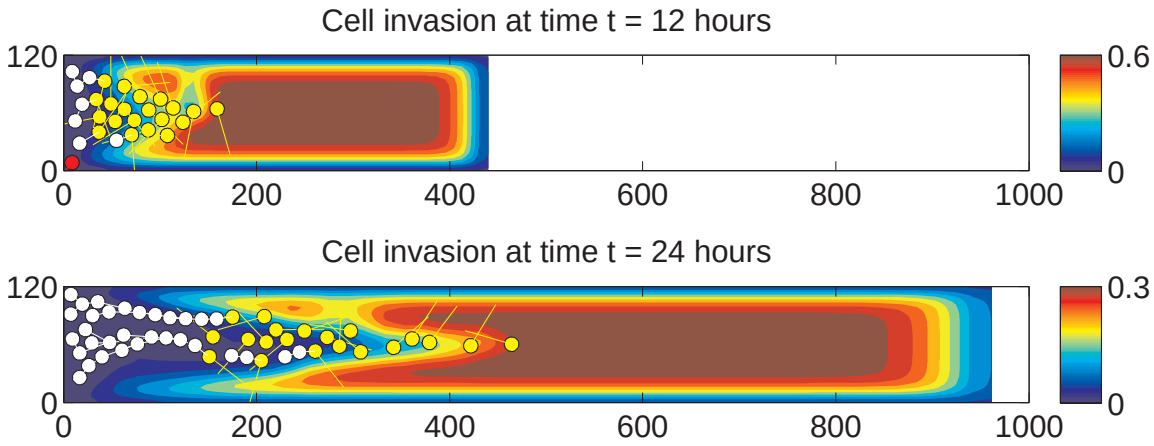


Figure 3.12: When gradient-dependent cell conversion is included in the model, cells in a reduced transplant from the back of a donor migratory stream in front of host migrating cells migrate out of the transplanted tissue. Host cells successfully migrate past the reduced transplant, as is found experimentally (see Section 3.4.4). The full video for this simulation may be found at http://tinyurl.com/cx966lh with the file name back_to_front_convert_2_reduced_transplant.avi.

neural tube are in a chemoattractant-depleted area and hence convert to follower cells. Follower cells may then attach to a moving cell chain and migrate further into the domain. The lack of cells near the domain entrance is in accordance with experimental data, which display a gap between the neural tube and the latter-most cells in the ablated stream (Figure 3.3).

Conversely, if cell conversion is gradient-dependent then no cells convert during an ablated migration. The reduced cell numbers result in the cells migrating in a narrow valley of consumed chemoattractant, so that the gradient is favourable in almost all directions. Hence the proportion of directions that have a favourable gradient can no longer be used as a signal of position within the stream. Thus the ablation experiments show evidence against the gradient-dependent conversion mechanism.

3.5.4 Generating a spatially structured migratory stream from an initially homogeneous population

Whether a cell's genetic profile and morphology is pre-determined as it enters the migratory domain or cells use environmental cues to determine their behaviour is an open question in the field. We may investigate this in our model by considering whether a heterogeneous population with leading phenotype cells at the front and trailing phenotypes at the back can be recovered if all cells enter the domain as a single phenotype.

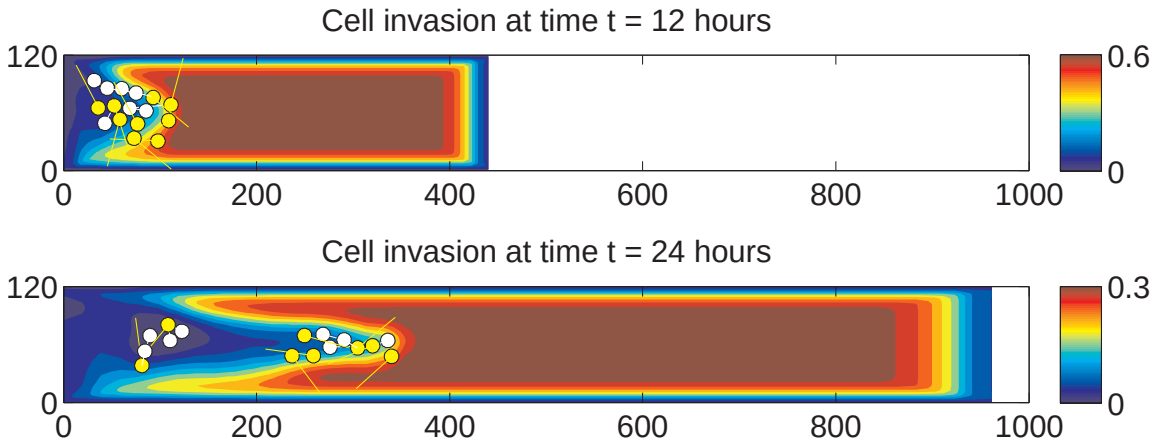


Figure 3.13: The extended model with cell conversion after a sustained failure to move (Section 3.5.1) gives a more heterogeneous population of cells whilst maintaining a spread out migratory stream (see Sections 3.2.1, 3.2.2 and 3.2.3). The full video for this simulation may be found at http://tinyurl.com/cx966lh with the file name ablation_convert_1.avi.

3.5.4.1 Time-based cell conversion

We first consider time-dependent cell conversion, where cells convert after a period of time of unsuccessful attempted movements. If all cells enter the domain with a leading phenotype then cell conversion does generate a heterogeneous migratory population (Figure 3.14). The migratory stream travels less far into the domain, however, and is not spatially well segregated, with leading cells present all the way through the stream.

Conversely if all cells are initially of the trailing subtype, then there is much less cell conversion during migration (Figure 3.15). However, cells still migrate successfully into the domain, and move further than in the initially-all-leading model. This model demonstrates that very few leading cells are required for successful migration, if the leading cells are in good positions to give directional information to the rest of the migratory stream. Using a conversion-based model is one way of ensuring a greater robustness to stochastic fluctuations in where cells enter the domain and how long it takes trailing cells to find other cells to provide directional guidance.

3.5.4.2 Gradient-based conversion

Now we consider an initially homogeneous population with gradient-based conversion, where cells convert based upon the shape of the chemoattractant profile near the cell. We assume in this model that the chemoattractant profile at the back of a migrating stream will be flat across the width of the domain, so that there are few directions in which the chemoattractant is favourable.

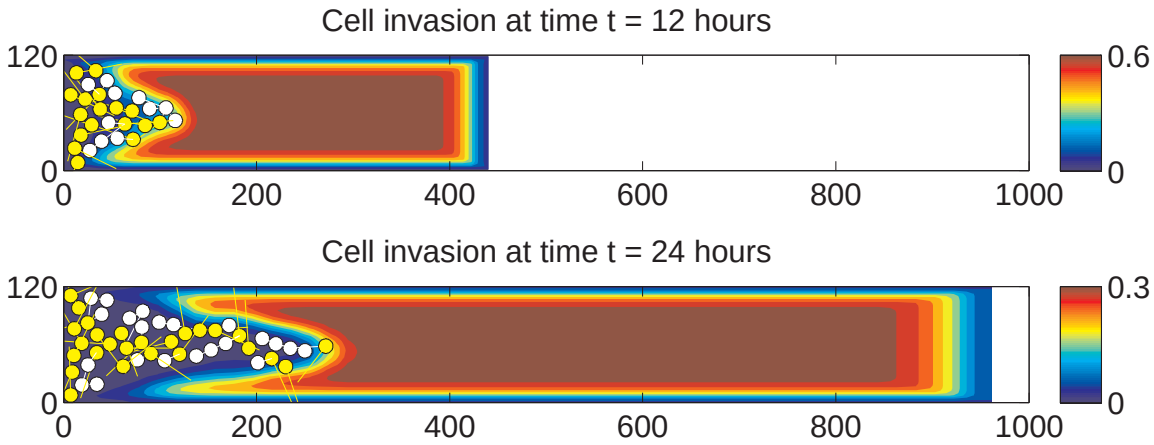


Figure 3.14: If all cells are initially of the leading subtype then time-based conversion can generate a heterogeneous migratory population but the subpopulations are not spatially segregated and the stream migrates less far than is seen experimentally. The full video for this simulation may be found at http://tinyurl.com/cx966lh with the file name without_followers_convert_1.avi.

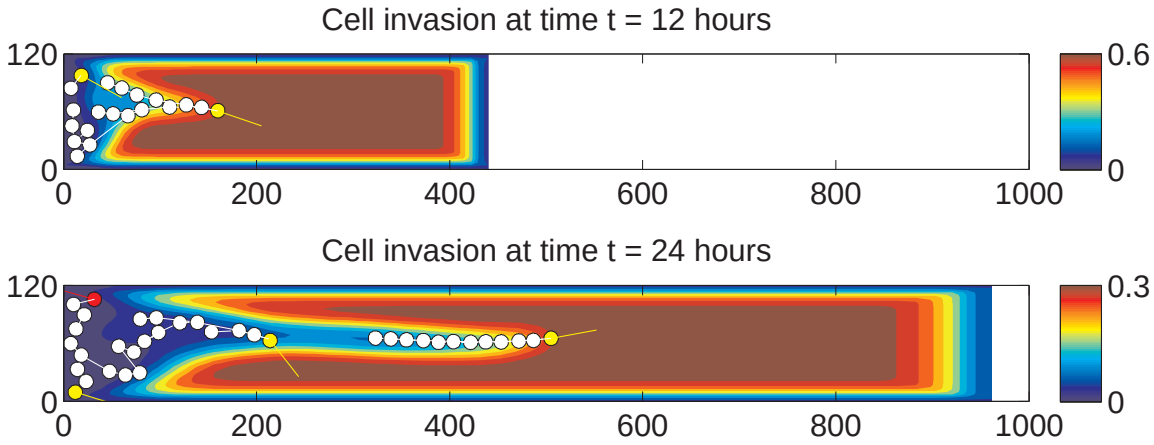


Figure 3.15: If all cells are initially of the trailing subtype then time-based conversion can generate a heterogeneous migratory population with very few cells of the leading subtype. Trailing-type cells successfully find chains and the stream migrates almost as far as is seen experimentally. The full video for this simulation may be found at http://tinyurl.com/cx966lh with the file name all_followers_convert_1.avi.

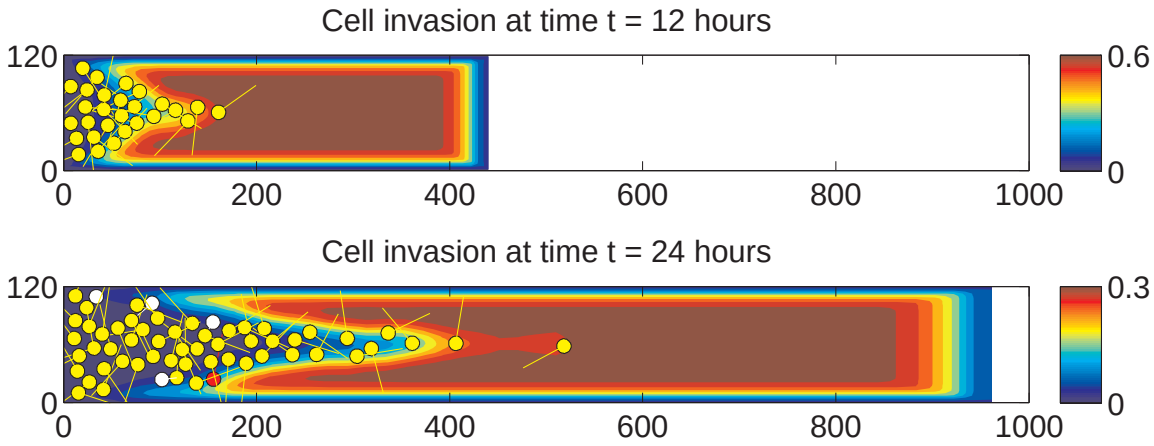


Figure 3.16: If all cells are initially of the leading subtype then gradient-based conversion generates a heterogeneous population, but with very few trailing cells. There is some clustering next to the domain entrance. The full video for this simulation may be found at http://tinyurl.com/cx966lh with the file name without_followers_convert_2.avi.

Whilst this is true in a mature migratory stream, initially all cells are in similar chemoattractant profiles, and the value at which conversion should occur to create spatially segregated populations changes throughout migration. As a result, attempting to initialise a migratory stream as either all of leading subtype (Figure 3.16) or all of trailing subtype (Figure 3.17) does not reproduce a spatially segregated stream. Whilst cells do migrate, there is some clustering next to the domain entrance, and there are very few cells of the trailing subtype at the end of the simulation. We conclude that gradient-based cell conversion of this type is not sufficient to use with a homogeneous initial population without further refinements to the model.

3.6 Discussion and conclusions

This chapter tested our model hypotheses and assumptions via a series of experiments. With the exception of model extensions in response to experimental results (Section 3.5), all modelling predictions were made prior to the experiments being carried out to prevent bias in our analysis. We demonstrated that dynamic feedback between cells and chemoattractant results in a system that is robust to a reduced population of cells, in that the remaining cells migrate successfully and colonise the length of the domain. With the addition of conversion between cell types, the phenotypic profile of the ablated stream is also reproduced.

The second experiment transplanted CNCCs and tissue from the leading portion of a migratory

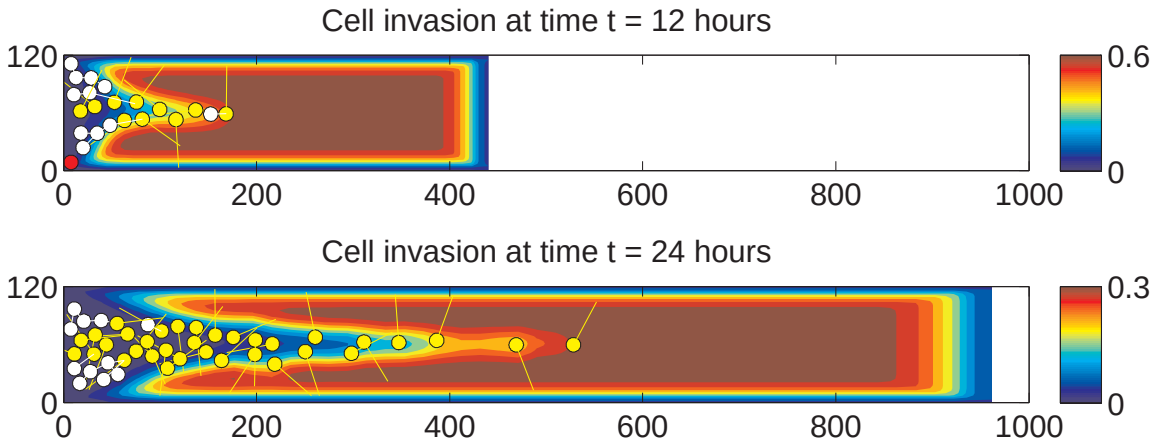


Figure 3.17: If all cells are initially of the trailing subtype then gradient-based conversion generates a heterogeneous population, but with very few trailing cells. There is some clustering next to the domain entrance. The full video for this simulation may be found at http://tinyurl.com/cx966lh with the file name all_followers_convert_2.avi.

stream to the neural tube of a host embryo. We predicted that the cells would not leave the transplanted tissue, since chemoattractant levels are higher in the transplant than in the trailing regions of the host domain. As was predicted, very few cells migrated from the transplants and in ten out of 13 embryos the cells failed to migrate at all. Our model therefore provides a possible explanation for the experimental results, which would otherwise have been regarded as failed experiments.

Lastly we considered the transplantation of cells from the trailing region of a donor embryo ahead of the migratory stream in the host. The failure of the original model to predict the experimental data led us to include conversion between the two cell subpopulations. In addition, to overcome the limitations of modelling in two spatial dimensions, we reduced the size of the transplanted domain. The extended model compared well with the experimental data and we conclude that cells are not fixed as one type but are instead able to change their molecular profile to respond to a changing microenvironment. Future modelling will aim to explore the circumstances under which cells change profiles and the mechanisms by which trailing cells receive guidance information from chemotactic cells.

To consider whether our model requires predetermined subpopulations of cells, we explored different cell conversion mechanisms with initially homogeneous populations to see if spatially segregated subpopulations could be regained. We demonstrated that a heterogeneous population of successfully migrating cells can be generated from an initially homogeneous population using time-dependent conversion. In addition, if all cells are assumed to enter the domain as the trailing subtype, time-

dependent conversion gives a small population of leading cells with trailing cells in chains that migrate almost as far into the domain as our non-converting model. In contrast, we showed that conversion based on the local chemoattractant gradient profile from a homogeneous initial population is not successful at generating long distance migration with spatially segregated subpopulations.

One advantage of using an IBM is the ability to include mechanisms explicitly at an individual level. However, stochastic models that take a significant period of time to simulate can be difficult to analyse, since any particular simulation may not be representative of the model in general. In addition, continuum-level analyses which give rise to information on the wavespeed of the migrating population and the specific dependence of the model on its parameter values are difficult to determine without running large numbers of simulations. In Chapters 4 to 7 we derive PDE approximations for increasingly complex IBMs with the final aim of being able to perform this analysis on our full model.

Chapter 4

Analysis of a simple 1D individual-based model

Although we aim ultimately to build a framework that allows us to analyse the complex hybrid IBM from Chapters 2 and 3, we will begin with much simpler models, building up systematically to more complicated mechanisms and modelling frameworks. We will apply the techniques learned at each level to extend the model further. Starting from models that we can simulate quickly also allows us to test the accuracy of our approximations in a wide region of parameter space. It is important to investigate the strengths and limitations of the approximations used so that it is clear under what circumstances the derived equations are applicable.

This chapter investigates the effects that the volume of an individual (volume exclusion) has on the evolution of a one-dimensional motile cell population. Whilst many migratory systems are highly crowded and individuals may encounter each other repeatedly, the usual continuum equations for modelling these systems represent individuals as point particles. One reason for the continued use of such equations is the lack of a consensus in the community as to how best to account for volume exclusion. In particular, most attempts to formulate a systematic derivation of a PDE description of movement with volume exclusion use on-lattice models with at most one cell permitted at each lattice point. One drawback of this formulation is that when deriving continuum equations from the model it is necessary to take the limit as the lattice spacing tends to zero, but this also results in the size of a cell (represented by the space taken by a lattice point) tending to zero. We present a framework for systematically deriving continuum equations from off-lattice IBMs in which the radius

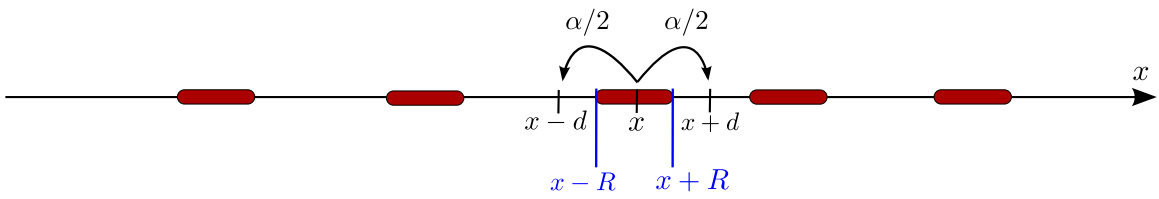


Figure 4.1: Schematic view of the one-dimensional IBM with volume exclusion. Each individual, shown in red, is represented by its centre point, x and occupies a region, $[x-R, x+R)$, around this point.

of an individual is small but finite, and thus is present explicitly in the derived equations. The work beginning this chapter has been published in Physical Review E (Dyson *et al.*, 2012).

4.1 Model description

The first model we analyse is a one-dimensional off-lattice IBM with volume exclusion. Each individual is defined by a position x in the domain $[B_L, B_R]$ and has radius, R, thus each cell occupies an interval $[x-R, x+R)$. Individuals move by hopping with rate $\alpha/2$ a distance $d > 0$ to the left or right. An attempted move is aborted if it would require moving through any point that is already occupied by another individual. For example, a move to the right from position x is aborted if there is another cell centre in the interval $[x+2R, x+d+2R)$. In addition, movement is aborted if it would result in any part of an individual being outside the domain, so that all cell centres must remain in the region $[B_L + R, B_R - R]$. The model is shown schematically in Figure 4.1.

4.2 Simulation algorithm

To enable efficient simulation of the model described in Section 4.1, we will use a modified version of the Gillespie Algorithm (Gillespie, 1977). This algorithm was originally used to simulate spatially homogeneous chemical reaction systems by using the underlying reaction rates to find the distribution of next-reaction times. To simulate the system we draw a random number from the distribution of next-reaction times and determine which reaction will occur. This can be modified to simulate our model by regarding attempted cellular movement as a 'reaction' (Baker *et al.*, 2010; Simpson *et al.*, 2011). To derive the distribution of next-movement times, τ, we consider the probability, $P_M(t)$, that there is an attempted movement in $[t, t+dt)$. For our model, $P_M(t) = N(t)\alpha dt$, where $N(t)$ is the number of cells at time t, and each cell moves left (or right) with probability $\alpha dt/2$. Let

$f(N(t), s)$ be the probability that, given the number of cells at time t, the next movement happens in the time interval $[t + s, t + s + ds)$. We also define $G(N(t), s)$, the probability that no movement is attempted in the time interval $[t, t + s)$. Then

$$\begin{aligned} f(N(t), s) &= \text{Prob}\left(\text{no move in } [t, t + s)\right) \text{Prob}\left(\text{a move in } [t + s, t + s + ds)\right), \\ &= G(N(t), s)N(s)\alpha ds, \\ &= G(N(t), s)N(t)\alpha ds, \end{aligned} \tag{4.1}$$

since there were no attempted movements during $[t, t + s)$ and hence no new cells can have been added during this time. Taking $\sigma > 0$ then we may characterise $G(N(t), \sigma)$ further by noting that

$$\begin{aligned} G(N(t), \sigma + d\sigma) &= \text{Prob}\left(\text{no move in } [t, t + \sigma)\right) \text{Prob}\left(\text{no move in } [t + \sigma, t + \sigma + d\sigma)\right), \\ &= G(N(t), \sigma)(1 - N(t)\alpha d\sigma). \end{aligned}$$

Hence, rearranging and letting $d\sigma \to 0$

$$\frac{G(N(t), \sigma + d\sigma) - G(N(t), \sigma)}{d\sigma} = -N(t)\alpha G(N(t), \sigma), \tag{4.2}$$

$$\frac{1}{G}\frac{\partial G}{\partial \sigma} = -N\alpha, \tag{4.3}$$

$$\frac{\partial}{\partial \sigma}(\ln G) = -N\alpha \tag{4.4}$$

$$G(N(t), \sigma) = Ke^{-N(t)\alpha\sigma}, \tag{4.5}$$

and since $G(N, 0) = 1$

$$G(N(t), \sigma) = e^{-N(t)\alpha\sigma}, \tag{4.6}$$

and hence

$$f(N(t), s) = N(t)\alpha e^{-N(t)\alpha s} ds. \tag{4.7}$$

To find τ so that $t + \tau$ is the time at which the next reaction occurs, we need to draw a random number from the distribution given by $f(N(t), s)$. We draw from the distribution, $f(N(t), s)$, by

noting that if τ is distributed according to equation (4.7) and

$$F(\tau) = -\int_{\tau} f(N(t), \tau')d\tau', \tag{4.8}$$

$$= e^{-N(t)\alpha\tau}, \tag{4.9}$$

then $F(\tau)$ is uniformly distributed, since if a and b are chosen arbitrarily with $0 < a < b < 1$ then the probability that $F(\tau) \in (a, b)$ is given by

$$\int_a^b F(\tau)\ \mathrm{d}F(\tau) = \int_{F^{-1}(b)}^{F^{-1}(a)} \tau\ \mathrm{d}\tau, \tag{4.10}$$

$$= \int_{F^{-1}(b)}^{F^{-1}(a)} f(N(t), \tau)\ \mathrm{d}\tau, \tag{4.11}$$

$$= N(t)\alpha \int_{F^{-1}(b)}^{F^{-1}(a)} e^{-N(t)\alpha\tau}\ \mathrm{d}\tau, \tag{4.12}$$

$$= -\int_{F^{-1}(b)}^{F^{-1}(a)} \frac{\mathrm{d}F}{\mathrm{d}\tau}\ \mathrm{d}\tau, \tag{4.13}$$

$$= F(F^{-1}(b)) - F(F^{-1}(a)), \tag{4.14}$$

$$= b - a. \tag{4.15}$$

Hence if τ is distributed according to equation (4.7), then $r = F(\tau)$ is uniformly distributed over the interval $(0, 1)$. If we take a uniformly distributed random number, r, therefore, then

$$\tau = F^{-1}(r), \tag{4.16}$$

$$= \frac{1}{N(t)\alpha} \ln \frac{1}{r}, \tag{4.17}$$

is drawn from the distribution $f(N(t), s)$ and can be used as the time to the next cell movement.

To find which cell moves and in which direction we pick a uniformly distributed integer, i, from 1 to $2N$. If $i \leq N$ then cell i attempts to move left, whereas if $i > N$ then cell $i - N$ attempts to move right. If another cell is in the way of the desired move then the attempted movement is aborted.

4.3 Derivation of continuum equations

We wish to derive the predicted distribution of cell positions at any particular time, t. The probability density function (pdf), $C_i(x, t)$, for the position, x, of cell centre i at time t can be found from

the pdfs of the position of all the cells at $t - \Delta t$ by taking Δt small enough that at most one cell movement occurs during each time step of length Δt. Then, considering the three different places that cell i could have been at the previous time step:

$$\begin{aligned}\text{Prob}(\text{cell}_i\colon \{[x, x+\Delta x), t+\Delta t\}) = \text{Prob}(\text{cell}_i\colon \{[x, x+\Delta x), t\} \to \{[x, x+\Delta x), t+\Delta t\}) \\ + \text{Prob}(\text{cell}_i\colon \{[x+d, x+d+\Delta x), t\} \to \{[x, x+\Delta x), t+\Delta t\}) \\ + \text{Prob}(\text{cell}_i\colon \{[x-d, x-d+\Delta x), t\} \to \{[x, x+\Delta x), t+\Delta t\}),\end{aligned}$$

where $\text{Prob}(\text{cell}_i\colon \{[x, x+\Delta x), t\})$ is the probability that the centre of cell i is in the region $[x, x+\Delta x)$ at time t and $\text{Prob}(\text{cell}_i\colon \{[x, x+\Delta x), t\} \to \{[x, x+\Delta x), t+\Delta t\})$ is the probability that the centre of cell i was in the region $[x, x+\Delta x)$ at both times t and $t+\Delta t$. Hence

$$\begin{aligned}C_i(x, t+\Delta t) = \underbrace{C_i(x,t)}_{\text{cell i at } (x,t)}\Big[\overbrace{1-\alpha\Delta t}^{\text{no move attempted}} + \underbrace{\frac{\alpha\Delta t}{2}\left(P_L^i(x,t) + P_R^i(x,t)\right)}_{\text{attempted move fails}}\Big] \\ + \overbrace{C_i(x+d,t)}^{\text{cell i at } x+d \text{ tries to move}}\frac{\alpha\Delta t}{2}\underbrace{\big[1 - P_L^i(x+d,t)\big]}_{\text{attempted move succeeds}} + \overbrace{C_i(x-d,t)}^{\text{cell i at } x-d \text{ tries to move}}\frac{\alpha\Delta t}{2}\underbrace{\big[1 - P_R^i(x-d,t)\big]}_{\text{attempted move succeeds}}, \quad (4.18)\end{aligned}$$

where $P_L^i(x,t)$ and $P_R^i(x,t)$ are the probabilities of a cell, other than cell i, being present in the regions $(x-d, x]$ and $[x, x+d)$, respectively, given that cell i is at x at time t. At this point we could expand $P_L^i(x,t)$ and $P_R^i(x,t)$ over the different possible positions of pairs of cells, resulting in an expression requiring knowledge of the position of triplets of cells. Triplets of cells can be expanded in terms of quartets of cells and so on. In order to make progress with the resulting infinite sequence of equations, it is necessary to make a closure assumption. We take the simplest possible approximation: that the cell positions are approximately independent when averaging over numerous simulations of the system (see Gillespie, 2009; Baker *et al.*, 2010; Erban and Othmer, 2004). Then, if $d < 4R$, so that at most one cell centre may be in the region integrated over (see

Figure 4.2):

$$P_R^i(x,t) = \sum_{j\neq i} \int_{2R}^{2R+d} C_j(x+\bar{x},t)\ \mathrm{d}\bar{x}, \tag{4.19}$$

$$P_L^i(x,t) = \sum_{j\neq i} \int_{-2R-d}^{-2R} C_j(x+\bar{x},t)\ \mathrm{d}\bar{x}, \tag{4.20}$$

$$P_R^i(x-d,t) = \sum_{j\neq i} \int_{2R-d}^{2R} C_j(x+\bar{x},t)\ \mathrm{d}\bar{x}, \tag{4.21}$$

$$P_L^i(x+d,t) = \sum_{j\neq i} \int_{-2R}^{-2R+d} C_j(x+\bar{x},t)\ \mathrm{d}\bar{x}. \tag{4.22}$$

If $2R+d$ is small compared to the length scale on which C changes then we can expand $P_L^i(x,t)$ and $P_R^i(x,t)$ in a Taylor series to obtain

$$P_R^i(x,t) = \sum_{j\neq i}\left(dC_j + \sum_{n=1}^{\infty}\sum_{k=1}^{n+1} \frac{(2R)^{n+1-k}d^k}{k!(n+1-k)!}\frac{\partial^n C}{\partial x^n}\right), \tag{4.23}$$

$$= \sum_{j\neq i}\left(dC_j + \frac{d}{2}(4R+d)\frac{\partial C_j}{\partial x} + \frac{d}{6}(12R^2+6Rd+d^3)\frac{\partial^2 C_j}{\partial x^2} + \mathcal{O}(R^n d^m)\right), \tag{4.24}$$

$$P_L^i(x,t) = \sum_{j\neq i}\left(dC_j + \sum_{n=1}^{\infty}\sum_{k=1}^{n+1} \frac{(-1)^n(2R)^{n+1-k}d^k}{k!(n+1-k)!}\frac{\partial^n C}{\partial x^n}\right), \tag{4.25}$$

$$= \sum_{j\neq i}\left(dC_j - \frac{d}{2}(4R+d)\frac{\partial C_j}{\partial x} + \frac{d}{6}(12R^2+6Rd+d^3)\frac{\partial^2 C_j}{\partial x^2} + \mathcal{O}(R^n d^m)\right), \tag{4.26}$$

where C_j is evaluated at x and $n+m \geq 4$. Substituting equations (4.24) and (4.26) into equation (4.18), and taking the limit as $\Delta t \to 0$, we find that

$$\frac{\partial C_i}{\partial t} = \frac{\alpha d^2}{2}\frac{\partial^2 C_i}{\partial x^2} + \frac{\alpha d^2}{2}\sum_{m=0}^{\infty}\frac{2^{2m+2}R^{2m+1}}{(2m+1)!}\frac{\partial}{\partial x}\left(C_i\sum_{j\neq i}\frac{\partial^{2m+1}C_j}{\partial x^{2m+1}}\right) + \mathcal{O}(d^3), \tag{4.27}$$

$$= \frac{\alpha d^2}{2}\frac{\partial^2 C_i}{\partial x^2} + \frac{\alpha d^2}{2}(4R-d)\frac{\partial}{\partial x}\left(C_i\sum_{j\neq i}\frac{\partial C_j}{\partial x}\right) + \mathcal{O}(d^{2+m}R^n), \tag{4.28}$$

for $n+m \geq 2$. We wish to now let the jump length, d, tend to zero. However, as the distance moved in a jump changes, the rate of jumps occurring, α, must change correspondingly so that movement still occurs. Thus as individuals move less far, they must move more often to maintain migration. This is widely used in the literature, for example in Codling *et al.* (2008); Painter and Sherratt

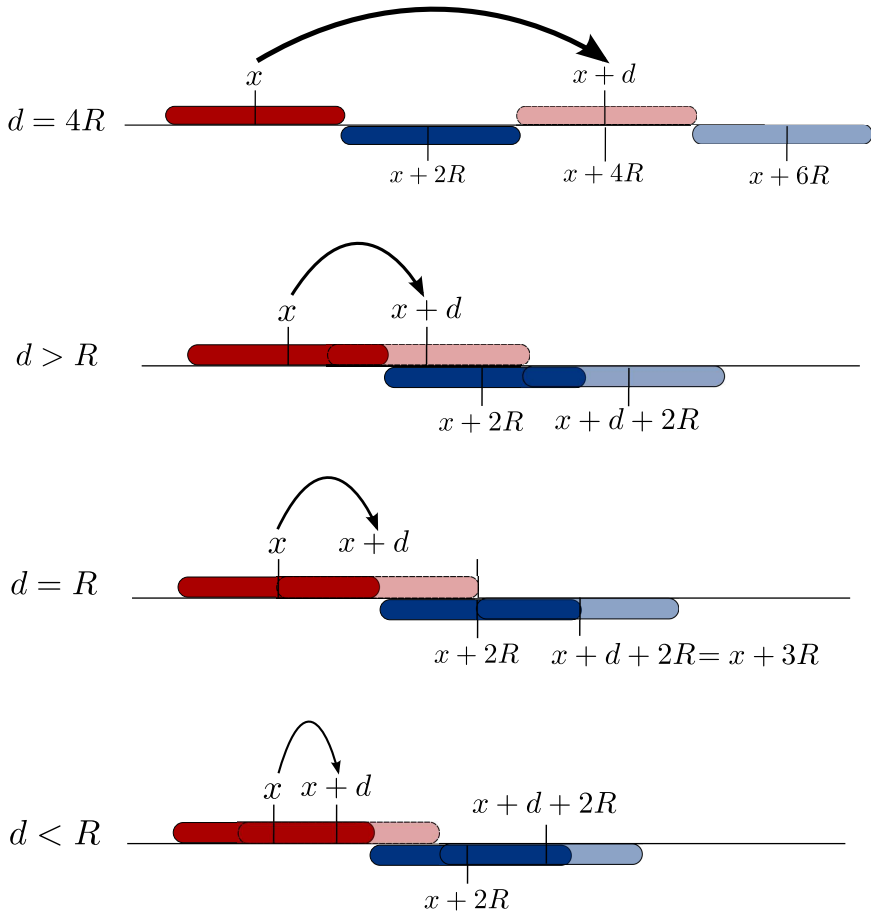


Figure 4.2: Different relative sizes of R and d result in different configurations when a cell attempts to move. Here the solid red cell with centre at x is attempting to move to the right to the position shown by the light red cell with centre at $x + d$. The solid blue cell with centre at $x + 2R$ shows the left-most position another cell could occupy to prevent movement, whilst the light blue cell at $x + d + 2R$ is the rightmost position another cell could occupy to prevent movement. Hence for the red cell to move, we require the interval $[x + R, x + R + d)$ to be clear of any part of a cell, thus requiring no cell centre to be in $[x + 2R, x + 2R + d)$.

(2003). As $d \to 0$

$$\frac{\partial C_i}{\partial t} = \hat{\alpha}\frac{\partial^2 C_i}{\partial x^2} + 4R\hat{\alpha}\frac{\partial}{\partial x}\left(C_i \sum_{j \neq i} \frac{\partial C_j}{\partial x}\right) + \mathcal{O}(R^2), \tag{4.29}$$

with $\hat{\alpha} = \lim_{d \to 0, \alpha \to \infty} \alpha d^2/2$ held constant.

4.3.1 Initial conditions

Initial conditions for C_i are given directly by the distribution of initial conditions in the IBM. If all the cells have initial positions drawn from the same distribution then $C_i(x,t) = C_j(x,t)\ \forall i, j$ and

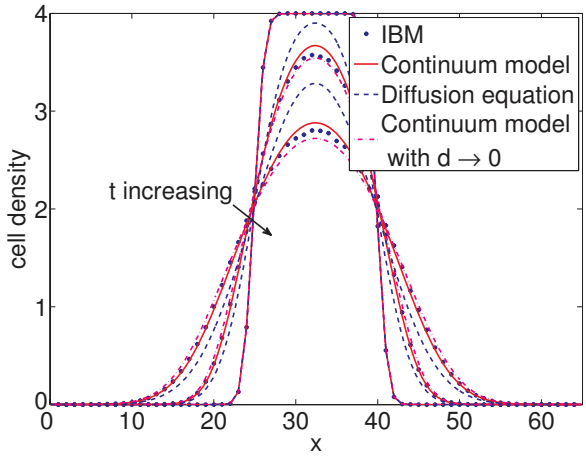


Figure 4.3: Comparison between the IBM and the continuum models with $d > 0$ (equation (4.30)) and as $d \to 0$ (equation (4.31)), plotted at $t = 0$, 200 and 600 for $x \in [0, 65]$. The solution to the diffusion equation is also shown for comparison. Parameters are $N = 60$ and $\mathrm{d}x = 0.05$, with other parameter values given in Table 4.1. For more details about the simulations see Section 4.4.

$\sum_{j\neq i} \partial C_j/\partial x = (N-1)\partial C_i/\partial x$, where N is the number of cells in the domain. Hence, summing equation (4.28) over all i gives

$$\frac{\partial C}{\partial t} = \frac{\alpha d^2}{2}\frac{\partial}{\partial x}\left(\left(1 + (4R - d)\frac{(N-1)}{N}C\right)\frac{\partial C}{\partial x}\right) + \mathcal{O}\left(d^{2+m}R^n\right), \tag{4.30}$$

for $n + m \geq 2$, where $C = \sum_{i=1}^{N} C_i$ is the total cell density, giving an indication of the distribution of all cells over the domain. Taking the limit as $d \to 0$,

$$\frac{\partial C}{\partial t} = \hat{\alpha}\frac{\partial}{\partial x}\left(\left(1 + 4R\frac{(N-1)}{N}C\right)\frac{\partial C}{\partial x}\right) + \mathcal{O}(R^2), \tag{4.31}$$

with $\hat{\alpha} = \lim_{d\to 0} \alpha d^2/2$ held constant as before.

In fact, even if initial cell positions are drawn from a variety of distributions we may still average these to use as the initial condition for $C = \sum_{i=1}^{N} C_i$ so that $C(0) = \sum_{i=1}^{N} C_i(0)$. C will then satisfy equation (4.31), as before, since we only ever consider the sum of the individual cell positions. The label of each cell in the summation is irrelevant, therefore, and so we may take the cells in different orders for each run of the numerical simulation. The initial condition of cell i is then the same for all i, so that we reduce to the case with identical initial conditions. This is analogous to Simpson *et al.* (2010b, 2011), where a one-dimensional system is derived by averaging over vertical columns of cells. The initial condition is thus averaged over all the cells without regard for whether particular cells may draw their initial position from different distributions.

4.3.2 Limiting equations

Equation (4.31) gives the limiting continuum model for finite-sized individuals moving with diffusivity $\hat{\alpha}$. To examine the accuracy of our derived equations under different parameter regimes, however, simulations of the IBM must be carried out with $d > 0$, and if d and R are of the same order of magnitude, it may be more appropriate to use equation (4.30) (whilst neglecting the higher order terms in the expansion). In this case, if $[1 + (4R - d)(N - 1)/N\ C(x,t)] < 0$ for any x, then equation (4.30) is a diffusion equation with negative diffusivity and is thus ill-posed. Indeed if $[1 + (4R - d)(N - 1)/N\ C(x,t)] < 0$ initially for any x then the negative diffusivity will lead to aggregation, and thus $[1+(4R-d)(N-1)/N\ C(x,t)] < 0$ for later time. In contrast, equation (4.31) is well-posed for all possible initial conditions since $1 + 4R(N - 1)/N\ C(x,t)$ is always positive for positive cell densities. The potential ill-posedness of equation (4.30) reflects the assumption in equations (4.19) to (4.22) that there can be at most one cell in an interval of length d, which is not satisfied if $d \geq 4R$. However, when $d \geq 4R$, equations (4.19) to (4.22) are still valid first order approximations for $C_i \ll 1$ since, using moment-closure assumptions:

$$P_R^i(x,t) = \int_{2R}^{2R+d} \left(\overbrace{1 - \overbrace{\prod_{j \neq i} \left(1 - C_j(x + \bar{x}, t)\right)}^{\text{prob(no cells present)}}}^{\text{prob(at least one cell present)}} \right) \mathrm{d}\bar{x}, \tag{4.32}$$

$$\approx \int_{2R}^{2R+d} \left(\sum_{j \neq i} C_j(x + \bar{x}, t) - \sum_{k \neq j \neq i} C_j(x + \bar{x}, t) C_k(x + \bar{x}, t) + \mathcal{O}\left(C_j^3\right) \right) \mathrm{d}\bar{x}. \tag{4.33}$$

Note that taking $R = 0$ in equation (4.30) does not reduce to simple diffusion, unlike in equation (4.31), where d also tends to zero. Volume exclusion still has an effect when $R = 0$ since, whilst cells 'occupy' no space, they still prevent other individuals from moving past for $d > 0$ since cells still sense the presence of others before moving.

4.3.3 Boundary and initial conditions

Appropriate boundary conditions must be considered carefully, since the centre of a cell cannot be within a distance R of the boundary. In essence we have no-flux conditions at $B_L + R$ and $B_R - R$, where B_L and B_R are the left and right boundaries, respectively. No-flux conditions neglect effects close to the boundary, however, where P_L^i and P_R^i may be different compared to the interior of the domain, and for now we simply choose a domain large enough to ensure cells do not reach the

Parameter	Typical value
d	0.15
R	0.15
B_L	0
B_R	80
total time	750
total cell number	30
dx	0.1
number of simulations	1000
P_c	0.05
α	2.22
σ	0.15
L_a	0.1
ϵ	0.1

Table 4.1: Typical simulation parameters.

boundaries.

4.4 Simulations

To compare the total cell density profile of C with model simulations, we average cell positions over many repetitions of the simulation to obtain the mean cell profile. Realisations of the IBM were performed using the Gillespie algorithm (see Section 4.2 and Gillespie, 1977) with initial conditions as follows: the centre of the first individual is drawn from a normal distribution, $\mathcal{N}((B_R - B_L)/2 - NR/0.8, 10)$, and subsequent cells are placed to the right with centres at intervals of $2R/0.8$ apart so that the cells are initially in the middle of the domain at a density of 0.8.

Note that, whilst all simulations shown here begin with evenly spaced cells, using an uneven distribution of cells does not change our averaged simulation (result not shown). Taking instead a random distribution of initial positions to a high density results in failed initialisations of individuals due to volume exclusion restricting overlapping cells. However, for each failed initialisation another random number must be drawn until the initial position is successful and this increases simulation times without giving additional useful data. We will therefore take evenly spaced cells.

We solve the PDEs using the Numerical Algorithms Group (NAG) routine d03pc (see Appendix A.1) with initial conditions determined by the average initial distribution from simulations, linearly interpolated onto a mesh with spacing dx.

Simulations with $R = 0.1$ and $d = 0.15$ show better agreement with either of equation (4.30) or equation (4.31) than with the diffusion equation (Figure 4.3). In this case the averaged simulations

lie between the predictions with $d \to 0$ and with $d = 0.15$, reflecting the complex nature of the error in our approximations (see Section 4.6).

4.5 Flux boundary conditions

Since our motivating biological model requires cells to enter the domain from the neural tube (see Chapters 2 and 3), we wish to be able to include an influx of cells into one end of the domain. There are two potential methods for incorporating a flux boundary condition into our derivation: either by finding the total number of cells predicted by our equation at any time t and ensuring this increases at a set rate; or by using the same method that was used within the domain whilst taking the differences due to the boundary into account.

4.5.1 Prescribing the total cell population over time

The change in the total number of cells is given by

$$\frac{\partial}{\partial t}\int_{B_L+R}^{B_R-R} C(x,t)\ \mathrm{d}x = \int_{B_L+R}^{B_R-R} \frac{\partial C}{\partial t}\ \mathrm{d}x, \tag{4.34}$$

$$= \int_{B_L+R}^{B_R-R} \left[\hat{\alpha}\frac{\partial}{\partial x}\left(\left(1+4R\frac{(N-1)}{N}C\right)\frac{\partial C}{\partial x}\right)+\mathcal{O}(R^2)\right]\ \mathrm{d}x, \tag{4.35}$$

$$= \left[\hat{\alpha}\left(1+4R\frac{(N-1)}{N}C\right)\frac{\partial C}{\partial x}\right]_{B_L+R}^{B_R-R} + (B_R-B_L-2R)\mathcal{O}(R^2). \tag{4.36}$$

Hence for an influx of cells at a rate P_c at $x = B_L + R$, whilst keeping a no-flux boundary condition at $x = B_R - R$, we require

$$-\hat{\alpha}\left(1+4R\frac{(N-1)}{N}C\right)\frac{\partial C}{\partial x}\bigg|_{x=B_L+R} + (B_R-B_L-2R)\mathcal{O}(R^2) = P_c\left(1-\int_{B_L+R}^{B_L+3R} C(x,t)\ \mathrm{d}x\right), \tag{4.37}$$

where the rate P_c is modified by volume exclusion, since if there is another individual occupying the section next to $x = B_L$ then a new cell cannot enter the domain. Using a Taylor expansion to

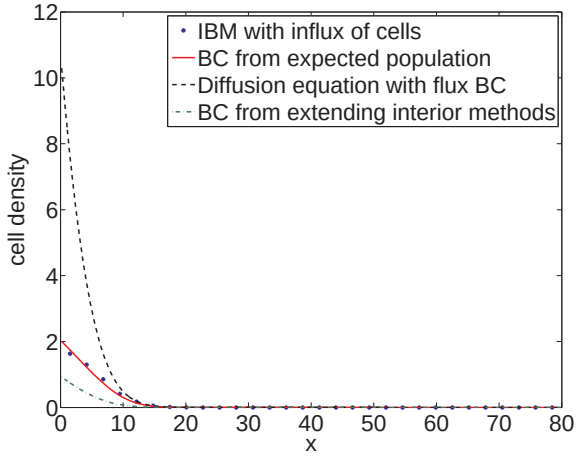


Figure 4.4: The evolution of the expected population size can be used to derive a boundary condition (BC) with an influx of cells (equation (4.41)) that compares well with averaged simulations of the IBM. Using the same methods as the interior of the domain does not result in the same boundary condition (equation (4.45)), and does not compare well to averaged simulations of the IBM. Parameter values are given in Table 4.1. For more details about the simulations see Section 4.4.

simplify the right-hand side, we obtain

$$-\hat{\alpha}\left(1+4R\frac{(N-1)}{N}C\right)\frac{\partial C}{\partial x}\bigg|_{x=B_L+R} = P_c\left(1-\int_0^{2R}\left(C(B_L+R,t)+x\frac{\partial C}{\partial x}(B_L+R,t)+\mathcal{O}(x^2)\right)\,\mathrm{d}x\right), \tag{4.38}$$

$$= P_c\left(1-2RC(B_L+R,t)+\mathcal{O}(R^2)\right). \tag{4.39}$$

Note that for large population numbers (N), which result after long simulation time, $(N-1)/N \approx 1$. Since numerical solutions taking $(N-1)/N \approx 1$ are much faster when the number of cells changes over time and are very similar to those which use a coupled ODE to approximate N (result not shown), we will use

$$-\hat{\alpha}\left(1+4RC\right)\frac{\partial C}{\partial x}\bigg|_{x=B_L+R} = P_c\left(1-\int_0^{2R} C(B_L+R,t)+x\frac{\partial C}{\partial x}(B_L+R,t)+\mathcal{O}(x^2)\,\mathrm{d}x\right), \tag{4.40}$$

$$= P_c\left(1-2RC(B_L+R,t)+\mathcal{O}(R^2)\right), \tag{4.41}$$

in our simulations, and we will neglect terms of order $\mathcal{O}(R^2)$ in our simulations.

4.5.2 Extending our original method

The methods used for deriving PDEs within the domain (Section 4.3) are difficult to extend to give a simple boundary condition, since the boundary influences the movement of cells that are within three radii of the boundary. If the centre of cell i is to the left of $B_L + 3R$, for example, there is not enough space for another individual to the left and hence all attempted movements of cell i to the left whilst remaining within the domain must succeed. These boundary layer effects have also been observed in systems without volume exclusion (Singer and Schuss, 2008). If we wish to derive a condition that is only taken at the boundary and assume that these effects close to the boundary are negligible, then taking $P_L^i(B_L + R) = 0$ in equation (4.18) and noting there cannot be a cell at $(B_L + R - d)$:

$$\begin{aligned} C_i(B_L + R, t + \Delta t) = C_i(B_L + R, t)\Big[1 - \alpha\Delta t/2 + \frac{\alpha\Delta t}{2} P_R^i(B_L + R, t)\Big] \\ + C_i(B_L + R + d, t)\frac{\alpha\Delta t}{2}\left[1 - P_L^i(B_L + R + d)\right] \\ + P_c\left(1 - \int_{B_L+R}^{B_L+3R} C(x,t)\ \mathrm{d}x\right). \end{aligned} \tag{4.42}$$

Taylor expanding in d and R, and using equations (4.24) and (4.26) as before yields

$$d\frac{\partial C_i}{\partial t}\bigg|_{x=B_L+R} = \left\{\frac{\alpha d^2}{2}\left(1 + 4R\frac{N-1}{N}C\right)\frac{\partial C}{\partial x} + dP_c\left(1 - 2RC\right) + \mathcal{O}(d^{2+m}R^n)\right\}\bigg|_{x=B_L+R}, \tag{4.43}$$

as $\Delta t \to 0$, where $n + m \geq 2$. To maintain a flux boundary condition as $d \to 0$, we require $P_c \to \infty$ so that $\hat{P}_c = dP_c$ is constant. In addition, we keep $\hat{\alpha} = \alpha d^2/2$ constant, and assuming that $\partial C_i/\partial t = \mathcal{O}(1)$:

$$-\hat{\alpha}\left(1 + 4R\frac{(N-1)}{N}C\right)\frac{\partial C}{\partial x}\bigg|_{x=B_L+R} = \hat{P}_c\left(1 - 2RC + \mathcal{O}(R^2)\right)\bigg|_{x=B_L+R}. \tag{4.44}$$

As before we may now take $(N-1)/N = 1$ so that

$$-\hat{\alpha}\left(1 + 4RC\right)\frac{\partial C}{\partial x}\bigg|_{x=B_L+R} = \hat{P}_c\left(1 - 2RC + \mathcal{O}(R^2)\right)\bigg|_{x=B_L+R}. \tag{4.45}$$

We have shown, therefore, that the two methods of derivation do not give the same boundary condition, since in equation (4.41), P_c does not depend on the jump distance, d. We expect that the first method (Section 4.5.1) will be more accurate, since it attempts to maintain the correct total

number of cells in the domain, and does not explicitly rely on solutions close to the boundary where we know that our approximations are less accurate. In numerical solutions to the two equations we take $d > 0$ and compare averaged simulations to show that equation (4.41) is indeed much more accurate than using equation (4.45) (Figure 4.4). When using a flux boundary condition, the presence of volume exclusion makes a huge difference to the concentration profile and total population size. It is therefore highly important to consider the reduction in influx into the domain caused by crowding.

4.6 Exploring parameter space

In this section, we investigate the accuracy of our continuum predictions (equations (4.30) and (4.31)) by comparing them with data averaged over multiple simulations of the model. To determine whether our derived equations are better predictors than naïvely using the diffusion equation, we will also compare the diffusion equation with averaged simulations of the model. Finally we also examine the varying importance of excluded volume in equation (4.31) by plotting the difference between our equation and the diffusion equation.

The differences between our equation, the diffusion equation and averaged simulations of the model will vary as the model parameters change. We therefore explore parameter space as the number of cells, N, the radius of a cell, R, and the distance moved during a jump, d, change. We take these parameters in pairs so that the result may be visualised as a three-dimensional surface plot.

To determine the aforementioned differences, we use the relative difference which, for two normalised density functions, f_1 and f_2, over a region $[B_L, B_R]$, is

$$D(f_1, f_2) = \frac{1}{B_R - B_L} \int_{B_L}^{B_R} \frac{(f_1(x) - f_2(x))^2}{f_1(x) + f_2(x)} \, \mathrm{d}x. \tag{4.46}$$

It is important to note that our numerical solutions to PDEs will only be correct to some given accuracy. Our numerical technique has an error given by $|E_i| = acc(1 + C(x_i))$ at each lattice point x_i, and we use $acc = 10^{-5}$. Hence if we find the difference between our prediction, C, and data S

then the error, $E(S,C)$ in $D(S,C)$ can be estimated by Taylor expanding in acc

$$E = \frac{1}{B_R - B_L}\int_{B_L}^{B_R} \frac{(S(x) - C(x) - acc(1 + C(x)))^2}{S(x) + C(x) + acc(1 + C(x))}\, \mathrm{d}x - D(S,C), \tag{4.47}$$

$$= -\frac{acc}{B_R - B_L}\int_{B_L}^{B_R} \frac{(S(x) - C(x))(3S(x) + C(x))(1 + C(x))}{(S(x) + C(x))^2}\, \mathrm{d}x + \mathcal{O}(acc^2), \tag{4.48}$$

so that if $(S(x)-C(x))(3S(x)+C(x))$ is of the same order of magnitude as $(S(x)+C(x))^2$ then the error in $D(S,C)$ is of order $acc = 10^{-5}$. Hence values of D that are orders of magnitude greater than acc may be considered to be due to approximations in our equations or fluctuations in the averaged simulations. The fluctuations in averaged simulations may be considered using the Central Limit Theorem, so that the difference between the true mean value and that found by n samples converges to a normal distribution with variance $1/n\sigma^2$. For our simulations, n may be considered to be the total number of cell paths averaged over, so that for 30 cells and 1000 repetitions, $1/n = \mathcal{O}(10^{-5})$. In addition, since all compared values of D are averaged over the same number of repetitions, the error in each should be of the same order of magnitude and thus be comparable.

4.6.1 Likelihood of successful movement

The first pair of parameters we consider is the radius of a cell, R, and the distance moved during each cell jump, d. If we hold the number of cells, N, constant then the occupied space ($2RN$) and the distance sensed in each attempted move (d) determine the likelihood of an attempted move being successful. In turn, if attempted moves are less likely to succeed then we would expect the total excluded volume effect over the domain to be larger. However, when $4R < d$ we find that the excluded volume term in equation (4.30) becomes negative, reducing the diffusivity of the population. The reduction in diffusion results from the playoff between reduced rates of successful movement due to volume exclusion and the resultant increase in the proportion of successful cell jumps towards unoccupied areas of the domain. The system is best approximated by the diffusion equation when $4R = d$ (Figure 4.5(c)).

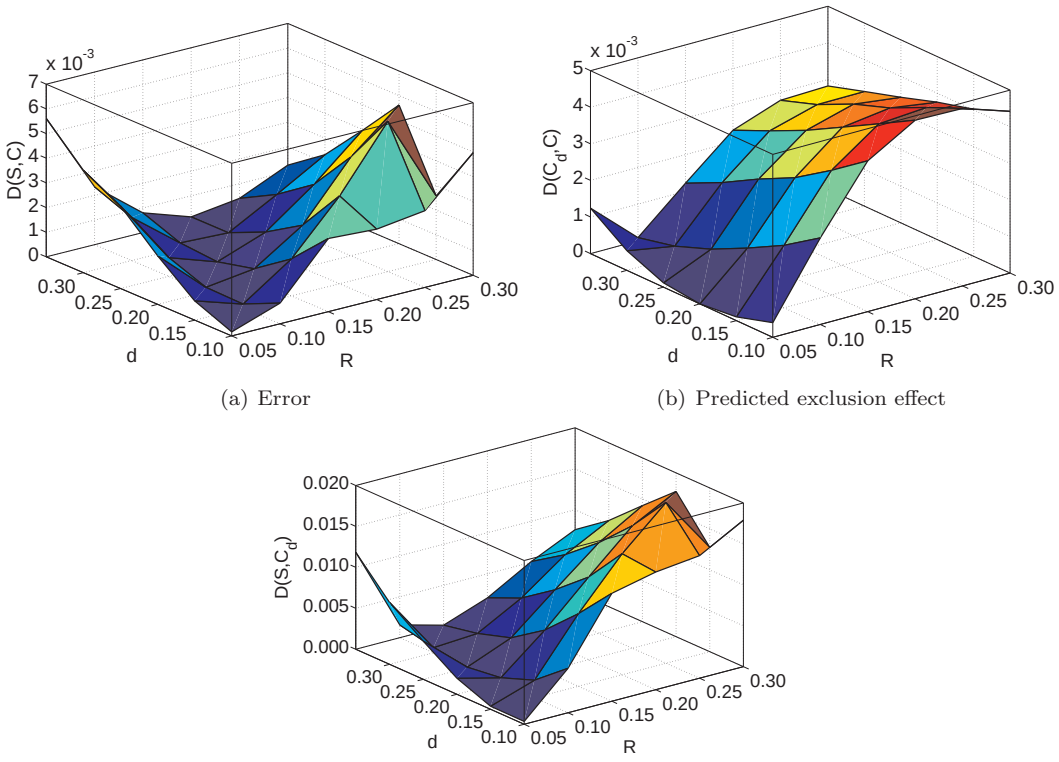
Varying parameter values changes the volume exclusion effects, but the error due to our approximations is also affected. Two different types of approximations are taken during our derivation: firstly the moment closure approximation that the positions of pairs of cells are approximately independent; and secondly the truncation of Taylor expansions in powers of $(2R + d)$. The error in our predictive equation is a combination of higher order terms from each of these approximations. Whilst we may

compute the next order term in our Taylor series expansions, this is highly complex and depends non-linearly on both R and d. In addition it is unclear how to make a higher order moment closure approximation (although these have been attempted, see Baker and Simpson, 2010; Singer, 2004) and this would be necessary to reliably analyse higher order terms. It is difficult to predict, therefore, the error introduced by our approximations and how this will vary as our parameters change.

However, we may find the error in our approximations as R and d change by determining the difference between a numerical solution to equation (4.30) and averaged simulations (Figure 4.5(a)). For small values of d and R, the errors in our predictions are very small (Figure 4.5(a)), however the effect of excluded volume is also reduced (Figure 4.5(b)). It is clear from Figure 4.5 that the greatest excluded volume effect with the minimum error can be found by increasing d whilst keeping R small. For 30 of the 36 parameter combinations tried, equation (4.30) (C) was more accurate than the diffusion equation (C_d), when compared to data (S) averaged over many simulations. In addition, the diffusion equation was only very slightly more accurate than equation (4.30), with $D(S, C_d) - D(S, C) > -10^{-4}$ for all parameter combinations that were tried.

4.6.2 Population crowding

The second pair of parameters we vary is the number of cells, N, and the radius of a cell, R. Taken in combination, these parameters give an indication of the total space occupied by cells ($2RN$) and hence the level of crowding due to the cell population. As R and N increase we would expect the effect of volume exclusion to become more pronounced. This expectation is supported by equations (4.30) and (4.31) since increasing R and N increases the exclusion term $4R(N-1)/N$. We note that as $N \to \infty$, $(N-1)/N \to 1$, reflecting the decreasing relative impact of each additional cell as the existing numbers increase. Unfortunately, whilst increasing R and N gives a greater excluded volume effect, the inaccuracies introduced by our approximations also become more significant. As with $R-d$ space, the error surface for our predictions is complex and difficult to find analytically. The accuracy of both the moment closure approximation and the truncated Taylor expansions depend on the size of R and N, since higher order terms in the Taylor expansions are also multiplied by $(N-1)/N$. Since $(N-1)/N$ is bounded above by 1 as N increases, however, we expect the major change in errors when varying the number of individuals to be due to higher order correlations between cells that are not accounted for in our moment closure approximation. In $R-N$ space the exclusion effects and the approximation errors seem to be highly correlated (Figure 4.6). However, for small

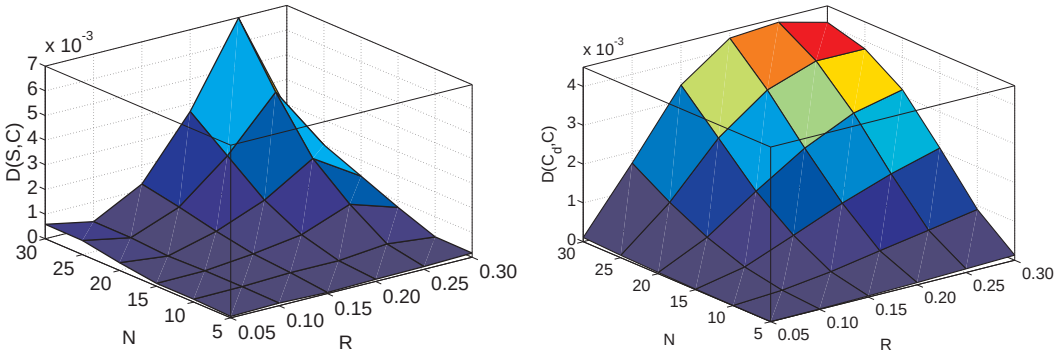


(a) Error

(b) Predicted exclusion effect

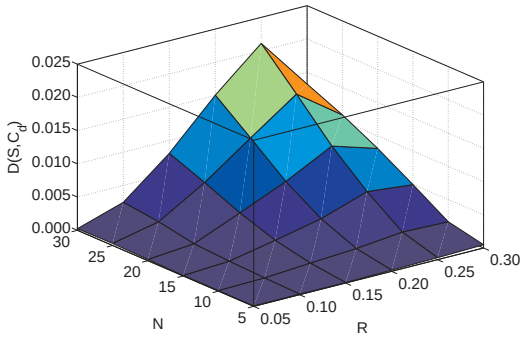
(c) Exclusion effect

Figure 4.5: Exploring the effect of changing the likelihood of successful movement (by changing d and R whilst keeping $\hat{\alpha} = \alpha d^2/2 = 0.025$ constant) on the error introduced by our approximations ($D(C, S)$), our prediction of the effect of volume exclusion ($D(C, C_d)$) and the actual effect of volume exclusion ($D(S, C_d)$). Here S is the average distribution of simulations of the IBM, C is the solution to equation (4.28), C_d is the solution to the diffusion equation and the relative difference, D, is given in equation (4.46). Parameter values are given in Table 4.1. For more details about the simulations see Section 4.4.



(a) Error

(b) Predicted exclusion effect



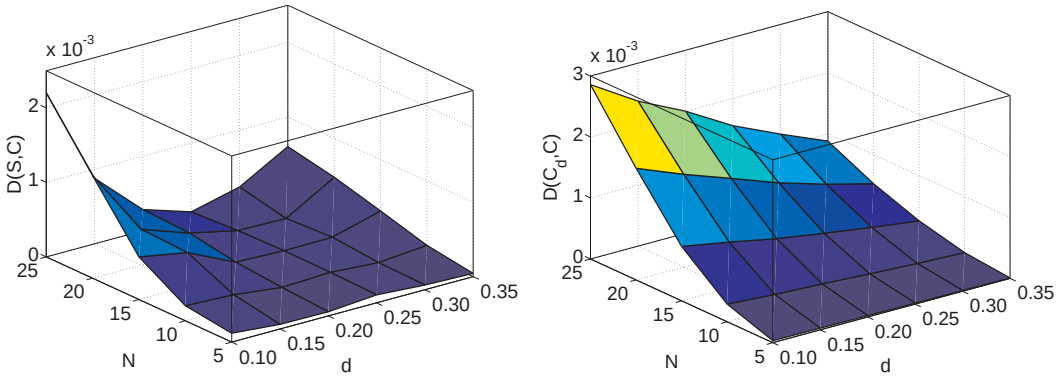
(c) Exclusion effect

Figure 4.6: Exploring the effect of changing the occupied space (by changing N and R) on the error introduced by our approximations ($D(C, S)$), our prediction of the effect of volume exclusion ($D(C, C_d)$) and the actual effect of volume exclusion ($D(S, C_d)$). Here S is the average distribution of simulations of the IBM, C is the solution to equation (4.28), C_d is the solution to the diffusion equation and the relative difference, D, is given in equation (4.46). Parameter values are given in Table 4.1. For more details about the simulations see Section 4.4.

R the error is relatively small, even as N increases and we note that for $R > 0.05$ equation (4.28) has a smaller error than the diffusion equation.

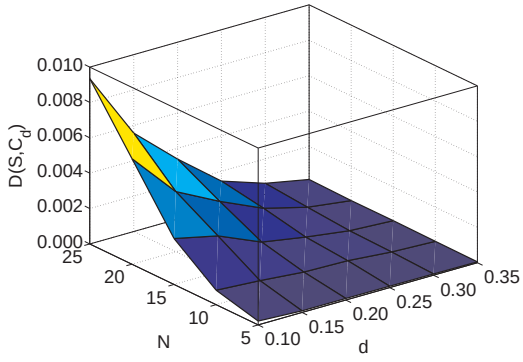
4.6.3 The final pair of parameters

Finally we vary the number of cells, N, and the distance moved in a jump, d, in combination (Figure 4.7). Here again, increasing N leads to an increase in both the total effect of volume exclusion and the error in our approximations, as we would expect from the greater chance of cell-cell encounters and the greater correlations that we would expect between cell positions. For the value of R taken, decreasing d gives a higher exclusion effect and greater error term and this is true for all considered values of N.



(a) Error

(b) Predicted exclusion effect



(c) Exclusion effect

Figure 4.7: Exploring the effect of changing the number of cells, N, and the distance moved in a jump, d, on the error introduced by our approximations ($D(C, S)$), our prediction of the effect of volume exclusion ($D(C, C_d)$) and the actual effect of volume exclusion ($D(S, C_d)$). Here S is the average distribution of simulations of the IBM, C is the solution to equation (4.28), C_d is the solution to the diffusion equation and the relative difference, D, is given in equation (4.46). Parameter values are given in Table 4.1. For more details about the simulations see Section 4.4.

4.7 Distributed distance moved

One advantage of using an off-lattice IBM is that we may consider models where the distance moved in a jump may vary over time and between individuals. For instance, we would expect that there would be a distribution of time-varying speeds within any population of cells. We could represent this in our model either by varying the rate of movement for different individuals, or by using a distribution of cell jump distances. Using a variable jump distance will affect the overall behaviour of the population, since it enables individuals to move arbitrarily close together and, in one dimension, prevents each cell from essentially moving on a lattice. In particular, using a normally distributed jump distance makes the model more similar to the classic depiction of Brownian dynamics, with (Barkai and Silbey, 2010; Bruna and Chapman, 2012) or without (Codling *et al.*, 2008) hard core exclusion effects.

4.7.1 Derivation of continuum equations

It is straightforward to extend our method to consider variable jump distances. With the pdf of d given by $f(u_d)$, for $u_d \in \mathbb{R}$, so that $\text{Prob}(d \in [u_d, u_d + \Delta u_d)) = f(u_d)$, we may use the probabilistic expansion:

$$\text{Prob}(A = a) = \int_{b \in \text{B}} \text{Prob}(A = a | B = b)\ \text{Prob}(B = b)\ \text{d}b, \tag{4.49}$$

to expand the probability

$$\text{Prob}(\text{cell}_i\text{: } \{[x, x + \Delta x), t + \Delta t\}) = \int_{-\infty}^{\infty} f(u_d) \text{Prob}(\text{cell}_i\text{: } \{[x, x + \Delta x), t + \Delta t\} \,|d = u_d)\ \text{d}u_d. \tag{4.50}$$

Then, similarly to equation (4.18), if Δt is small enough for only one cell movement to happen during $(t, t + \Delta t)$,

$$\begin{aligned} C_i(x, t + \Delta t) = \int_{-\infty}^{\infty} f(u_d) \{ & \underbrace{C_i(x,t)}_{\text{cell i at } (x,t)} \Big[\overbrace{1 - \alpha \Delta t}^{\text{no move attempted}} + \underbrace{\alpha \Delta t P^i(x, u_d, t)}_{\text{attempted move fails}} \Big] \\ & + \overbrace{C_i(x - u_d, t) \alpha \Delta t}^{\text{cell i at } x - u_d \text{ tries to move}} \underbrace{\big[1 - P^i(x - u_d, u_d, t)\big]}_{\text{attempted move succeeds}} \Big]\ \text{d}u_d, \end{aligned} \tag{4.51}$$

where the probability, P^i, that a cell other than cell i prevents movement, now depends on the jump distance u_d:

$$P^i(x, u_d, t) = \sum_{j \neq i} \int_{2R}^{2R+u_d} C_j(x + \bar{x}, t) \, \mathrm{d}\bar{x}. \tag{4.52}$$

Hence

$$\begin{aligned}\frac{\partial C_i}{\partial t} = \alpha \int_0^\infty (f(u_d) + f(-u_d))(C_i(x - u_d, t) - C_i(x, t)) \, \mathrm{d}u_d \\ + \alpha \sum_{j \neq i} \int_0^\infty \Bigg[f(u_d) \left(C_i(x, t) \int_{2R}^{2R+u_d} C_j(x + \bar{x}, t) \, \mathrm{d}\bar{x} - C_i(x - u_d, t) \int_{2R-u_d}^{2R} C_j(x + \bar{x}, t) \, \mathrm{d}\bar{x} \right) \\ + f(-u_d) \left(C_i(x, t) \int_{-2R-u_d}^{-2R} C_j(x + \bar{x}, t) \, \mathrm{d}\bar{x} - C_i(x + u_d, t) \int_{-2R}^{-2R+u_d} C_j(x + \bar{x}, t) \, \mathrm{d}\bar{x} \right) \Bigg] \, \mathrm{d}u_d,\end{aligned} \tag{4.53}$$

separating the integral in equation (4.51) into $u_d \geq 0$ and $u_d < 0$.

Now if the jump direction is unbiased, so that $f(u_d) = f(-u_d)$, and if $f(u_d)$ is exponentially small for $u_d \notin [-\epsilon, \epsilon]$, where $\epsilon \ll 1$ then, since $|u_d| < \epsilon \ll 1$ within the integral, then we may expand in terms of u_d as in Section 4.3:

$$\begin{aligned}\frac{\partial C_i}{\partial t} = \alpha \int_0^\epsilon u_d^2 f(u_d) \, \mathrm{d}u_d \frac{\partial^2 C_i}{\partial x^2} + \alpha \sum_{m=0}^\infty \frac{2^{2m+2} R^{2m+1}}{(2m+1)!} \frac{\partial}{\partial x} \left(C_i \sum_{j \neq i} \frac{\partial^{2m+1} C_j}{\partial x^{2m+1}} \right) \int_0^\epsilon u_d^2 f(u_d) \, \mathrm{d}u_d \\ + \mathcal{O}\left(\int_0^\epsilon f(u_d)(u_d^3) \, \mathrm{d}u_d \right).\end{aligned} \tag{4.54}$$

Extending the integrals to infinity, whilst introducing only exponentially small error terms,

$$\frac{\partial C_i}{\partial t} = \frac{\alpha \langle d^2 \rangle}{2} \frac{\partial^2 C_i}{\partial x^2} + \frac{\alpha \langle d^2 \rangle}{2} \sum_{m=0}^\infty \frac{2^{2m+2} R^{2m+1}}{(2m+1)!} \frac{\partial}{\partial x} \left(C_i \sum_{j \neq i} \frac{\partial^{2m+1} C_j}{\partial x^{2m+1}} \right) + \mathcal{O}(\langle d^3 \rangle). \tag{4.55}$$

For example, if the jump distances are normally distributed with variance σ^2 ($d \sim \mathcal{N}(0, \sigma^2)$), then $\int_0^\infty f(u_d) u_d^3 \, \mathrm{d}u_d = \sigma^3 \sqrt{2/\pi}$ and $\langle d^2 \rangle = \sigma^2$, so that

$$\frac{\partial C_i}{\partial t} = \frac{\alpha \sigma^2}{2} \frac{\partial^2 C_i}{\partial x^2} + \frac{\alpha \sigma^2}{2} \sum_{m=0}^\infty \frac{2^{2m+2} R^{2m+1}}{(2m+1)!} \frac{\partial}{\partial x} \left(C_i \sum_{j \neq i} \frac{\partial^{2m+1} C_j}{\partial x^{2m+1}} \right) + \mathcal{O}(\sigma^3). \tag{4.56}$$

As before, if the initial distribution is the same for all C_i, then summing equation (4.56) over all i

$$\frac{\partial C}{\partial t} = \frac{\alpha\sigma^2}{2}\frac{\partial^2 C}{\partial x^2} + \frac{\alpha\sigma^2}{2}\frac{N-1}{N}\sum_{m=0}^{\infty}\frac{2^{2m+2}R^{2m+1}}{(2m+1)!}\frac{\partial}{\partial x}\left(C_i\sum_{j\neq i}\frac{\partial^{2m+1}C_j}{\partial x^{2m+1}}\right) + \mathcal{O}(\sigma^3), \tag{4.57}$$

$$= \frac{\alpha\sigma^2}{2}\frac{\partial^2 C}{\partial x^2} + \frac{\alpha\sigma^2}{2}\frac{N-1}{N}\left(4R - 2\sigma\sqrt{\frac{2}{\pi}}\right)\frac{\partial}{\partial x}\left(C\frac{\partial C}{\partial x}\right) + \mathcal{O}(R^n\sigma^{2+m}), \tag{4.58}$$

where $C = \sum_{i=1}^{N} C_i = NC_1$ and $n + m \geq 2$. Equation (4.58) is of the same form as for the fixed distance model (equation (4.28)), with d^3 replaced by $2\sigma^3\sqrt{2/\pi}$. Using a normally distributed distance with $\sigma = d$ thus increases the expected effect of volume exclusion. However, taking $\sigma \to 0$ whilst holding $\hat{\alpha} = \alpha\sigma^2/2$ constant, as before, regains equation (4.29). In addition, we note that taking a sum of dirac delta functions

$$f_d(u_d) = (\delta(-d) + \delta(d))/2,$$

so that individuals move a fixed distance to the left or right, recovers equation (4.28).

Numerical solutions of this model also compare well with averaged realisations of the IBM (Figure 4.8) for a range of σ, R and α. The shape of the R-N surfaces are very similar to the fixed movement distance case, and have no significant qualitative differences (Figure 4.9). However, the error in our approximations is somewhat reduced, and the effects of exclusion are also reduced since we use the same fixed value of σ that we used for d in Figure 4.6 rather than a scaled value. Similarly, the shape of the d-N surfaces are similar to those for a fixed movement distance but with slightly lower values (result not shown). Whilst the shape of the R-d surfaces is very similar to that for a fixed movement distance, the position of minimum volume exclusion effect is shifted in the distributed movement distance case (Figure 4.10). This is due to the additional factor in the exclusion term that is introduced by $\langle d^3\rangle = 2\sigma^3\sqrt{2/\pi}$, which moves the parameter regime in which no first-order exclusion effects are observed to $4R = 2\sigma\sqrt{2/\pi}$. In particular, using a normally distributed jump length requires σ to be greater than the value of d required in the constant jump model to reduce rather than enhance diffusivity.

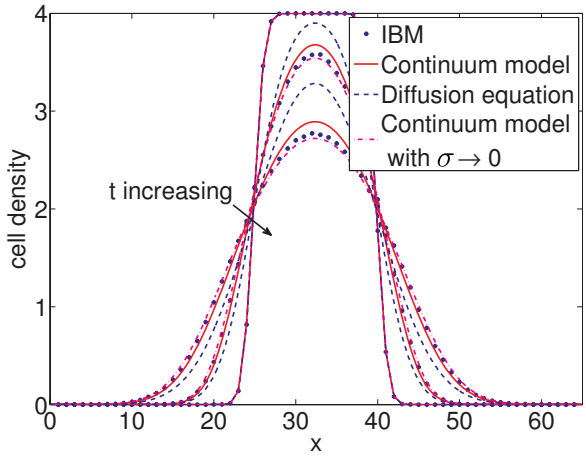


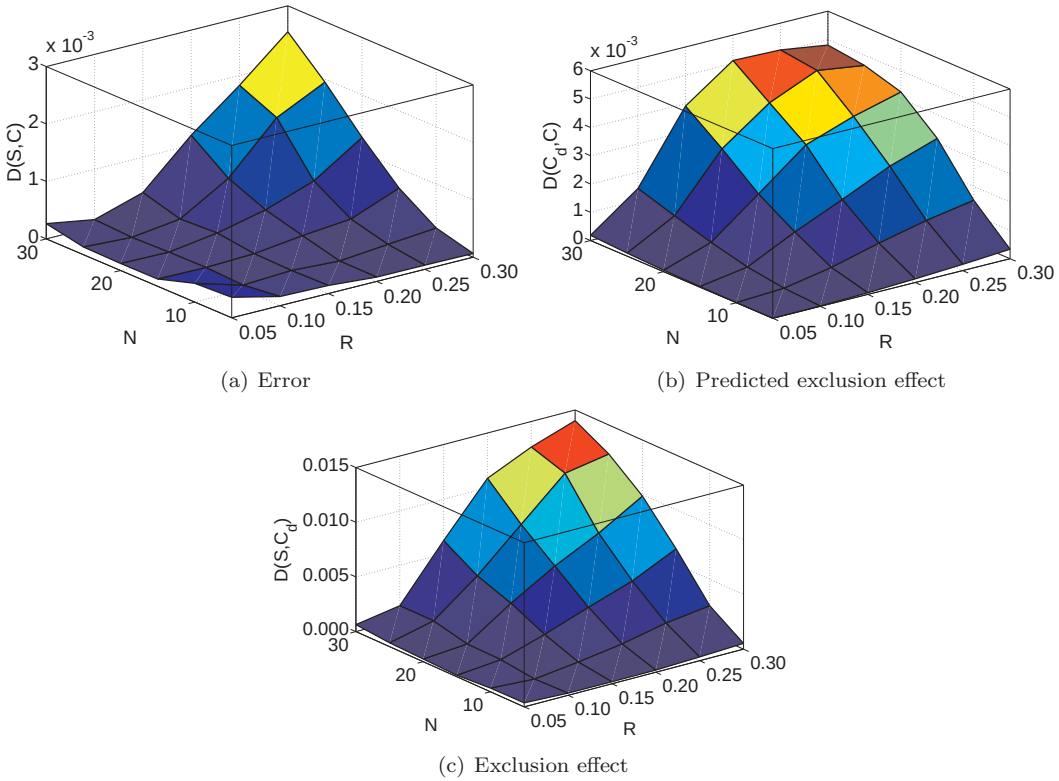
Figure 4.8: Comparison between the IBM and the continuum models with $d \sim \mathcal{N}(0, \sigma^2)$ (equation (4.56)) and as $\sigma \to 0$ (equation (4.31)), plotted at $t = 0$, 200, and 600 for $x \in [0, 65]$. The solution to the diffusion equation is also shown for comparison. Parameters are $N = 60$ and $\mathrm{d}x = 0.05$, with other parameter values given in Table 4.1. For more details about the simulations see Section 4.4.

4.8 Domain growth

Domain growth is rarely incorporated into on-lattice models of movement, since growth necessitates using a time-varying lattice. It is difficult to determine an accurate way to evolve the lattice as the tissue grows whilst maintaining a reasonable level of detail (discussed in Baker *et al.*, 2010; Yates *et al.*, 2012). A big advantage of our off-lattice model is the ability to sidestep issues of lattice size and spacing, and move directly to a domain growth description using continuous space. The derivation of continuum domain growth equations then follows naturally from our earlier derivation (Section 4.3), by considering typical domain growth arguments used in PDE based modelling (Landman *et al.*, 2003). We discuss this in detail in this section.

4.8.1 Model description

We consider uniform domain growth, in which we fix $B_L = 0$ and take $B_R = L(t)$, so that for each time step of length Δt a cell at position $X(t)$ at time t is advected by domain growth to $X(t+\Delta t) = L(t+\Delta t)X(t)/L(t)$ at time $t+\Delta t$. We note that whilst domain contraction does occur in some biological and physical situations, a reduction in domain size may force individuals closer to each other than is permitted under volume exclusion rules. Since it is unclear how volume exclusion should be implemented in the case of domain contraction and, furthermore, there is no contraction in our system of interest, we will only consider domain expansion here.

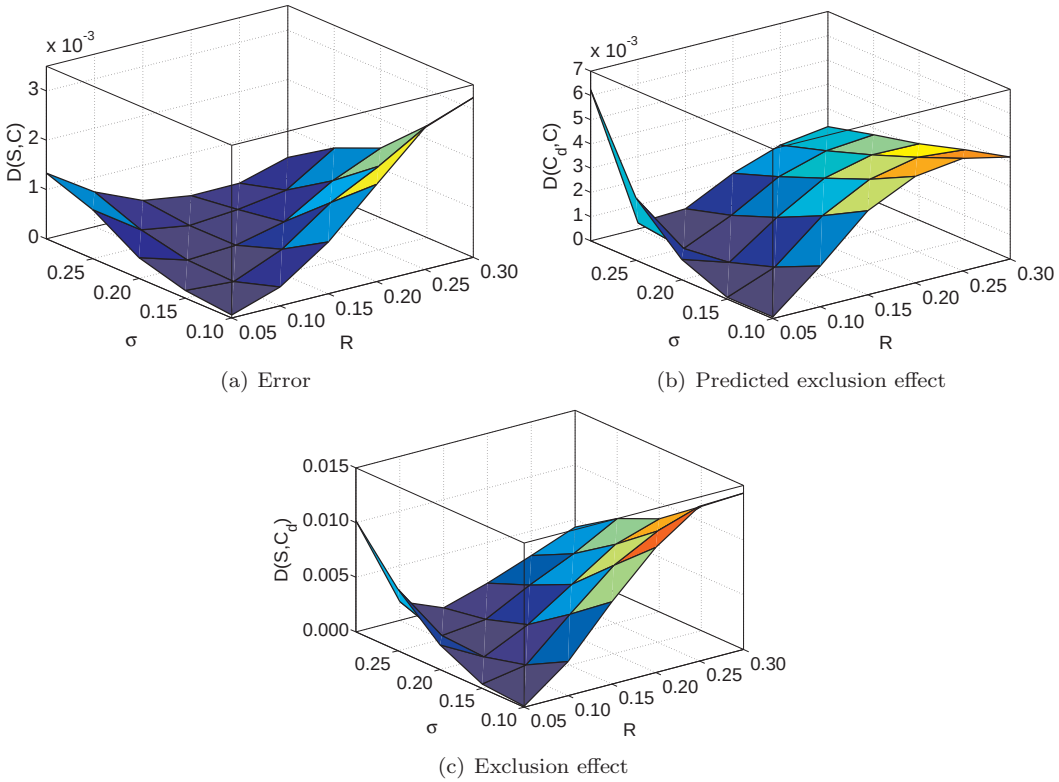


(a) Error

(b) Predicted exclusion effect

(c) Exclusion effect

Figure 4.9: Changing the occupied space (by changing N and R) with normally distributed jump distance (Section 4.7). We show (a) the error introduced by our approximations ($D(C, S)$), (b) our prediction of the effect of volume exclusion ($D(C, C_d)$) and (c) the actual effect of volume exclusion ($D(S, C_d)$). Here S is the average distribution of simulations of the IBM, C is the solution to equation (4.56), C_d is the solution to the diffusion equation and the relative difference, D, is given in equation (4.46). Parameter values are given in Table 4.1. For more details about the simulations see Section 4.4.



(a) Error

(b) Predicted exclusion effect

(c) Exclusion effect

Figure 4.10: Changing the likelihood of successful movement (by changing d and R whilst keeping $\hat{\alpha} = \alpha d^2/2 = 0.025$ constant) with normally distributed jump distance (Section 4.7). We show (a) the error introduced by our approximations ($D(C, S)$), (b) our prediction of the effect of volume exclusion ($D(C, C_d)$) and (c) the actual effect of volume exclusion ($D(S, C_d)$). Here S is the average distribution of simulations of the IBM, C is the solution to equation (4.58), C_d is the solution to the diffusion equation and the relative difference, D, is given in equation (4.46). Parameter values are given in Table 4.1. For more details about the simulations see Section 4.4.

4.8.2 Derivation of continuum equations

We will derive continuum equations using a fixed jump distance as an example, so that it is clear how domain growth changes the derivation. An exactly analogous derivation may be used in the case of variable jump distances.

When considering domain growth, it is necessary to recall that for a continuous pdf, $C_i(x,t)$, the probability of a cell centre being at any particular point x is zero. Instead we consider the probability of a cell centre being in an interval $[x, x+\delta x)$, for $\delta x \ll 1$:

$$\text{Prob(cell centre } i \text{ in } [x, x+\delta x)) = \int_0^{\delta x} C_i(x+\bar{x}, t)\ \mathrm{d}\bar{x}, \tag{4.59}$$

$$\approx \delta x C_i(x,t). \tag{4.60}$$

On a stationary domain, the segment δx is of constant size as time increases and does not appear in the derived equations. However, when the domain size is increasing, a segment of size δx at time $t+\Delta t$ only contains cell centres from a smaller segment at time t. Thus it is important to consider how δx varies in time.

For any small time increment, Δt, $L(t+\Delta t) = L(t) + \Delta t L'(t) + \mathcal{O}(\Delta t^2)$. Hence a spatial coordinate $X(t)$ at time t is given at time $t+\Delta t$ by

$$\begin{aligned} X(t+\Delta t) &= \frac{L(t+\Delta t)}{L(t)} X(t), \\ &= X(t)\left(1 + \frac{\Delta t L'}{L}\right) + \mathcal{O}(\Delta t^2). \end{aligned} \tag{4.61}$$

Thus a segment of size δx at time t grows to approximately $\delta x\,(1+\Delta t L'/L)$ at $(t+\Delta t)$ and a segment of size δx at time $(t+\Delta t)$ must have been of size

$$\delta x \frac{1}{1+\frac{\Delta t L'}{L}} \approx \delta x \left(1 - \frac{\Delta t L'}{L}\right), \tag{4.62}$$

at time t. Hence equation (4.18) is modified to give

$$\int_0^{\delta x} C_i(x+\bar{x}, t+\Delta t)\,\mathrm{d}\bar{x} =$$

$$\underbrace{\int_0^{\delta x} C_i\left(\frac{x+\bar{x}}{1+\frac{\Delta t L'}{L}}, t\right)\,\mathrm{d}\bar{x}}_{\text{cell } i \text{ was in the region that became } [x, x+\delta x)}\Bigg[\overbrace{1-\alpha\Delta t}^{\text{no move attempted}} + \underbrace{\frac{\alpha\Delta t}{2}\left(P_L^i\left(\frac{x}{1+\frac{\Delta t L'}{L}}, t\right) + P_R^i\left(\frac{x}{1+\frac{\Delta t L'}{L}}, t\right)\right)}_{\text{attempted move fails}}\Bigg]$$

$$+\underbrace{\alpha\Delta t\int_0^{\delta x} C_i\left(\frac{x+d+\bar{x}}{1+\frac{\Delta t L'}{L}}, t\right)\,\mathrm{d}\bar{x}}_{\text{cell } i \text{ in region that became } [x+d, x+d+\delta x) \text{ tries to move}}\overbrace{\left[1-P_L^i\left(\frac{x+d}{1+\frac{\Delta t L'}{L}}, t\right)\right]}^{\text{attempted move succeeds}}$$

$$+\underbrace{\alpha\Delta t\int_0^{\delta x} C_i\left(\frac{x-d+\bar{x}}{1+\frac{\Delta t L'}{L}}, t\right)\,\mathrm{d}\bar{x}}_{\text{cell } i \text{ in region that became } [x-d, x-d+\delta x) \text{ tries to move}}\overbrace{\left[1-P_R^i\left(\frac{x-d}{1+\frac{\Delta t L'}{L}}, t\right)\right]}^{\text{attempted move succeeds}}, \tag{4.63}$$

assuming that Δt is small enough that at most one cell movement occurs during the interval $(t, t+\Delta t)$. Using the expansion in equation (4.60)

$$\begin{aligned}
&\delta x C_i(x, t+\Delta t)\,\mathrm{d}\bar{x} = \\
&\quad \delta x\left(1-\frac{\Delta t L'}{L}\right) C_i\left(\frac{x}{1+\frac{\Delta t L'}{L}}, t\right)\left[1-\alpha\Delta t+\frac{\alpha\Delta t}{2}\left(P_L^i\left(\frac{x}{1+\frac{\Delta t L'}{L}}, t\right)+P_R^i\left(\frac{x}{1+\frac{\Delta t L'}{L}}, t\right)\right)\right] \\
&\quad +\alpha\Delta t\delta x\left(1-\frac{\Delta t L'}{L}\right) C_i\left(\frac{x+d}{1+\frac{\Delta t L'}{L}}, t\right)\left[1-P_L^i\left(\frac{x+d}{1+\frac{\Delta t L'}{L}}, t\right)\right] \\
&\quad +\alpha\Delta t\delta x\left(1-\frac{\Delta t L'}{L}\right) C_i\left(\frac{x-d}{1+\frac{\Delta t L'}{L}}, t\right)\left[1-P_R^i\left(\frac{x-d}{1+\frac{\Delta t L'}{L}}, t\right)\right].
\end{aligned} \tag{4.64}$$

Now we may Taylor expand C_i and P_i in Δt, divide by δx, and note that all but two of the extra terms introduced are of order $(\Delta t)^2$, so that

$$\begin{aligned}
C_i(x, t+\Delta t) = C_i(x,t)\Big[1-\alpha\Delta t+\frac{\alpha\Delta t}{2}\left(P_L^i(x,t)+P_R^i(x,t)\right)\Big] + C_i(x+d,t)\frac{\alpha\Delta t}{2}\left[1-P_L^i(x+d,t)\right] \\
+ C_i(x-d,t)\frac{\alpha\Delta t}{2}\left[1-P_R^i(x-d,t)\right] - \Delta t\frac{\partial}{\partial x}\left(\frac{L'x}{L}C_i(x,t)\right)+\mathcal{O}(\Delta t^2).
\end{aligned} \tag{4.65}$$

This is a similar equation to the non-growing case (equation (4.18)), but with an extra term for the dilution due to growth. The derivation is exactly analogous to PDE-based modelling approach, and the dilution due to growth in the stochastic system results from the change in size of the infinitesimal

segment δx. Hence for fixed jump distance, d, letting $\Delta t \to 0$ and using Section 4.3 gives:

$$\frac{\partial C}{\partial t} + \frac{\partial}{\partial x}\left(\frac{L'x}{L}C\right) = \frac{\alpha d^2}{2}\frac{\partial}{\partial x}\left(\left(1 + (4R-d)\frac{(N-1)}{N}C\right)\frac{\partial C}{\partial x}\right) + \mathcal{O}\left(d^{2+m}R^n\right), \tag{4.66}$$

for $n+m \geq 2$ and $C = \sum_{i=1}^{N} C_i$. Taking the limit as $d \to 0$ with $\hat{\alpha} = \alpha d^2/2$ constant gives

$$\frac{\partial C}{\partial t} + \frac{\partial}{\partial x}\left(\frac{L'x}{L}C\right) = \hat{\alpha}\frac{\partial}{\partial x}\left(\left(1 + 4R\frac{(N-1)}{N}C\right)\frac{\partial C}{\partial x}\right) + \mathcal{O}(R^2). \tag{4.67}$$

For normally-distributed jump distances, using Section 4.7.1,

$$\frac{\partial C}{\partial t} + \frac{\partial}{\partial x}\left(\frac{L'x}{L}C\right) = \frac{\alpha\sigma^2}{2}\frac{\partial^2 C}{\partial x^2} + \frac{\alpha\sigma^2}{2}\frac{N-1}{N}\left(4R - 2\sigma\sqrt{\frac{2}{\pi}}\right)\frac{\partial}{\partial x}\left(C\frac{\partial C}{\partial x}\right) + \mathcal{O}(\sigma^{2+m}R^n), \tag{4.68}$$

and as $\sigma \to 0$ with $\hat{\alpha} = \alpha\sigma^2/2$ constant, we recover equation (4.67).

To solve numerically, we rescale onto a fixed domain with $X = x/L$ and $T = t$. Then

$$\frac{\partial C}{\partial t} = \frac{\partial C}{\partial T} + \frac{\partial C}{\partial X}\left(-\frac{XL'}{L}\right), \qquad \frac{\partial C}{\partial x} = \frac{1}{L}\frac{\partial C}{\partial X}, \tag{4.69}$$

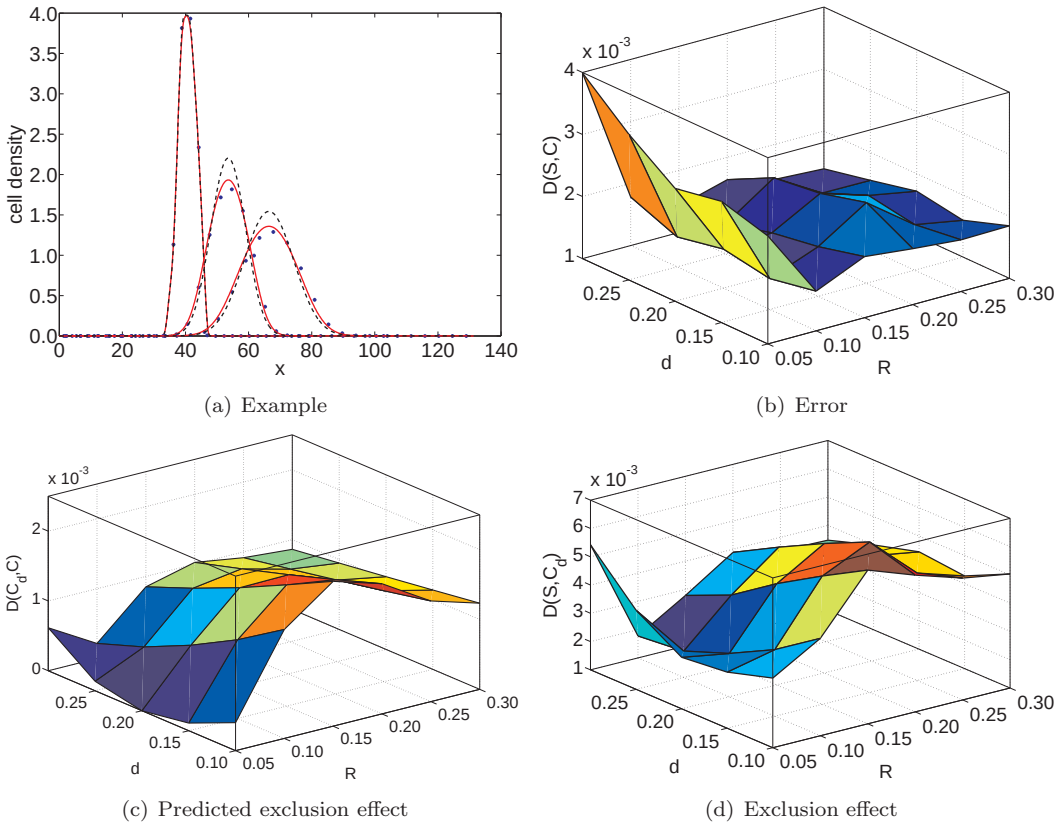
and, for a fixed jump distance, d, rescaling equation (4.66),

$$\frac{\partial C}{\partial T} + \frac{L'}{L}C = \frac{\alpha d^2}{2L^2}\frac{\partial}{\partial X}\left(\left(1 + (4R-d)\frac{(N-1)}{N}C\right)\frac{\partial C}{\partial X}\right) + \mathcal{O}\left(d^{2+m}R^n\right). \tag{4.70}$$

4.8.3 Simulations

Simulations of the extended model with domain growth compare well with numerical solutions to equation (4.70) (Figure 4.11(a)). However, the inclusion of an expanding domain does reduce the effects of volume exclusion when compared to a static migratory domain (Figures 4.11 to 4.13). Volume exclusion effects are reduced since there are fewer cell-cell interactions due to the general dilution of the cell concentration, so that cells encounter each other less frequently. In addition, however, the error between equation (4.70) and averaged simulations is also reduced, and there is still a clear distinction in our simulations between the diffusion equation with domain growth and equation (4.70).

We anticipate that biased migratory mechanisms may mitigate the diminishing effect of exclusion, as individuals encounter other cells more often in spite of the dilution effects of domain expansion (see



(a) Example

(b) Error

(c) Predicted exclusion effect

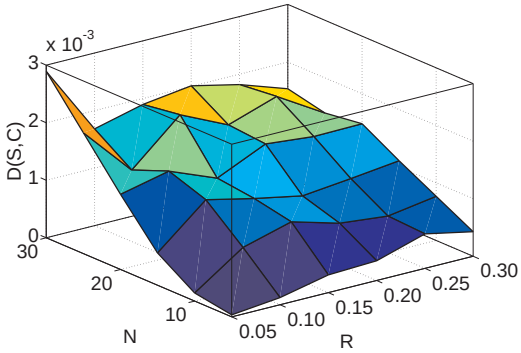
(d) Exclusion effect

Figure 4.11: Changing the likelihood of successful movement (by changing d and R whilst keeping $\hat{\alpha} = \alpha d^2/2 = 0.025$ constant) on a growing domain (Section 4.8). We show (a) an example simulation for $R = d = 0.15$, (b) the error introduced by our approximations ($D(C, S)$), (c) our prediction of the effect of volume exclusion ($D(C, C_d)$) and (d) the actual effect of volume exclusion ($D(S, C_d)$). Here S is the average distribution of simulations of the IBM, C is the solution to equation (4.68), C_d is the solution to the diffusion equation with domain growth and the relative difference, D, is given in equation (4.46). Parameter values are given in Table 4.1. For more details about the simulations see Section 4.4.

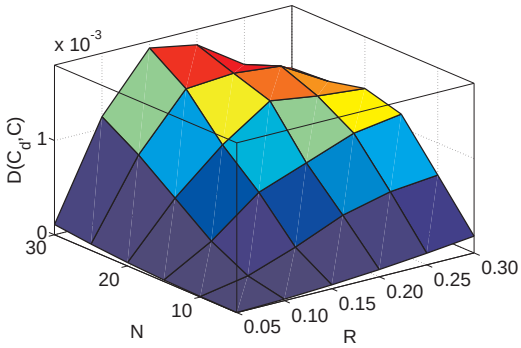
Section 4.9 and Chapter 6). We will therefore next consider how to extend our analysis technique to include the effects of biased migration.

4.9 Biased movement

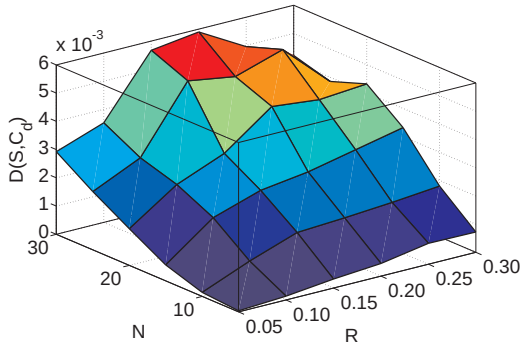
As a first attempt at investigating chemotactic movement, as modelled in Chapters 2 and 3, we will consider introducing a simple bias to the movement rates used due to the changing gradient of a given chemoattractant function, $M(x, t)$. An initial approximation to chemotaxis is to bias the cell movement rates so that cells move right with rate $\alpha/2 + \epsilon \partial M/\partial x$ and left with rate $\alpha/2 - \epsilon \partial M/\partial x$, where the factor ϵ is used to scale the magnitude of $M(x, t)$ so that all rates remain positive.



(a) Error

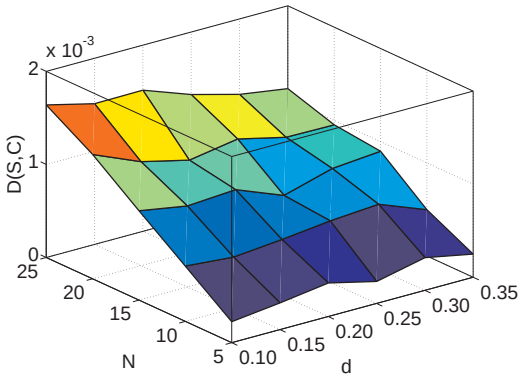


(b) Predicted exclusion effect

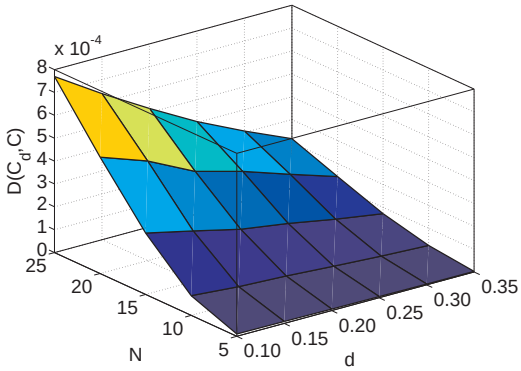


(c) Exclusion effect

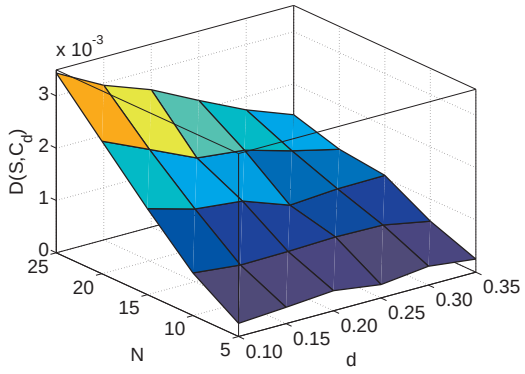
Figure 4.12: Changing the likelihood of successful movement (by changing R and N) on a growing domain (Section 4.8). We show (a) the error introduced by our approximations ($D(C, S)$), (b) our prediction of the effect of volume exclusion ($D(C, C_d)$) and (c) the actual effect of volume exclusion ($D(S, C_d)$). Here S is the average distribution of simulations of the IBM, C is the solution to equation (4.68), C_d is the solution to the diffusion equation with domain growth and the relative difference, D, is given in equation (4.46). Parameter values are given in Table 4.1. For more details about the simulations see Section 4.4.



(a) Error



(b) Predicted exclusion effect



(c) Exclusion effect

Figure 4.13: Exploring the effect of changing the number of cells, N, and the distance moved in a jump, d, on a growing domain (Section 4.8). We show (a) the error introduced by our approximations ($D(C, S)$), (b) our prediction of the effect of volume exclusion ($D(C, C_d)$) and (c) the actual effect of volume exclusion ($D(S, C_d)$). Here S is the average distribution of simulations of the IBM, C is the solution to equation (4.68), C_d is the solution to the diffusion equation with domain growth and the relative difference, D, is given in equation (4.46). Parameter values are given in Table 4.1. For more details about the simulations see Section 4.4.

Hence cells have movement that is biased up the gradient of chemoattractant. There are two main drawbacks to this method of incorporating chemotaxis: firstly, it requires that $\alpha/2 > |\epsilon \partial M/\partial x|$ so that all movement rates are positive; and secondly, we have not allowed for any feedback from the cells to the chemoattractant. We cannot analyse the system described in Chapters 2 and 3 using the form of biased movement introduced in this section, since the mechanisms are different. In particular, in Chapters 2 and 3 chemotaxis does not directly reduce movement rates, instead affecting the probability of successful movement so that movement is only allowed in favourable directions. In contrast, the model presented here uses a predetermined gradient and directly increases the rate of movement towards higher chemoattractant concentrations. Nevertheless, it is a useful first step towards coupled chemotaxis, and enables us to determine whether volume exclusion is important in a system incorporating biased movement before looking in more depth at coupled systems and hybrid models in Chapter 6.

4.9.1 Derivation of continuum equations

The derivation of the continuum equation follows as in Section 4.3, with $\alpha/2 \pm \epsilon \partial M/\partial x$ in the place of $\alpha/2$. The pdf for the position of the ith cell centre is therefore given by:

$$\begin{aligned} C_i(x,t+\Delta t) = \underbrace{C_i(x,t)}_{\text{cell } i \text{ at } (x,t)}\Big[\overbrace{1-\alpha\Delta t}^{\text{no move attempted}} + \underbrace{\left(\frac{\alpha}{2} - \epsilon\frac{\partial M}{\partial x}(x,t)\right)\Delta t P_L^i(x,t) + \left(\frac{\alpha}{2} + \epsilon\frac{\partial M}{\partial x}(x,t)\right)\Delta t P_R^i(x,t)}_{\text{attempted move fails}}\Big] \\ + \overbrace{C_i(x-d,t)\left(\frac{\alpha}{2} + \epsilon\frac{\partial M}{\partial x}(x-d,t)\right)\Delta t}^{\text{cell } i \text{ at } (x-d,t) \text{ tries to move right}}\underbrace{(1-P_R^i(x-d,t))}_{\text{attempted move succeeds}} \\ + \overbrace{C_i(x+d,t)\left(\frac{\alpha}{2} - \epsilon\frac{\partial M}{\partial x}(x+d,t)\right)\Delta t}^{\text{cell } i \text{ at } (x+d,t) \text{ tries to move left}}\underbrace{(1-P_L^i(x+d,t))}_{\text{attempted move succeeds}}. \end{aligned} \tag{4.71}$$

As before we may expand $C_i(x \pm d, t)$ and $\partial M/\partial x(x \pm d, t)$ to find that $C = \sum_i C_i$ satisfies

$$\begin{aligned} \frac{\partial C}{\partial t} = 2d\epsilon\frac{\partial}{\partial x}\left(\left(d\frac{N-1}{N}C - 1\right)C\frac{\partial M}{\partial x}\right) + \frac{\alpha d^2}{2}\frac{\partial^2 C}{\partial x^2} + \frac{N-1}{N}\frac{\alpha d^2}{2}(4R-d)\frac{\partial}{\partial x}\left(C\frac{\partial C}{\partial x}\right) \\ + \mathcal{O}\left(d^{2+n}R^m + \epsilon d^2\right), \end{aligned} \tag{4.72}$$

where $n+m \geq 1$ and if, for simplicity, we consider $M(x,t) = x$ then

$$\frac{\partial C}{\partial t} = 2d\epsilon \frac{\partial}{\partial x}\left(\left(d\frac{N-1}{N}C - 1\right)C\right) + \frac{\alpha d^2}{2}\frac{\partial^2 C}{\partial x^2} + \frac{N-1}{N}\frac{\alpha d^2}{2}(4R-d)\frac{\partial}{\partial x}\left(C\frac{\partial C}{\partial x}\right) + \mathcal{O}\left(d^{2+n}R^m\right). \tag{4.73}$$

As in Section 4.3, we take the limit as $d \to 0$ and allow α to scale with d so that $\hat{\alpha} = \lim_{d\to 0} \alpha d^2/2$ is finite. In addition, we require the bias term to scale with d, so that $\hat{\epsilon} = \lim_{d\to 0} 2d\epsilon$ is finite. Hence

$$\frac{\partial C}{\partial t} = -\hat{\epsilon}\frac{\partial C}{\partial x} + \hat{\alpha}\frac{\partial}{\partial x}\left(\left(1 + 4R\frac{(N-1)}{N}C\right)\frac{\partial C}{\partial x}\right) + \mathcal{O}\left(R^2\right). \tag{4.74}$$

However, as noted in Section 4.3.3, all simulations must be carried out with $d > 0$, and hence $2d^2\epsilon\partial/\partial x(C^2)(N-1)/N$ may be significant and will be included in our numerical solutions.

Indeed, when simulating the system the neglected volume exclusion term $2d^2\epsilon\partial/\partial x(C^2)(N-1)/N$ is found to be important for correctly predicting the speed of an initialised group of cells (result not shown). We would expect biased movement to increase the overall impact of volume exclusion, since cells will move together in similar directions and thus maintain a more cohesive group. In such a group, cells will encounter each other more often and we would expect for this to result in a volume exclusion term derived from the biased movement.

When we consider the effect of exclusion predicted by taking the difference between equation (4.74) and the same equation without exclusion, there is a slight dependence on the radius of a cell, as predicted (Figure 4.14(c)). This very small dependence of the effects of volume exclusion to changes in R is validated when we consider the difference between the simulated system with exclusion and the solution to the equivalent naïve equation without exclusion (Figure 4.14(d)).

As we had anticipated, the overall effect of volume exclusion is much greater than in the case without bias, giving $D(S, C_d)$ up to 0.5 compared to at most 0.02 without bias. The reason for this greater exclusion effect is partially due to the decrease in speed of the peak of the cell density profile (Figure 4.14(a)). When individuals cannot move past each other, the speed of propagation in the direction of the bias is much slower. In addition, individuals at the back of the group move successfully much less frequently and hence the back of the concentration profile is steeper than the front of the wave, where the biased movement is into unoccupied areas. The difference in movement rates between the back and the front of the wave makes the profile of the distribution more skewed than when volume exclusion is not considered. This speed and profile difference is accentuated

for larger cell jump distances (Figures 4.14(c) and 4.14(d)) and for higher cell population numbers (Figures 4.15(b) and 4.15(c)) since these lead to a higher proportion of attempted cell movements being prevented by other individuals, giving an even higher volume exclusion effect. The increased volume exclusion effects for higher N and d is particularly seen in d-N surface plots (Figure 4.16), where the usual shape of the surface is changed so that increasing d no longer decreases the effects of exclusion (as in Figure 4.7(c)), but instead increases exclusion effects (Figure 4.16(c)).

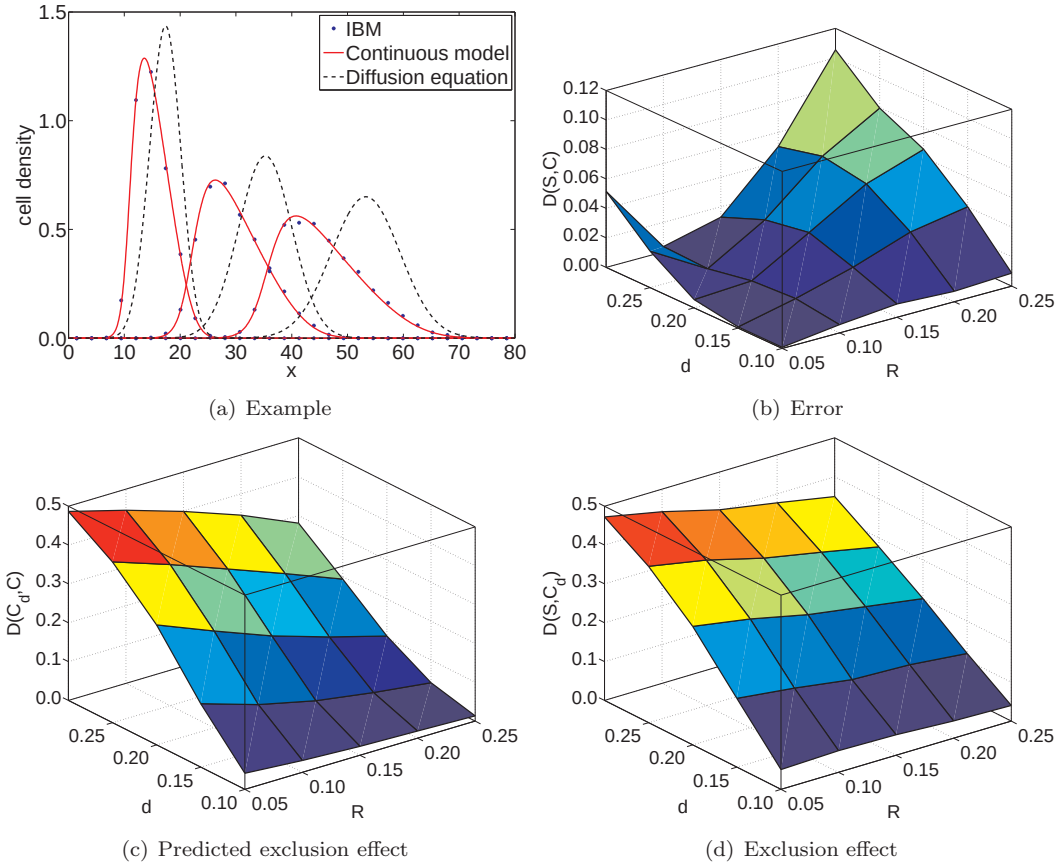
The greater effect of excluded volume compared to low errors between our predictions and the simulated model make the biased movement an example of a mechanism where our analysis technique is very valuable compared to the naïve equations that might otherwise be used. Since the chemotactic system that we wish to analyse will also, in effect, bias individual cell movements, we expect that volume exclusion will also be important in that analysis (Chapter 6).

4.10 Discussion

In this chapter we have demonstrated our technique for analysing one-dimensional IBMs by deriving PDE approximations to averaged simulations. This technique allows us to systematically derive continuum equations from underlying microscopic rules rather than using large-scale building blocks to construct model equations. In particular, we have shown the importance of cell volume and their occupation of space. Understanding the role of cell crowding in biological systems is highly important, since there may be many cells present in a small domain volume with a high degree of cell-cell interaction.

Our framework enables the systematic derivation of continuum equations in which the size of an individual is present explicitly in the equations. Since the radius is explicitly included, these equations could also in the future be used to investigate cells that change size over time, perhaps in response to environmental pressures. For example, the radius of an individual could depend on the total population size, so that individuals are compressed in more crowded conditions. In addition, it is straightforward to consider multiple populations of cells, with differing sizes and movement rates. It is more interesting, however, to consider multiple cell populations in two dimensions (see Section 5.5), since in the one-dimensional formulation individuals cannot move past each other, thus restricting the interaction of separate populations.

We have shown that there are parameter regimes in which the error due to our approximations is



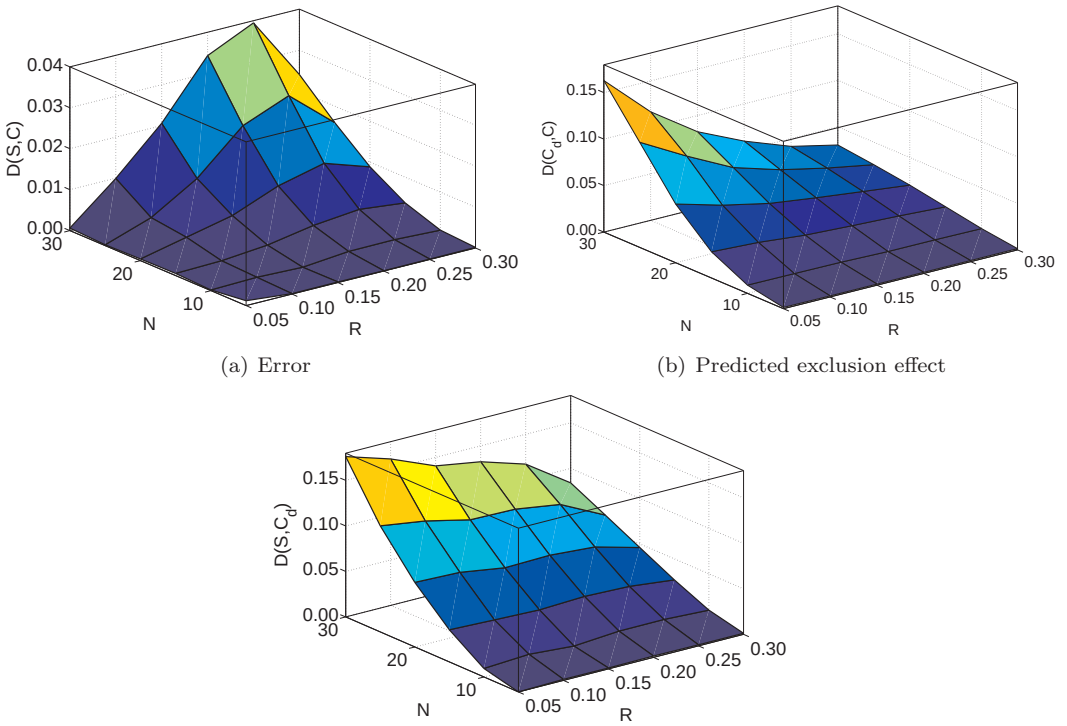
(a) Example

(b) Error

(c) Predicted exclusion effect

(d) Exclusion effect

Figure 4.14: Changing the likelihood of successful movement (by changing d and R whilst keeping $\hat{\alpha} = \alpha d^2/2 = 0.025$ constant) with biased movement (Section 4.9). We show (a) an example simulation for $R = d = 0.15$, (b) the error introduced by our approximations ($D(C, S)$), (c) our prediction of the effect of volume exclusion ($D(C, C_d)$) and (d) the actual effect of volume exclusion ($D(S, C_d)$). Here S is the average distribution of simulations of the IBM, C is the solution to equation (4.73), C_d is the solution to the corresponding equation without volume exclusion and the relative difference, D, is given in equation (4.46). Parameter values are given in Table 4.1. For more details about the simulations see Section 4.4, with the difference that the centre of the first cell is given an initial position drawn from $\mathcal{N}(10, 0.1)$ to give room to migrate in the direction of bias.

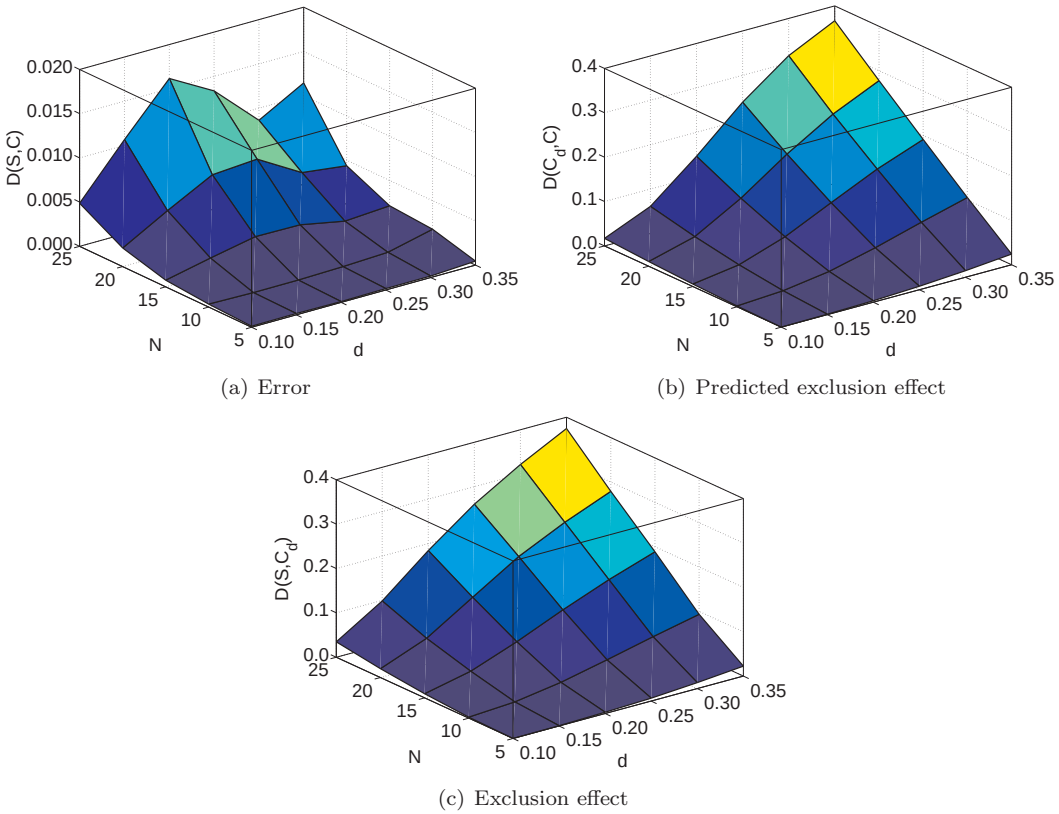


(a) Error

(b) Predicted exclusion effect

(c) Exclusion effect

Figure 4.15: Changing the occupied space (by changing R and N) with biased movement (Section 4.9). We show (a) the error introduced by our approximations ($D(C, S)$), (b) our prediction of the effect of volume exclusion ($D(C, C_d)$) and (c) the actual effect of volume exclusion ($D(S, C_d)$). Here S is the average distribution of simulations of the IBM, C is the solution to equation (4.73), C_d is the solution to the corresponding equation without volume exclusion and the relative difference, D, is given in equation (4.46). Parameter values are given in Table 4.1. For more details about the simulations see Section 4.4, with the difference that the centre of the first cell is given an initial position drawn from $\mathcal{N}(10, 0.1)$ to give room to migrate in the direction of bias.



(a) Error

(b) Predicted exclusion effect

(c) Exclusion effect

Figure 4.16: Exploring the effect of changing the number of cells, N, and the distance moved in a jump, d, with biased movement (Section 4.9). We show (a) the error introduced by our approximations ($D(C, S)$), (b) our prediction of the effect of volume exclusion ($D(C, C_d)$) and (c) the actual effect of volume exclusion ($D(S, C_d)$). Here S is the average distribution of simulations of the IBM, C is the solution to equation (4.73), C_d is the solution to the corresponding equation without volume exclusion and the relative difference, D, is given in equation (4.46). Parameter values are given in Table 4.1. For more details about the simulations see Section 4.4, with the difference that the centre of the first cell is given an initial position drawn from $\mathcal{N}(10, 0.1)$ to give room to migrate in the direction of bias.

low compared to the impact of volume exclusion. Hence we conclude that it is important to use our equations in modelling systems when a detailed examination of cell movement in a crowded environment is required. When biased movement is included, our equations accurately reproduce both the speed of a group of cells and the wavefront profile as time increases, which are not well represented by the continuum equations that are derived when volume exclusion is not taken into account.

However, we note that one-dimensional models are a special case for volume exclusion since individuals must always maintain their initial order in the domain, which is not the case in systems with higher spatial dimensions. Truly one-dimensional biological systems are very rare, so that most one-dimensional models consider a two- or three-dimensional situation that has been averaged.

To address these questions about higher dimensional systems, and enable analysis of our biologically motivated model from Chapters 2 and 3, we will extend our analytical technique to two-dimensional systems in Chapter 5.

Chapter 5

Reaching for the second dimension

In the previous chapter we began by investigating simple diffusive off-lattice IBMs with volume exclusion and progressed to consider domain growth and simple biased movement. Since our ultimate aim is to be able to analyse the complex two-dimensional hybrid IBM developed in Chapters 2 and 3, we must first extend the models in Chapter 4 into two spatial dimensions before considering the dynamic feedback between the two modelling levels (see Chapter 6). Two-dimensional modelling gives greater insight into the movement of populations where individuals can migrate around one another and form two-dimensional groups of cells. Indeed, there are many biological systems and population crowding situations that are inherently two-dimensional and thus cannot be accurately represented by a one-dimensional system.

Analysing volume exclusion in a two-dimensional off-lattice model brings new challenges to the techniques used in Chapter 4, requiring the integration of pdfs over the area newly occupied during cell movement. We achieve this by extending the area of integration to one that is more easily integrated over, and then Taylor expand in a similar way to Chapter 4. In this way we again reduce the governing equations of cell density to a local form where the evolution of the density at a point depends only upon the density and derivatives of the density at that point. We then extend the basic two-dimensional model to include domain growth, biased movement and multiple cell populations with differing parameter values.

5.1 Model description

We begin by considering a simple two-dimensional model, analogous to the one in Section 4.1. Cells are defined by a position $\mathbf{x} = (x, y) \in [B_L, B_R] \times [B_B, B_T]$ and occupy a disc of radius R around this point. Individuals move by hopping with rate α a distance $d > 0$ in a random direction θ. An attempted move is aborted if it would result in two individuals overlapping (*i.e.* if the centres of two cells would be closer than $2R$ apart). If an individual attempts to move out of the domain, it is reflected back from the boundary, so that all parts of a cell remain in the region. This model is simulated using the Gillespie Algorithm, similarly to Section 4.2.

5.2 Derivation of continuum equations

If $C_i(\mathbf{x}, t)$ is the pdf of the position, $\mathbf{x} = (x, y)$, of the ith cell centre at time t, then we consider the probability that the centre of cell i is in the region $[x, x + \delta x) \times [y, y + \delta y)$ for $\delta x, \delta y \ll 1$:

$$\int_x^{x+\delta x} \int_y^{y+\delta y} C_i(\mathbf{x}, t)\, \mathrm{d}x\, \mathrm{d}y \approx \delta x \delta y C_i(\mathbf{x}, t). \tag{5.1}$$

We use a similar method to Section 4.3, and expand probabilities by integrating over the direction, θ, chosen by cell i (see Section 4.7.1). Hence for $\Delta t \ll 1$, dividing by δx and δy,

$$\begin{aligned} C_i(\mathbf{x}, t + \Delta t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \{ & \underbrace{C_i(\mathbf{x}, t)}_{\text{cell } i \text{ at } (\mathbf{x}, t)} \Big[\overbrace{1 - \alpha \Delta t}^{\text{no move attempted}} + \underbrace{\alpha \Delta t P_i(\mathbf{x}, \theta, t)}_{\text{attempted move fails}} \Big] \\ & + \overbrace{\alpha \Delta t C_i(\mathbf{x} - \mathbf{d}_\theta, t)}^{\text{cell } i \text{ at } \mathbf{x} - \mathbf{d}_\theta \text{ tries to move}} \underbrace{[1 - P_i(\mathbf{x} - \mathbf{d}_\theta, \theta, t)]}_{\text{attempted move succeeds}} \} \, \mathrm{d}\theta + \mathcal{O}(\Delta t^2), \end{aligned} \tag{5.2}$$

where $\mathbf{d}_\theta = (d \cos\theta, d \sin\theta)$ and $P_i(\mathbf{x}, \theta, t)$ is the probability that there is another cell preventing the movement of cell i from position $\mathbf{x}$ in direction θ, given that cell i is at position $\mathbf{x}$. Hence

$$\begin{aligned} \frac{\partial C_i}{\partial t} = & \frac{\alpha}{2\pi} \int_{-\pi}^{\pi} [C_i(\mathbf{x} - \mathbf{d}_\theta, t) - C_i(\mathbf{x}, t)] \, \mathrm{d}\theta \\ & + \frac{\alpha}{2\pi} \int_{-\pi}^{\pi} [P_i(\mathbf{x}, \theta, t) C_i(\mathbf{x}, t) - P_i(\mathbf{x} - \mathbf{d}_\theta, \theta, t) C_i(\mathbf{x} - \mathbf{d}_\theta, t)] \, \mathrm{d}\theta, \end{aligned} \tag{5.3}$$

as $\Delta t \to 0$.

5.2.1 Finding $P_i(\mathbf{x}, \theta, t)$

Next we need to find an expression for the probability $P_i(\mathbf{x}, \theta, t)$ of cell movement being prevented given that cell i is at position $\mathbf{x}$. As before we will make a closure assumption to make progress with the system of equations by assuming that the positions of two randomly chosen cells are approximately independent. Hence $P_i(\mathbf{x}, \theta, t)$ is given by the probability of a cell occupying the region A_i shown in Figure 5.1. Hence

$$P_i(\mathbf{x}, \theta, t) = \sum_{j \neq i} \int_{A_i} C_j(\mathbf{x}, t) \, \mathrm{d}\mathbf{x}. \tag{5.4}$$

To determine the integral over A_i, we change variables to polar coordinates, centred at $\mathbf{x}$. To simplify the algebra in finding an analytic form of $P(\mathbf{x}, \theta, t)$, we will extend the excluded area, A_i, so that movement is prevented if cell i would move through any occupied area thus including the blue areas in Figure 5.1. For a small distance moved this only introduces a small error since the additional area integrated over is very small (of order $\mathcal{O}(d^3/R)$), whilst preventing the limits of ϕ depending on d and R. In addition, it can be argued that requiring an unoccupied route to the cell's new position is more realistic. Thus

$$P_i(\mathbf{x}, \theta, t) = \sum_{j \neq i} \int_0^d \int_{-\pi/2}^{\pi/2} 2R \cos\phi \; C_j(x + r\cos\theta + 2R\cos(\theta + \phi), y + r\sin\theta + 2R\sin(\theta + \phi), t) \; \mathrm{d}\phi \; \mathrm{d}r, \tag{5.5}$$

where, once again, $\mathbf{x} = (x, y)$. We may now Taylor expand this expression in R, d and $r \in [0, d]$, analogous to the technique in Section 4.3. Since

$$\int_{-\pi}^{\pi} \cos\theta \; \mathrm{d}\theta = 0, \qquad \int_{-\pi}^{\pi} \sin\theta \; \mathrm{d}\theta = 0, \tag{5.6}$$

$$\int_{-\pi}^{\pi} \sin\theta\cos\theta \; \mathrm{d}\theta = 0, \qquad \int_{-\pi}^{\pi} \cos^2\theta \; \mathrm{d}\theta = \pi, \tag{5.7}$$

$$\int_{-\pi}^{\pi} \sin^2\theta \; \mathrm{d}\theta = \pi, \qquad \int_{-\pi/2}^{\pi/2} \cos\phi\sin\phi \; \mathrm{d}\phi = 0, \tag{5.8}$$

$$\int_{-\pi/2}^{\pi/2} \cos\phi \; \mathrm{d}\phi = 2, \qquad \int_{-\pi/2}^{\pi/2} \cos^2\phi \; \mathrm{d}\phi = \frac{\pi}{2}, \tag{5.9}$$

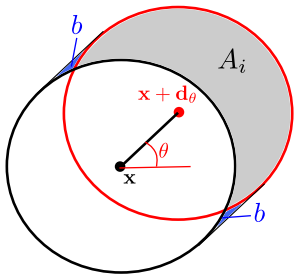


Figure 5.1: $P(\mathbf{x},\theta,t)$ is the probability of a cell occupying the grey shaded region, A_i, shown here. The black circle represents the current position of the cell, with centre at $\mathbf{x}$, whilst the red circle indicates the attempted movement in direction θ, so that the cell centre moves to $\mathbf{x}+\mathbf{d}_\theta$. Since the considered cell occupies the black circle, it is known to be free of other cells. Hence to find whether the movement is prevented, we must find the probability of another cell occupying the grey area, A_i. To simplify the algebra involved, we extend the region of integration to also include b, the blue shaded regions.

we find that the excluded term,

$$E_i = \frac{\alpha}{2\pi}\int_0^{2\pi} \left[P_i(\mathbf{x},\theta,t)C_i(\mathbf{x},t) - P_i(\mathbf{x}-\mathbf{d}_\theta,\theta,t)C_i(\mathbf{x}-\mathbf{d}_\theta,t)\right]\,\mathrm{d}\theta, \tag{5.10}$$

is given by

$$E_i = 2\alpha R^2 d^2 \pi (R\pi - d)(N-1)\left\{C_i\left(\frac{\partial^2 C_i}{\partial x^2}+\frac{\partial^2 C_i}{\partial y^2}\right)+\left[\left(\frac{\partial C_i}{\partial x}\right)^2+\left(\frac{\partial C_i}{\partial y}\right)^2\right]\right\}, \tag{5.11}$$

where N is the number of cells, since if the initial conditions for all the cells are the same, then $C_i = C_j\ \forall i,j$. Hence $C=\sum_i C_i$ satisfies

$$\frac{\partial C}{\partial t} = \frac{\alpha d^2}{4}\nabla^2 C + \frac{\alpha d^4}{64}\nabla^2\left(\nabla^2 C\right) + \frac{N-1}{N}\alpha R d^2 (R\pi - d)\nabla\cdot(C\nabla C) + \mathcal{O}\left((d+R)^5\right), \tag{5.12}$$

and as $d\to 0$ with $\hat{\alpha}=\alpha d^2/4$ held constant (as in Section 4.3)

$$\frac{\partial C}{\partial t} = \hat{\alpha}\nabla\cdot\left(\left(1+\frac{N-1}{N}4\pi R^2 C\right)\nabla C\right). \tag{5.13}$$

We may compare equation (5.13) with the equivalent equation in one dimension (equation (4.31)), which is strikingly similar. Both the one- and two-dimensional models result in the usual diffusion equation with an extra term representing the effects of volume exclusion, dependent on both the concentration and the gradient of C, so that the form of both equations (4.31) and (5.13) is the same. Indeed, the multiplying factor ($4R$ in equation (4.31) and $4\pi R^2$ in equation (5.13)) in both models is the area around an individual in which no other individual can be present due to volume

exclusion.

5.2.2 Initial and boundary conditions

As with the one-dimensional models we must consider carefully how to derive appropriate boundary conditions. Individuals are reflected from the boundaries when their centre is within R of the boundary and so, analogous to Section 4.3.3 we take no flux boundary conditions as a first approximation.

5.2.3 Simulations

Realisations of the IBM were performed using the Gillespie algorithm (see Section 4.2, Gillespie, 1977), with initial conditions as follows. The centre coordinates of the bottom left individual are drawn from a normal distribution, $\mathcal{N}(10 - \sqrt{2RN}/0.8, 1)$, and the remaining cells are placed in a regular grid in the positive x- and y-directions from this first cell with grid spacing $2R/0.8$ to give a cell density of 0.8. For the PDEs we use the Numerical Algorithms Group routine d03ra, which uses the method of lines to reduce the problem to a system of ordinary differential equations (ODEs) and solves the resulting system using a backwards differentiation formula (see Appendix A.2). Initial conditions are determined by the average initial distribution from the simulations, linearly interpolated onto a mesh with spacings dx and dy. All numerical solutions to the PDEs neglect terms referred to as higher order in the equations. Since we are not considering any particular physical or biological problem, we will take all parameter values and variables to be nondimensional and merely note that dimensional systems should be nondimensionalised before using our technique, so that the size of higher order terms may be sensibly compared. Default nondimensional parameter values are given in Table 5.1, with other parameter values given in figures or captions.

5.2.3.1 Flux boundary conditions

We may also derive an appropriate boundary condition when cells are inserted into the domain at a rate P_c. All inserted cells are initialised at $x = R$ with a randomly selected y-coordinate. We may use a similar technique to Section 4.5.1, and integrate over the domain of migration, $\mathcal{D}$, minus strips

Parameter	Typical value
d	0.15
R	0.15
B_L	0
B_R	20
B_B	0
B_T	20
total time	100
total cell number	16
$\mathrm{d}x$	0.5
$\mathrm{d}y$	0.5
number of simulations	2000
P_c	1
α	2.22
σ	0.15
L_a	0.1
ϵ	0.2
ϕ	$\pi/4$

Table 5.1: Typical nondimensional simulation parameters.

of width R next to the boundaries, $\mathcal{D}\backslash\mathcal{R}$. Hence the rate of change of the total cell population is

$$\frac{\partial}{\partial t}\int\limits_{\mathcal{D}\backslash\mathcal{R}} C(\mathbf{x},t)\,\mathrm{d}V = \int\limits_{\mathcal{D}\backslash\mathcal{R}} \frac{\partial C}{\partial t}\,\mathrm{d}V, \tag{5.14}$$

$$= \int\limits_{\mathcal{D}\backslash\mathcal{R}} \hat{\alpha}\nabla\cdot\left(\left(1+\frac{N-1}{N}4R^2C\right)\nabla C\right)\,\mathrm{d}V, \tag{5.15}$$

$$= \int_{\partial(\mathcal{D}\backslash\mathcal{R})} \hat{\alpha}\left(\left(1+\frac{N-1}{N}4R^2C\right)\nabla C\right)\cdot\mathbf{n}\,\mathrm{d}\mathbf{x}, \tag{5.16}$$

where $\mathbf{n}$ is the outward-pointing normal to the boundary. The domain boundary $\partial\left(\mathcal{D}\backslash\mathcal{R}\right)$ can be split into the boundary at $x = B_L + R$ and those at $x = B_R - R$, $y = B_B + R$ and $y = B_T - R$. Since no-flux boundary conditions are taken on the latter three boundaries, then the integral equation (5.16) is given by

$$\int\limits_{\partial(\mathcal{D}\backslash\mathcal{R})} \hat{\alpha}\left(\left(1+\frac{N-1}{N}4R^2C\right)\nabla C\right)\cdot\mathbf{n}\,\mathrm{d}\mathbf{x} =$$

$$-\frac{\alpha d^2}{4}(B_T - B_B - 2R)\left(1+\frac{N-1}{N}4R(R\pi - d)C\right)\left.\frac{\partial C}{\partial x}\right|_{x=R}. \tag{5.17}$$

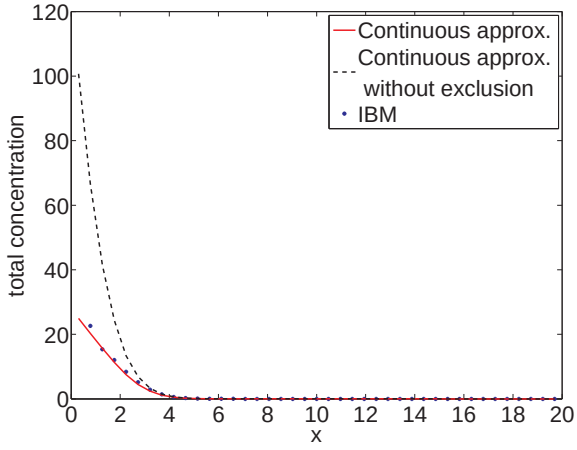


Figure 5.2: We use the expected population size to successfully define a flux boundary condition (see Section 5.2.3.1). Since a y-independent flux along the $x = 0$ boundary results in invariability of concentration over changes to the y-coordinate (result not shown), we integrate over y and plot the resultant one-dimensional profile. Parameter values are given in Table 5.1.

However, the change in total cell population is also given by the flux into the domain so that

$$-\frac{\alpha d^2}{4}(B_T - B_B - 2R)\left(1 + \frac{N-1}{N}4R(R\pi - d)C\right)\frac{\partial C}{\partial x}\bigg|_{x=R} = P_c\left(1 - \int_0^{2R}\int_{-\pi/2}^{\pi/2} rC(\mathbf{x} + \mathbf{r}_\theta^r, t)\ \mathrm{d}r\ \mathrm{d}\theta\right) \tag{5.18}$$

$$= P_c\left(1 - 2\pi R^2 C\right)\Big|_{x=R}, \tag{5.19}$$

Taylor expanding in terms of R. Numerical solutions to equation (5.13) and averaged model simulations with flux boundary condition equation (5.19) are invariant to changes in y-coordinate (result not shown) and thus we integrate numerical solutions to equation (5.13) with equation (5.19) over y to more easily compare with averaged simulations of the IBM. It is clear that volume exclusion has a significant effect, particularly for higher influxes of cells, since individuals prevent later cells from entering the domain (Figure 5.2). Hence using continuum equations without volume exclusion significantly over predicts the concentrations in the domain.

5.2.4 Results

Since an influx of cells give such a large discrepancy between simulations and predictions without volume exclusion, we will compare the usual diffusion equation with equation (5.13) when individuals are initialised in the domain as is described in Section 5.2.3 and no flux boundary conditions are

taken on all boundaries. Numerical solutions of equation (5.13) with no flux boundary conditions compare well with averaged model simulations (Figure 5.3) and still display a significant effect of volume exclusion (Figures 5.3(c) and 5.3(d)). In particular volume exclusion can enhance diffusivity so that the maximum density of cells in the domain is over-estimated when volume exclusion is not taken into account (Figure 5.3(d)), as is predicted by equation (5.12) for $(R\pi - d) > 0$.

5.2.5 Exploring parameter space

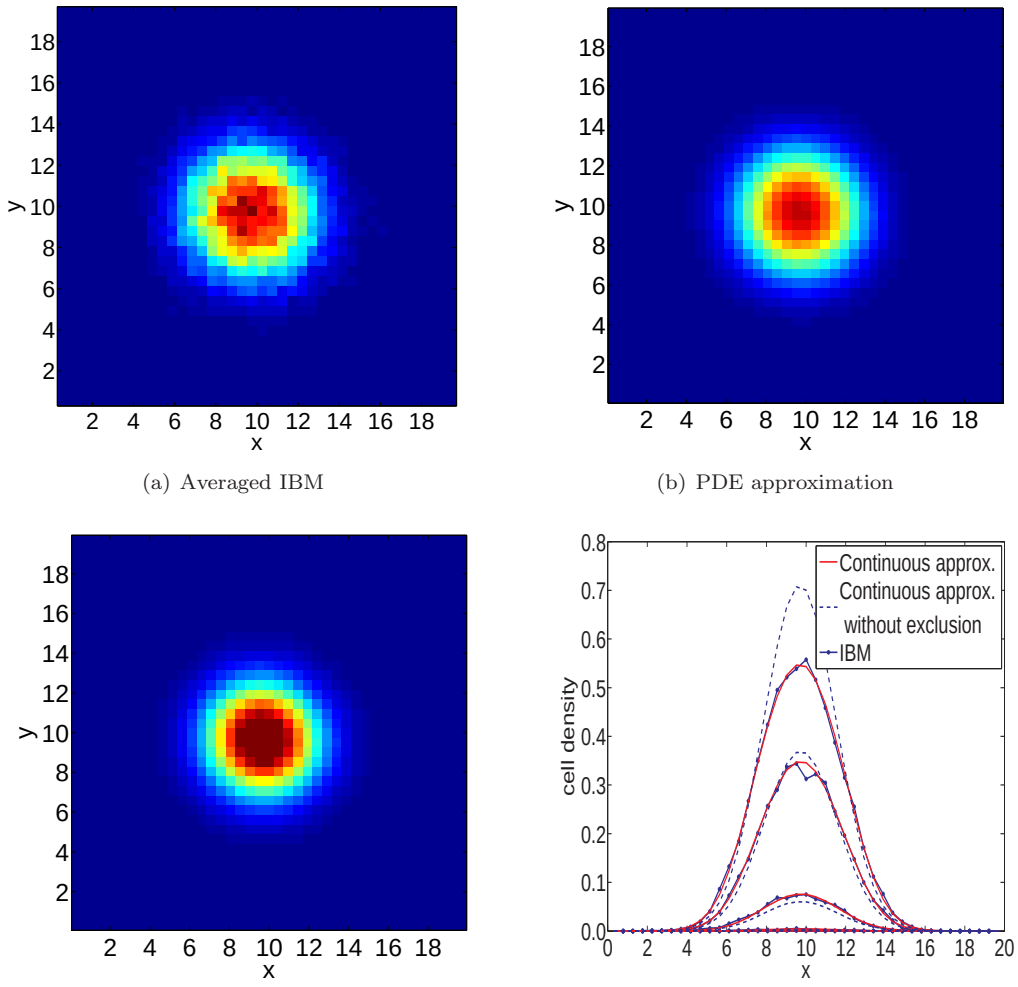
In this section we explore the accuracy of our continuum prediction (equation (5.12)) and the equivalent equation without volume exclusion, by comparing numerical solutions to these equations with averaged model simulations. We also determine the effect of volume exclusion on the cell density profile by comparing averaged simulations with numerical solutions to the diffusion equation (obtained by deriving an equivalent continuum equation without taking volume exclusion into account).

To find the relative difference between two normalised density functions $f_1(\mathbf{x}), f_2(\mathbf{x})$ over a domain, $\mathcal{D}$, we take

$$\mathcal{D}(f_1, f_2) = \int_{\mathcal{D}} \frac{(f_1(\mathbf{x}) - f_2(\mathbf{x}))^2}{f_1(\mathbf{x}) + f_2(\mathbf{x})} \, \mathrm{d}\mathbf{x}. \tag{5.20}$$

Numerical solutions of equation (5.13) continue to compare well to averaged realisations of the model as the cell radius, R, and distance, d, moved during a jump vary (Figure 5.4(a)). In this case the profile of the exclusion surface may be qualitatively predicted by plotting $|R(R\pi - d)|$ (result not shown). The magnitude of the exclusion term is a successful predictor of the effect of exclusion since the error in our predicted equation (Figure 5.4(a)) is an order of magnitude lower than the size of the predicted effect (Figure 5.4(c)), and is relatively constant as R and d change.

Conversely, as the number of cells, N, increases for large R, the error in our approximations also increases (Figure 5.4(b)). Since increasing cell numbers leads to higher correlations between the positions of cells initialised in one group, it is perhaps unsurprising that our moment closure approximation is less exact for larger populations. However, we note that the volume exclusion effect (Figure 5.4(f)) also increases, and for higher cell numbers and $R > 0.05$, equation (5.13) is a much better predictor of the population profile than if volume exclusion is not taken into account. Furthermore, greater population numbers decrease the ratio of the approximation error (Figure 5.4(b)) in equation (5.13) to the error resulting from neglecting exclusion effects (Figure 5.4(f)). Hence for



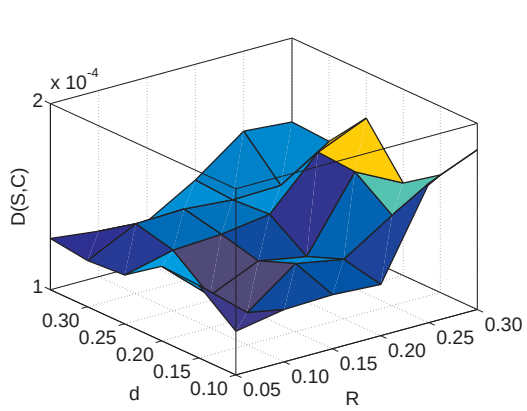
(a) Averaged IBM

(b) PDE approximation

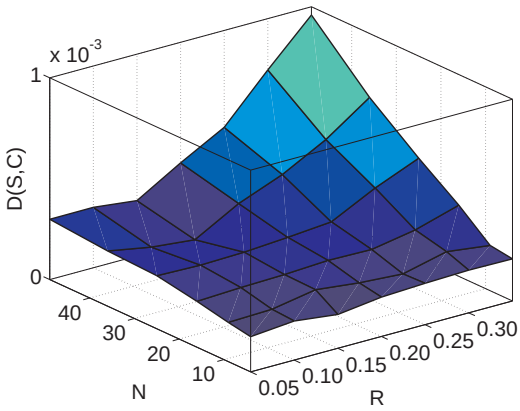
(c) PDE approximation without exclusion

(d) y-axis slices

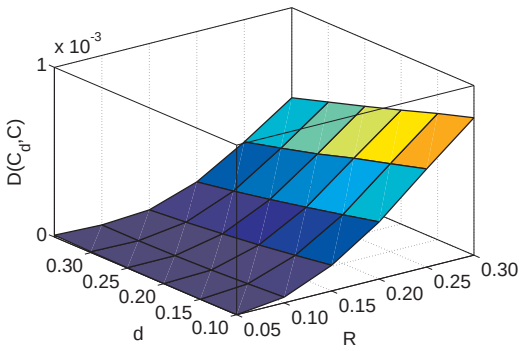
Figure 5.3: Comparison between the IBM and the continuum equation with or without volume exclusion (equation (5.13)). The distance moved is $d = 0.3$, the radius of a cell is $R = 0.3$ and other parameter values are given in Table 5.1. Other details about the simulations are given in Section 5.2.3.



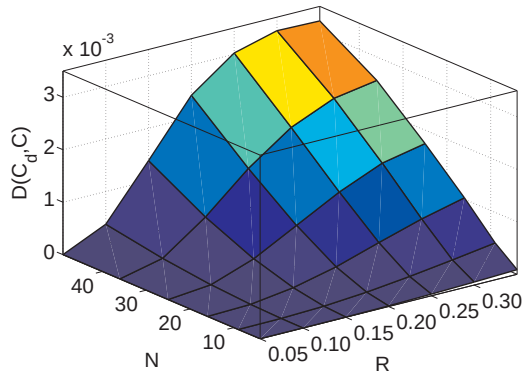
(a) Error over $d - R$ space



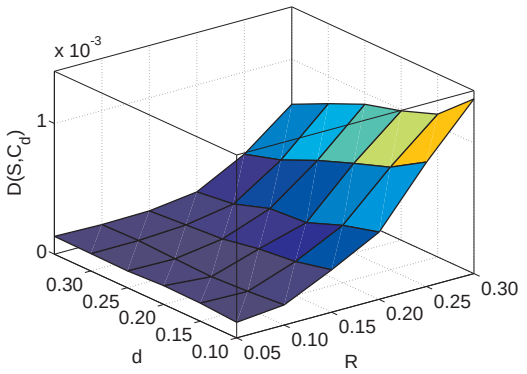
(b) Error over $N - R$ space



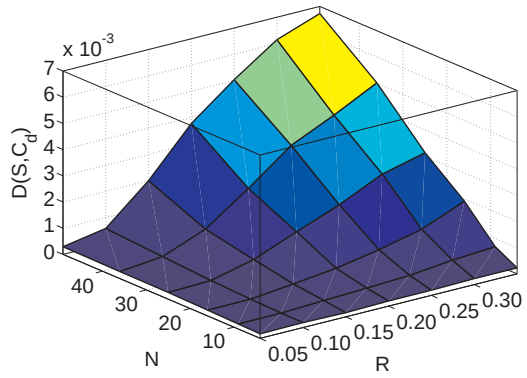
(c) Predicted exclusion effect over $d - R$ space



(d) Predicted exclusion effect over $N - R$ space



(e) Exclusion effect over $d - R$ space



(f) Exclusion effect over $N - R$ space

Figure 5.4: Exploring the effect of changing the likelihood of successful movement in the model by changing the distance moved, d, and the radius of a cell, R ((a), (c) and (e)) and changing the occupied space in the model by changing the number of cells, N, and the radius of a cell, R ((b), (d) and (f)). We consider the effect on: (a)-(b) the error introduced by our approximations (the difference between the solution to equation (5.13) and the average of our simulations); (c)-(d) our prediction of the effect of volume exclusion (the difference between the solution to our equation and the solution to the diffusion equation); and (e)-(f) the actual effect of volume exclusion (the difference between the average of our simulations and the solution to the diffusion equation). For more details, see Sections 4.6 and 5.2.5. Other parameter values are given in Table 5.1 with $\alpha = 0.05/d^2$ so that $\hat{\alpha} = 0.0125$ is constant as d changes. Other details about the simulations are given in Section 5.2.3.

larger cells and for greater cell populations equation (5.13) is a better average predictor than naïvely neglecting volume exclusion effects.

5.3 Distribution of jump distances

In one dimension (Section 4.1) fixing the jump distance, d, effectively restricts each individual to a lattice defined by the initial starting point of the individual, with spacing d. The equivalent two-dimensional model described in Section 5.1 does not have the same drawback as the one-dimensional model, since movement can occur in any direction so that there is a route of fixed distance jumps between any two points in the domain. However, using a fixed distance sometimes requires individuals to move through a more tortuous route to move close to each other since the direct movement is prevented by the presence of the other individual. In addition, there are situations that we may wish to model (such as crowd dynamics; Helbing *et al.*, 2000; Tanimoto *et al.*, 2010; Weng *et al.*, 2006) in which individuals may move differing distances, perhaps dependent on the space available to move in. We will therefore allow each jump to be of a distance, d, that is drawn from a probability distribution. Using distributed distances also enables more direct comparison with previous derivations of volume exclusion in two dimensions (Bruna and Chapman, 2012) and allows comparison with the equations derived in one dimension (see Section 4.7).

As with the one-dimensional case, we use the probabilistic expansion

$$\text{Prob}(\text{A} = a) = \int_{b \in \text{B}} \text{Prob}(\text{A} = a | \text{B} = b)\ \text{Prob}(\text{B} = b)\ \text{d}b, \tag{5.21}$$

and integrate over the different values that the jump distance may take. In this case, we assume the distance moved to be positive with distribution $f(u_d)$ for $u_d > 0$, so that the range of movement angles is $\theta \in [-\pi, \pi)$, as in Section 5.2. If $\mathbf{d}_\theta^{u_d} = (u_d \cos\theta, u_d \sin\theta)$, then

$$\begin{aligned}\frac{\partial C_i}{\partial t} = \frac{\alpha}{2\pi} \int_0^\infty \underbrace{f(u_d)}_{\substack{\text{Prob(moving } u_d\text{)}}} \int_{-\pi}^{\pi} \overbrace{[C_i(\mathbf{x} - \mathbf{d}_\theta^{u_d}, t) - C_i(\mathbf{x}, t)]}^{\text{Prob(successfully moving given } u_d\text{)}}\ \text{d}\theta\ \text{d}u_d \\ + \frac{\alpha}{2\pi} \int_0^\infty \underbrace{f(u_d)}_{\substack{\text{Prob of moving } u_d}} \int_{-\pi}^{\pi} \overbrace{[P_i(\mathbf{x}, \theta, t) C_i(\mathbf{x}, t) - P_i(\mathbf{x} - \mathbf{d}_\theta^{u_d}, \theta, t) C_i(\mathbf{x} - \mathbf{d}_\theta^{u_d}, t)]}^{\text{Prob(unsuccessfully moving given } u_d\text{)}}\ \text{d}\theta\ \text{d}u_d.\end{aligned} \tag{5.22}$$

Note that this equation is identical to equation (5.3), but integrated over all possible jump distances,

u_d. To be able to Taylor expand in terms of u_d, as in Section 4.7.1, we require that the distribution $f(u_d)$ is suitably narrow, so that there exists $0 < \epsilon \ll 1$ such that splitting the integral in equation (5.22) results in an integral over $[0, \epsilon]$ that is an order of magnitude greater than that over (ϵ, ∞). Proceeding as in Laplace's method of integration, we neglect the integral from (ϵ, ∞) to give

$$\begin{aligned}\frac{\partial C_i}{\partial t} = \frac{\alpha}{2\pi}\int_0^{\epsilon} f(u_d)\int_{-\pi}^{\pi}\left[P_i(\mathbf{x},\theta,t)C_i(\mathbf{x},t) - P_i(\mathbf{x}-\mathbf{d}_\theta^{u_d},\theta,t)C_i(\mathbf{x}-\mathbf{d}_\theta^{u_d},t)\right]\,\mathrm{d}\theta\,\mathrm{d}u_d \\ + \frac{\alpha}{2\pi}\int_0^{\epsilon} f(u_d)\int_{-\pi}^{\pi}\left[C_i(\mathbf{x}-\mathbf{d}_\theta^{u_d},t) - C_i(\mathbf{x},t)\right]\,\mathrm{d}\theta\,\mathrm{d}u_d,\end{aligned} \tag{5.23}$$

where $\epsilon \ll 1$. Then we may treat the integrand in exactly the same way as in Section 5.2 with d replaced by u_d, so that

$$\frac{\partial C}{\partial t} = \int_0^{\epsilon}\left[\frac{\alpha u_d^2}{4}\nabla^2 C + \frac{\alpha u_d^4}{64}\nabla^2\left(\nabla^2 C\right) + \frac{N-1}{N}\alpha R u_d^2(R\pi - u_d)\nabla\cdot(C\nabla C) + \mathcal{O}\left((u_d+R)^5\right)\right]\,\mathrm{d}u_d. \tag{5.24}$$

Since the integral over $[0, \epsilon]$ is an order of magnitude greater than that from (ϵ, ∞), as specified above, re-extending the integral to infinity only adds in higher order error terms, so that

$$\frac{\partial C}{\partial t} = \frac{\alpha\langle d^2\rangle}{4}\nabla^2 C + \frac{\alpha\langle d^4\rangle}{64}\nabla^2\left(\nabla^2 C\right) + \frac{N-1}{N}\alpha R(R\pi\langle d^2\rangle - \langle d^3\rangle)\nabla\cdot(C\nabla C)\,. \tag{5.25}$$

For example, if we take the jump distance, d, to be drawn from the Rayleigh distribution with parameter σ then

$$f(u_d) = \frac{u_d}{\sigma^2}e^{-u_d^2/2\sigma^2}. \tag{5.26}$$

Thus if θ is uniformly distributed then $X = d\cos(\theta)$ and $Y = d\sin(\theta)$ are normally distributed with mean zero and variance σ^2. Now if $\sigma \ll \epsilon \ll 1$ then

$$\int_{\epsilon}^{\infty} u_d e^{-u_d^2/2\sigma^2}\int_{-\pi}^{\pi}\left[P_i(\mathbf{x},\theta,t)C_i(\mathbf{x},t) - P_i(\mathbf{x}-\mathbf{d}_\theta^{u_d},\theta,t)C_i(\mathbf{x}-\mathbf{d}_\theta^{u_d},t) + C_i(\mathbf{x}-\mathbf{d}_\theta^{u_d},t) - C_i(\mathbf{x},t)\right]\,\mathrm{d}\theta\,\mathrm{d}u_d, \tag{5.27}$$

is of order $\mathcal{O}\left(\epsilon\exp(-\epsilon^2/\sigma^2)\right)$ and since $\epsilon/\sigma \gg 1$ this term is exponentially small compared to the equivalent integral over $[0, \epsilon]$. Hence we may restrict the integral to Taylor expand and then re-extend the integral to obtain equation (5.25). Now $\int_0^{\infty} f(u_d)u_d^3\,\mathrm{d}u_d = 3\sigma^3\sqrt{\pi/2}$ and $\int_0^{\infty} f(u_d)u_d^2\,\mathrm{d}u_d = 2\sigma^2$,

so that

$$\frac{\partial C}{\partial t} = \frac{\alpha\sigma^2}{2}\nabla^2 C + \frac{N-1}{N}\alpha\sigma^2 R\left(2R\pi - 3\sigma\sqrt{\frac{\pi}{2}}\right)\nabla\cdot(C\nabla C)\,. \tag{5.28}$$

The derived equation may be directly compared with the equation derived in the literature by Bruna and Chapman (2012), who describe hard-core particle interactions, using the stochastic differential equation

$$d\mathbf{X}_i \equiv \sqrt{2}d\mathbf{B}_i + \mathbf{f}_i dt, \qquad 1 \leq i \leq N, \tag{5.29}$$

where $\mathbf{X}_i$ is the position of particle i and we take the external forces on the particles, $\mathbf{f}_i$, to be zero. The $\mathbf{B}_i$ are independent two-dimensional Brownian motions ($d\mathbf{B}_i \sim \mathcal{N}(0, dt)$) so that the position of particle i evolves according to $d\mathbf{X}_i \sim \mathcal{N}(0, 2dt)$. Bruna and Chapman (2012) derive an equation for the pdf, $p(\mathbf{x}_1, t)$, of the position, $\mathbf{x}_1$, of the first particle which is given in two dimensions with no external forces as

$$\frac{\partial p}{\partial t}(\mathbf{x}_1, t) = \nabla_{x_1}\cdot\left\{\nabla_{x_1}\left[p + \frac{\pi}{2}(N-1)\epsilon^2 p^2\right]\right\}, \tag{5.30}$$

where ∇_{x_1} is the derivative with respect to $\mathbf{x}_1$ and ϵ is the diameter of a particle so that $\epsilon = 2R$, in our terminology. To enable comparisons we take $C_1 = C/N$ in equation (5.28) and $\hat{\alpha} = \alpha\sigma^2/2 = 1$ constant as $\sigma = \sqrt{2dt} \to 0$. Then

$$\frac{\partial C}{\partial t} = \nabla\cdot\left(\left(1 + (N-1)\pi(2R)^2 C\right)\nabla C\right), \tag{5.31}$$

$$= \nabla\cdot\left(\nabla\left(C + (N-1)\frac{\pi}{2}(2R)^2 C^2\right)\right), \tag{5.32}$$

exactly as in equation (5.30). Thus our equation is consistent with that derived by Bruna and Chapman (2012), whilst our method of derivation allows the inclusion of other distributions of jump distances and different migratory mechanisms not included by them.

Using distributed jump distances does not significantly alter the form of the derived equation (equation (5.28)) from equation (5.12), and only changes the magnitude of the excluded volume term when compared with a fixed jump of size σ. Hence the magnitude of the term is in general larger than for fixed jump distances and volume exclusion decreases the diffusivity for a smaller value of σ than for the equivalent fixed distance model. The change in magnitude of the exclusion term is

the same as in the one-dimensional case (equation (4.58)), increasing the magnitude to $3\sigma^3\sqrt{\pi/2}$. It is to be expected that the change in magnitude would be the same in any number of dimensions, since the increased magnitude results from comparing d^3 to the third moment of the Rayleigh distribution used for the jump distances, which is the same in both cases. Consequently, varying R and σ results in similar surfaces to Figures 5.4(a), 5.4(c) and 5.4(e), whilst shifting the position of minimum exclusion effect (Figures 5.6(c) and 5.6(e)).

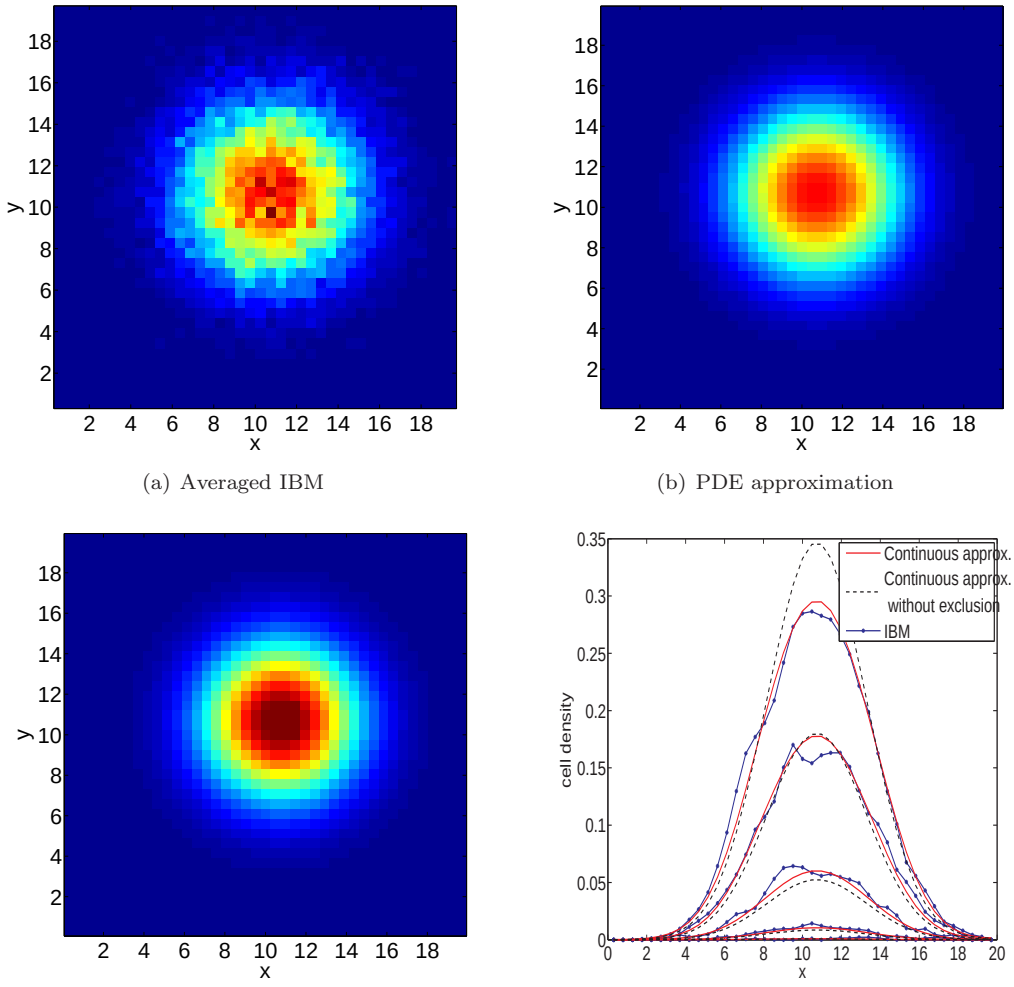
The surface profiles produced by varying R and N (Figures 5.6(b), 5.6(d) and 5.6(f)) are very similar to the equivalent fixed jump profiles (Figures 5.4(b), 5.4(d) and 5.4(f)). Interestingly, the approximation error in our derivation is not reduced by using a distributed jump distance, in contrast to the one-dimensional case (Section 4.7). The increased error in the one-dimensional case with fixed movement distances may be an artefact of the inability of cells to move arbitrarily close to one another in one dimension. Using a distributed jump distance therefore reduces the dependence of cell positions on their initial positions and thus may reduce the error due to our assumption of independence of cell positions. Conversely, all parts of the domain are accessible to cells in two dimensions even when using a fixed movement distance. We note that as $\sigma \to 0$ with $\hat{\alpha} = \lim_{\sigma\to 0} \alpha\sigma^2/2$ finite, equation (5.28) reduces to equation (5.13), as expected.

5.4 Domain growth

Domain growth is found in many biological systems, particularly during embryo development, which often necessitates large amounts of cell rearrangement and migration (Painter *et al.*, 2000; Dormann and Weijer, 2006; Zhang *et al.*, 2010). Growth is especially relevant during cell migration, since changes in the size and shape of the migratory domain impact on the speed of migration and dilutes cell concentrations. Since the system modelled in Chapters 2 and 3 involves uniform domain growth along one axis, we will consider uniform growth, defined by a given increase in the length and/or width of the domain. Let $B_L = B_B = 0$, $B_R = L_x(t)$ and $B_T = L_y(t)$, with $\mathbf{L}(t) = (L_x, L_y)(t)$ so that after a time step, $\Delta t \ll 1$, a coordinate $\mathbf{X}(t) = (X, Y)(t)$ will move to

$$\mathbf{X}(t+\Delta t) = \left(\frac{L_x(t+\Delta t)}{L_x(t)}, \frac{L_y(t+\Delta t)}{L_y(t)}\right) \cdot \mathbf{X}(t), \tag{5.33}$$

$$= \mathbf{X}(t) + \Delta t \left(\frac{L_x'}{L_x}, \frac{L_y'}{L_y}\right) \cdot \mathbf{X}(t). \tag{5.34}$$



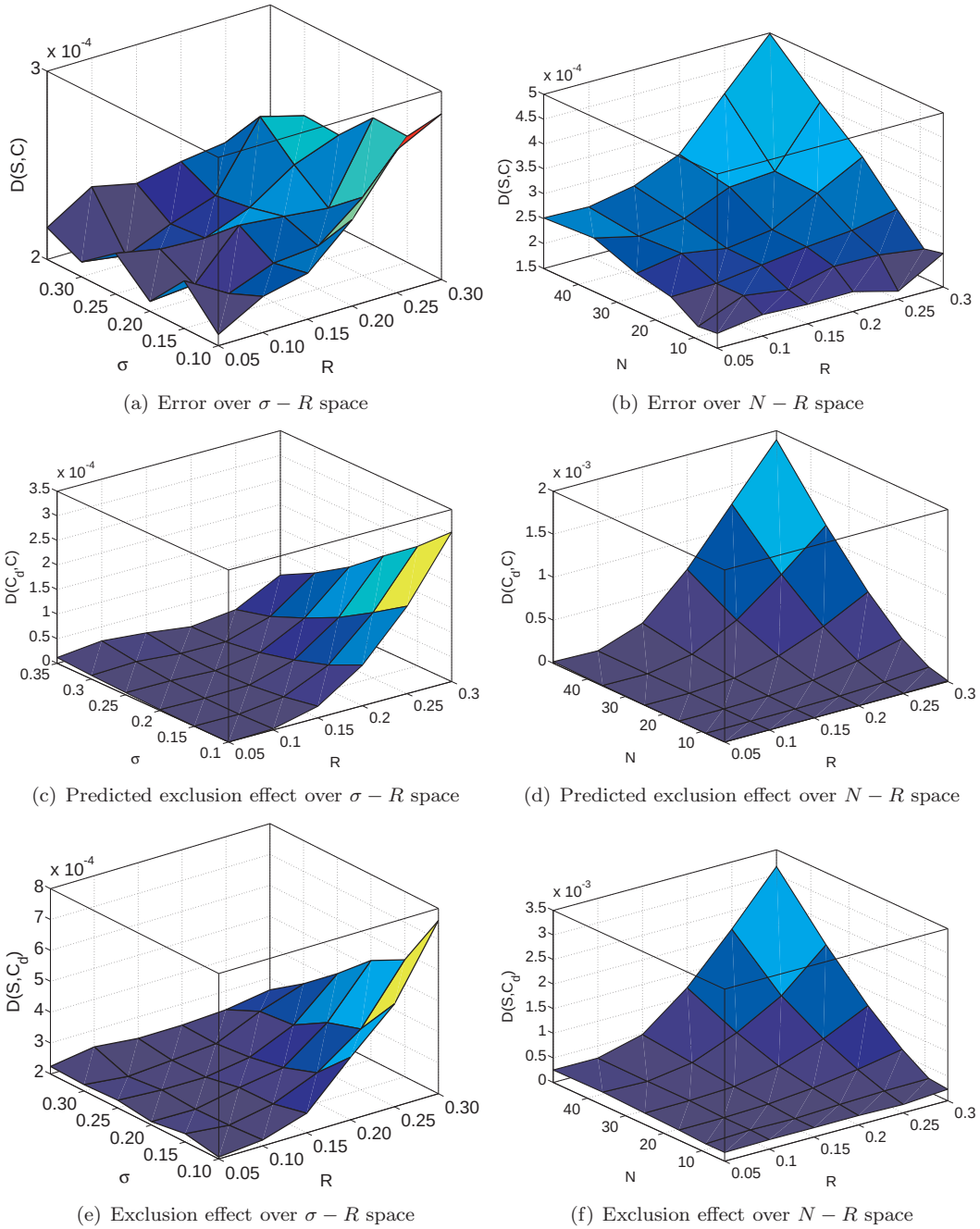
(a) Averaged IBM

(b) PDE approximation

(c) PDE approximation without exclusion

(d) y-axis slices

Figure 5.5: Comparison between the IBM with distributed movement distances and the continuum equation with or without volume exclusion (equation (5.28)). The distance moved is drawn from the Rayleigh distribution with parameter $\sigma = 0.1$, the radius of a cell is $R = 0.3$ and other parameter values are given in Table 5.1. Other details about the simulations are given in Section 5.2.3.



(a) Error over $\sigma - R$ space

(b) Error over $N - R$ space

(c) Predicted exclusion effect over $\sigma - R$ space

(d) Predicted exclusion effect over $N - R$ space

(e) Exclusion effect over $\sigma - R$ space

(f) Exclusion effect over $N - R$ space

Figure 5.6: Exploring the effect of changing the likelihood of successful movement in the model with distributed jump distances by changing the parameter, σ, of the Rayleigh jump distribution and the radius of a cell, R ((a), (c) and (e)) and changing the occupied space in the model by changing the number of cells, N, and the radius of a cell, R ((b), (d) and (f)). We consider the effect on: (a)-(b) the error introduced by our approximations (the difference between the solution to equation (5.28) and the average of our simulations); (c)-(d) our prediction of the effect of volume exclusion (the difference between the solution to our equation and the solution to the diffusion equation); and (e)-(f) the actual effect of volume exclusion (the difference between the average of our simulations and the solution to the diffusion equation). For more details, see Sections 4.6 and 5.2.5. Other parameter values are given in Table 5.1 with $\alpha = 0.05/\sigma^2$ so that $\hat{\alpha} = 0.025$ is constant as σ changes. Other details about the simulations are given in Section 5.2.3.

The probability that cell centre i is in the region $[x, x+\delta x) \times [y, y+\delta y)$ for $\delta x, \delta y \ll 1$ is

$$\int_x^{x+\delta x} \int_y^{y+\delta y} C_i(\mathbf{x}, t) \, \mathrm{d}x \, \mathrm{d}y \approx \delta x \delta y C_i(\mathbf{x}, t), \tag{5.35}$$

as before, but a patch of size $\delta x \times \delta y$ at time t grows to size $\delta x \left(1 + \Delta t L_x'/L_x\right) \times \delta y \left(1 + \Delta t L_y'/L_y\right)$ by $t + \Delta t$. Hence the probability that cell centre i is at time t in a region that grows to become $[x, x+\delta x) \times [y, y+\delta y)$ at time $t + \Delta t$ is

$$\delta x \frac{1}{1 + \frac{\Delta t L_x'}{L_x}} \times \delta y \frac{1}{1 + \frac{\Delta t L_y'}{L_y}} C_i(\mathbf{x}, t) \approx \delta x \left(1 - \frac{\Delta t L_x'}{L_x}\right) \times \delta y \left(1 - \frac{\Delta t L_y'}{L_y}\right) C_i(\mathbf{x}, t). \tag{5.36}$$

Note that taking the factor $\delta x/(1 + \Delta t L_x'/L_x) \times \delta y/(1 + \Delta t L_y'/L_y)$ ensures that the total integral of C_i over the whole domain, $\mathcal{D}(t)$, at time t remains constant as the domain increases in size, as required for a pdf. This can be seen since

$$\int_{\mathcal{D}(t+\Delta t)} C_i(x, y, t + \Delta t) \, \mathrm{d}y \, \mathrm{d}x = \int_{\mathcal{D}(t+\Delta t)} \frac{1}{1 + \frac{\Delta t L_x'}{L_x}} \frac{1}{1 + \frac{\Delta t L_y'}{L_y}} C_i \left(\frac{x}{1 + \frac{\Delta t L_x'}{L_x}}, \frac{y}{1 + \frac{\Delta t L_y'}{L_y}}, t \right) \, \mathrm{d}y \, \mathrm{d}x, \tag{5.37}$$

$$= \int_0^{L_x(t)} \int_0^{L_y(t)} C_i(X, Y, t) \, \mathrm{d}X \, \mathrm{d}Y, \tag{5.38}$$

by changing variables to

$$X = \frac{x}{1 + \frac{\Delta t L_x'}{L_x}}, \tag{5.39}$$

$$Y = \frac{y}{1 + \frac{\Delta t L_y'}{L_y}}. \tag{5.40}$$

Using equation (5.36),

$$\overbrace{\delta x \delta y}^{\text{new size of patch}} C_i(\mathbf{x}, t + \Delta t) = \overbrace{\delta x \delta y \left(1 - \frac{\Delta t L_x'}{L_x}\right) \left(1 - \frac{\Delta t L_y'}{L_y}\right)}^{\text{previous size of patch}} \Bigg[C_i(\mathbf{x}_L, t) + \frac{\alpha \Delta t}{2\pi} \int_{-\pi}^{\pi} \left[C_i(\mathbf{x}_L - \mathbf{d}_\theta^L, t) - C_i(\mathbf{x}_L, t) \right] \, \mathrm{d}\theta + \frac{\alpha \Delta t}{2\pi} \int_{-\pi}^{\pi} \left[P_i(\mathbf{x}_L, \theta, t) C_i(\mathbf{x}_L, t) - P_i(\mathbf{x}_L - \mathbf{d}_\theta^L, \theta, t) C_i(\mathbf{x}_L - \mathbf{d}_\theta^L, t) \right] \, \mathrm{d}\theta \Bigg] + \mathcal{O}(\Delta t^2), \tag{5.41}$$

where

$$\mathbf{x}_L = \left(\frac{x}{1 + \frac{\Delta t L_x'}{L_x}}, \frac{y}{1 + \frac{\Delta t L_y'}{L_y}} \right), \tag{5.42}$$

$$\mathbf{d}_\theta^L = \left(\frac{d}{1 + \frac{\Delta t L_x'}{L_x}} \cos\theta, \frac{d}{1 + \frac{\Delta t L_y'}{L_y}} \sin\theta \right). \tag{5.43}$$

Hence, expanding further in terms of Δt and dividing through by $\delta x \delta y \Delta t$

$$\begin{aligned} \frac{\partial C_i}{\partial t} = \frac{\alpha}{2\pi} \int_{-\pi}^{\pi} \left[C_i(\mathbf{x} - \mathbf{d}_\theta, t) - C_i(\mathbf{x}, t) \right] \, \mathrm{d}\theta - \overbrace{\left(\frac{\partial}{\partial x} \left(\frac{L_x' x}{L_x} C_i \right) + \frac{\partial}{\partial y} \left(\frac{L_y' y}{L_y} C_i \right) \right)}^{\text{terms due to domain growth}} \\ + \frac{\alpha}{2\pi} \int_{-\pi}^{\pi} \left[P_i(\mathbf{x}, \theta, t) C_i(\mathbf{x}_L, t) - P_i(\mathbf{x}_L - \mathbf{d}_\theta^L, \theta, t) C_i(\mathbf{x}_L - \mathbf{d}_\theta^L, t) \right] \, \mathrm{d}\theta, \end{aligned} \tag{5.44}$$

as $\Delta t \to 0$.

Thus $C = \sum_i C_i$ satisfies

$$\begin{aligned} \frac{\partial C}{\partial t} + \overbrace{\left(\frac{L_x'}{L_x} + \frac{L_y'}{L_y} \right) C}^{\text{dilution due to domain growth}} + \underbrace{\left(\frac{L_x' x}{L_x} \frac{\partial C}{\partial x} + \frac{L_y' y}{L_y} \frac{\partial C}{\partial y} \right)}_{\text{additional flux due to domain growth}} = \hat{\alpha} \nabla^2 C + \frac{N-1}{N} \alpha R d^2 (R\pi - d) \nabla \cdot (C \nabla C) \\ + \frac{\alpha d^4}{64} \nabla^2 \left(\nabla^2 C \right) + \mathcal{O}\left((d+R)^5 \right), \end{aligned} \tag{5.45}$$

To solve equation (5.45) numerically, we rescale onto a fixed domain with $X = x/L_x$, $Y = y/L_y$ and $T = t$, so that $\mathrm{d}C/\mathrm{d}t = \mathrm{d}C/\mathrm{d}T + (\mathrm{d}C/\mathrm{d}X)(XL_x'/L_x) + (\mathrm{d}C/\mathrm{d}Y)(YL_y'/L_y)$. Then as $d \to 0$ with $\hat{\alpha} = \lim_{d \to 0} \alpha d^2/4$ constant:

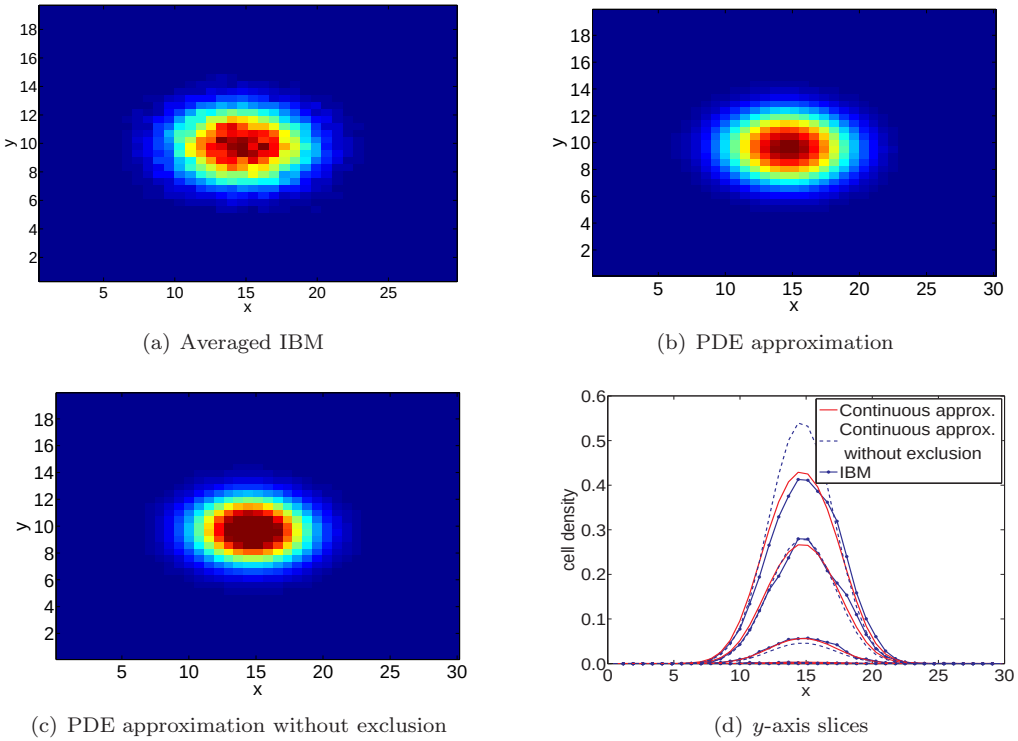
$$\frac{\partial C}{\partial T} + \overbrace{\left(\frac{L_x'}{L_x} + \frac{L_y'}{L_y} \right) C}^{\text{dilution}} = \hat{\alpha} \nabla_L \cdot \left(\left(1 + \frac{N-1}{N} 4R^2 C \right) \nabla_L C \right), \tag{5.46}$$

where

$$\nabla_L = \left(\frac{1}{L_x} \frac{\partial}{\partial X}, \frac{1}{L_y} \frac{\partial}{\partial Y} \right), \tag{5.47}$$

$$\nabla_L^2 = \nabla_L \cdot \nabla_L = \left(\frac{1}{L_x^2} \frac{\partial^2}{\partial X^2} + \frac{1}{L_y^2} \frac{\partial^2}{\partial Y^2} \right), \tag{5.48}$$

and the additional flux due to domain growth has been removed by rescaling onto a fixed domain.



(a) Averaged IBM

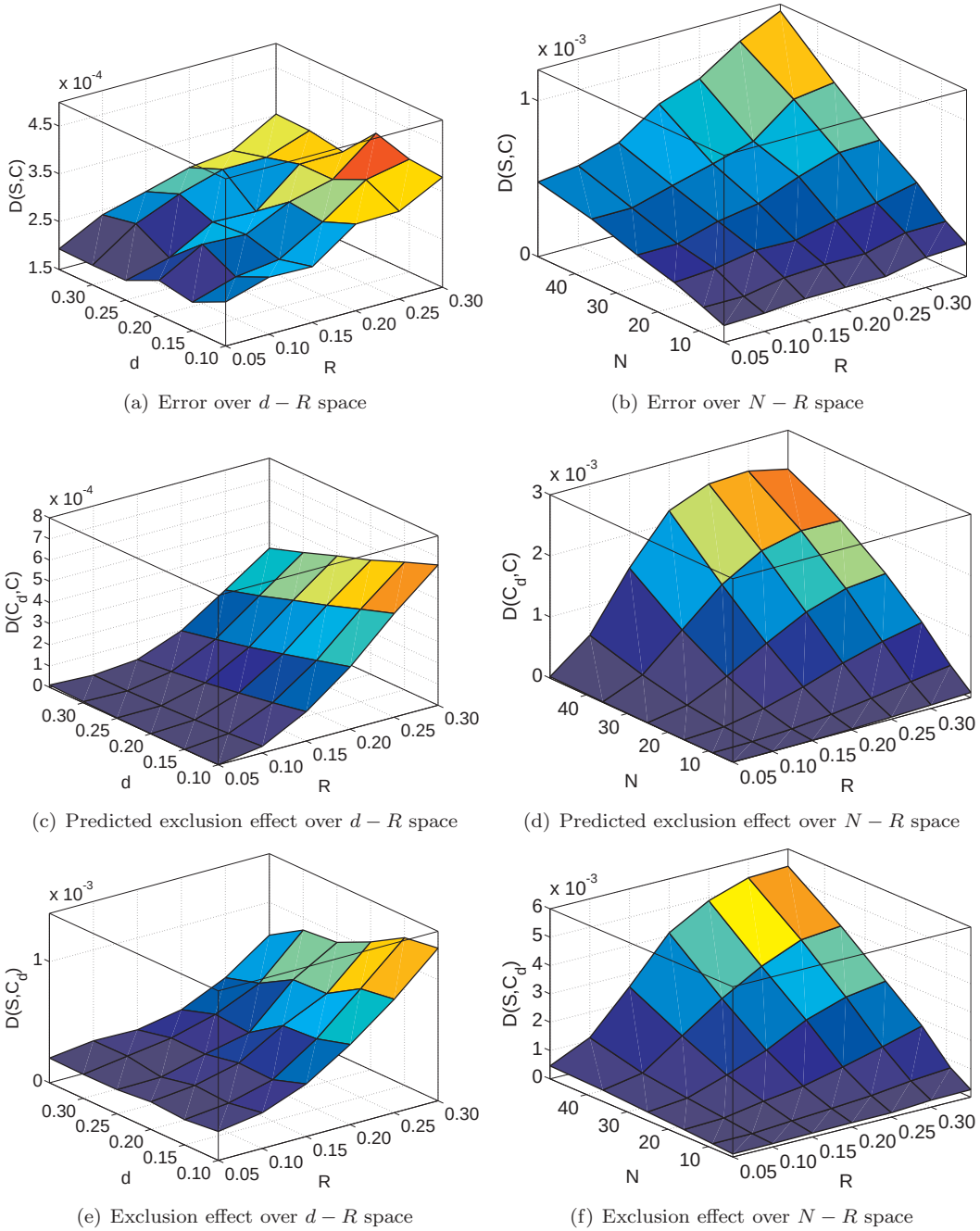
(b) PDE approximation

(c) PDE approximation without exclusion

(d) y-axis slices

Figure 5.7: Comparison between the IBM with domain growth along the x-axis and the continuum equation with or without volume exclusion (equation (5.48)). The domain grows parallel to the x-axis at a rate $L_a = 0.1$, the distance moved is $d = 0.3$, the radius of a cell is $R = 0.3$ and other parameter values are given in Table 5.1. Other details about the simulations are given in Section 5.2.3.

The differences between equation (5.48) and equation (5.13) are exactly analogous to the one-dimensional case (Section 5.4), comprising of a dilution term and an additional flux term that is removed by rescaling onto a fixed domain. Simulations of the IBM on a domain that expands parallel to the x-axis (so that $L_y' = 0$) compare well with numerical solutions to equation (5.48) (Figure 5.7). As in one dimension, the expanding domain does not significantly qualitatively or quantitatively change the exclusion effect surface as either R and d (Figures 5.8(a), 5.8(c) and 5.8(e)) or R and N (Figures 5.8(b), 5.8(d) and 5.8(f)) change. The volume exclusion effects are very slightly decreased by the dilution caused by the expansion of the domain (Figures 5.8(e) and 5.8(f)) whilst the error in our predicted profiles is slightly increased (Figures 5.8(a) and 5.8(b)).



(a) Error over $d - R$ space

(b) Error over $N - R$ space

(c) Predicted exclusion effect over $d - R$ space

(d) Predicted exclusion effect over $N - R$ space

(e) Exclusion effect over $d - R$ space

(f) Exclusion effect over $N - R$ space

Figure 5.8: Exploring the effect of changing the likelihood of successful movement in the model with domain growth by changing the distance moved, d, and the radius of a cell, R ((a), (c) and (e)) and changing the occupied space in the model by changing the number of cells, N, and the radius of a cell, R ((b), (d) and (f)). We consider the effect on: (a)-(b) the error introduced by our approximations (the difference between the solution to equation (5.48) and the average of our simulations); (c)-(d) our prediction of the effect of volume exclusion (the difference between the solution to our equation and the solution to the diffusion equation); and (e)-(f) the actual effect of volume exclusion (the difference between the average of our simulations and the solution to the diffusion equation). For more details, see Sections 4.6 and 5.4. Other parameter values are given in Table 5.1 with $\alpha = 0.05/d^2$ so that $\hat{\alpha} = 0.0125$ is constant as d changes. Other details about the simulations are given in Section 5.2.3.

5.5 Heterogeneous populations

It is straightforward to extend our method to consider multiple populations of cells, which may have different properties. This enables study of simple heterogeneous populations such as proteins in signalling pathways, predator-prey systems and the immune response. For example, we could consider the situation where one subpopulation of cells are large and slow moving, whilst the other subpopulation is small but quick. If C_j^i is the pdf for the ith member of population j, which consists of N_j cells of radius R_j, moving a distance d_j at a rate α_j, then

$$C_j^i(\mathbf{x}, t+\Delta t) = C_j^i(\mathbf{x},t)\Big(\overbrace{1-\alpha_j\Delta t}^{\text{does not try to move}} + \underbrace{\frac{\alpha_j\Delta t}{2\pi}\int_0^{2\pi} P_j^i(\mathbf{x},\theta,t)\ \mathrm{d}\theta}_{\text{movement prevented}}\Big) + \underbrace{\frac{\alpha_j\Delta t}{2\pi}\int_0^{2\pi}\left(1-P_j^i(\mathbf{x}-\mathbf{d}_\theta,\theta,t)\right)C_j^i(\mathbf{x}-\mathbf{d}_\theta,t)\ \mathrm{d}\theta}_{\text{successful moves to }\mathbf{x}}, \tag{5.49}$$

where

$$P_j^i(\mathbf{x},\theta,t) = \overbrace{\sum_{l\neq i, l=1}^{N_j}\int_0^{d_j}\int_{-\pi/2}^{\pi/2} 2R_j\cos\phi\ C_j^l(\mathbf{x}+\mathbf{r}_\theta(2R_j),t)\ \mathrm{d}\phi\ \mathrm{d}r}^{\text{movement prevented by other cells in population } j} + \overbrace{\sum_{k\neq j}\sum_{l=1}^{N_k}\int_0^{d_j}\int_{-\pi/2}^{\pi/2}(R_j+R_k)\cos\phi\ C_k^l(\mathbf{x}+\mathbf{r}_\theta(R_j+R_k),t)\ \mathrm{d}\phi\ \mathrm{d}r}^{\text{movement prevented by cells in other populations}}, \tag{5.50}$$

and $\mathbf{r}_\theta(R) = (r\cos\theta + R\cos(\theta+\phi), r\sin\theta + R\sin(\theta+\phi))$. Hence, expanding in terms of d_j and R_j $\forall j$,

$$\frac{\partial C_j^i}{\partial t} = \frac{\alpha_j d_j^2}{4}\nabla^2 C_j^i + \alpha_j R_j d_j^2(R_j\pi - d_j)\nabla\cdot\left(C_j^i\sum_{l\neq i}\nabla C_j^l\right) + \sum_{k\neq j}\frac{\alpha_j\pi}{4}d_j^2(R_j+R_k)^2\nabla\cdot\left(C_j^i\sum_{l=1}^{N_k}\nabla C_k^l\right) - N_k\frac{\alpha_j}{2}d_j^3(R_j+R_k)\nabla\cdot\left(\sum_{l=1}^{N_k}C_k^l\nabla C_j^i\right). \tag{5.51}$$

Taking $d_j\to 0$ with $\hat{\alpha}_j = \lim_{d_j\to 0}\alpha d_j^2/4$, then summing over all i, we obtain

$$\frac{\partial C_j}{\partial t} = \hat{\alpha}_j\nabla\cdot\left(\left(1+4\frac{N_j-1}{N_j}R_j(R_j\pi-d_j)C_j\right)\nabla C_j\right) + \hat{\alpha}_j\sum_{k\neq j}\pi(R_j+R_k)^2\nabla\cdot(C_j\nabla C_k), \tag{5.52}$$

where $C_j = \sum_{i=1}^{N_j} C_j^i$ and $C_k = \sum_{i=1}^{N_k} C_k^i$.

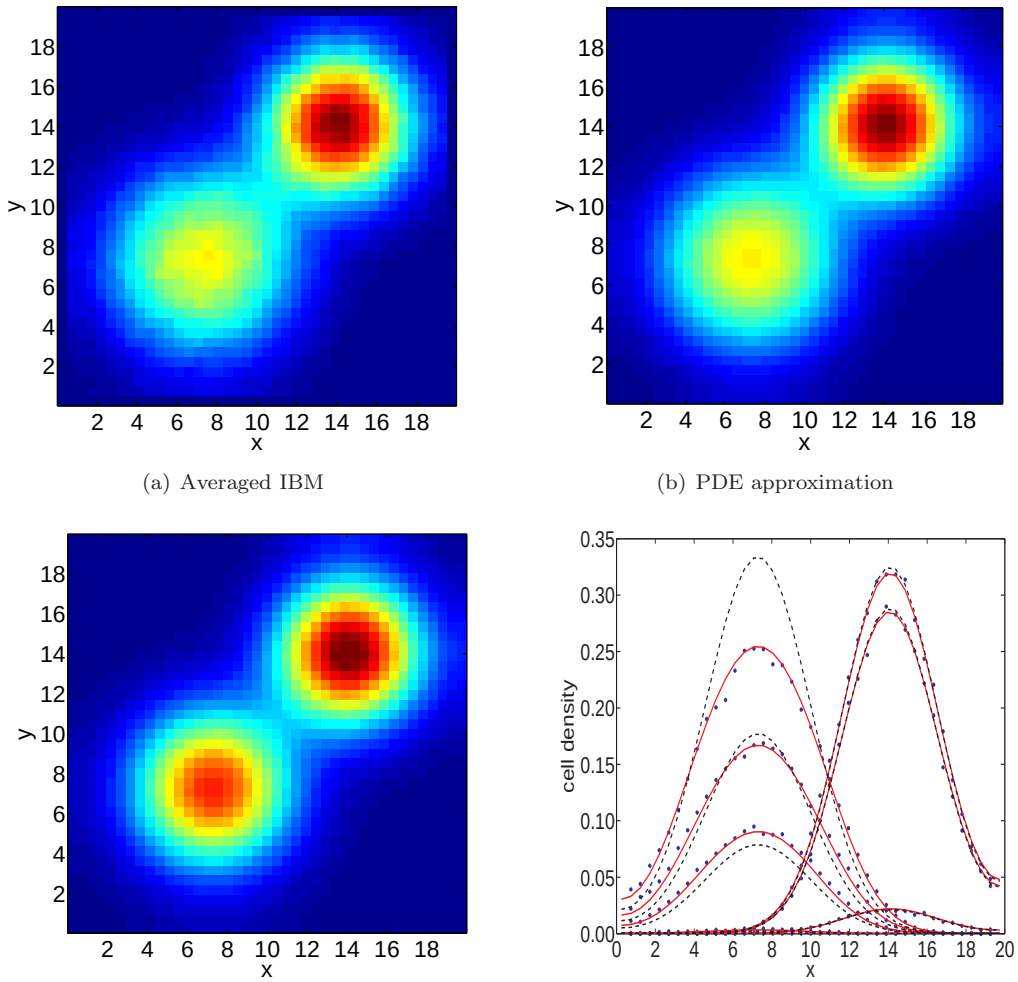
The addition of other cell populations leads to a chemotaxis-type term of individuals of population j, moving down concentration gradients of other populations. For two populations equation (5.52) is

$$\frac{\partial C_1}{\partial t} = \hat{\alpha}_1 \nabla \cdot \left(\left(1 + 4\frac{N_1 - 1}{N_1} R_1(R_1\pi - d_1)C_1 \right) \nabla C_1 \right) + \hat{\alpha}_1 \pi (R_1 + R_2)^2 \nabla \cdot (C_1 \nabla C_2)\,, \tag{5.53}$$

$$\frac{\partial C_2}{\partial t} = \hat{\alpha}_2 \nabla \cdot \left(\left(1 + 4\frac{N_2 - 1}{N_2} R_2(R_2\pi - d_2)C_2 \right) \nabla C_2 \right) + \hat{\alpha}_2 \pi (R_1 + R_2)^2 \nabla \cdot (C_2 \nabla C_1)\,. \tag{5.54}$$

Note that the volume exclusion in equation (5.53) due to individuals of population 2 only occurs through the spatial heterogeneity of C_2. Hence if population 2 is homogeneously distributed throughout the domain then the dispersal of C_1 is unaffected. This is because volume exclusion that is represented by C_2 directly is at higher order and thus is negligible as $d \to 0$. For example, if C_2 is spatially homogeneous, then the diffusion of C_1 would be reduced by $(R_1 + R_2)dC_2$, since we integrate over a region of that size.

Averaged simulations of a population of larger individuals (at the bottom left) and a population of smaller individuals (at the top right) are shown in Figure 5.9(a). We expect that the population of smaller individuals will be less affected by volume exclusion, both within their population (since less space is taken up by individuals), and between the two populations (since each smaller individual requires less space to move into), than the population of larger individuals (Figure 5.9). Conversely, whilst we expect volume exclusion to have a large effect on population dynamics between the larger individuals, we expect there to be less volume exclusion effects where the two populations meet, since an area that is mainly populated by smaller individuals will have more free space and thus have fewer interactions between cells. Hence the population of larger individuals is very different with volume exclusion in regions where there are no other individuals, but volume exclusion is less important at the overlap between the two populations (Figure 5.9(d)). Although the volume exclusion effect at the interface between populations is not significant, our equations do capture the interface well (Figure 5.9(d)) for a range of parameter combinations and initial conditions (results not shown), thus demonstrating the potential for examining the interactions of different populations of individuals. This is an area that could be helpful in the study of diffusion in crowded environments (anomalous diffusion) to find good representations of small, quick individuals moving amongst a large, slow population.



(a) Averaged IBM

(b) PDE approximation

(c) PDE approximation without exclusion

(d) y-axis slices

Figure 5.9: Comparison between the IBM with two populations of cells and the continuum equation with or without volume exclusion (equation (5.51)). Sixteen cells of each population were placed in square groups in different parts of the domain. The centre coordinates of the bottom left-most cell of population one are drawn from two normal distributions: $x \sim \mathcal{N}(20/3 - \sqrt{2RN}/0.8, 0.1)$ and $y \sim \mathcal{N}(10 - \sqrt{2RN}/0.8, 0.1)$, whilst the centre coordinates of the bottom left-most cell of population two are drawn from two normal distributions: $x \sim \mathcal{N}(40/3 - \sqrt{2RN}/0.8, 0.1)$ and $y \sim \mathcal{N}(10 - \sqrt{2RN}/0.8, 0.1)$. Subfigure (d) shows slices through the y-axis at the same time point, with population 1 on the left and population 2 on the right. Blue dots are averaged IBM simulations, black dashed lines are the continuous approximation without exclusion and red lines are the continuous approximation with exclusion. The distance moved is $d_1 = d_2 = 0.15$, the radius of a cell is $R_1 = 0.4$, $R_2 = 0.1$ and other parameter values are given in Table 5.1. Other details about the simulations are given in Section 5.2.3.

5.6 Biased movement

As in Section 4.9, we will make a simple first attempt at investigating chemotaxis by including a simple bias in movement rates, leaving the analysis of a fully coupled chemotactic system to Chapter 6. In two dimensions, we may choose a direction, ϕ, for a global bias, so that cells move in a direction specified by a randomly chosen angle $\theta \in [\phi - \pi/2, \phi + \pi/2]$ with rate $\alpha/2 + \epsilon$, and, with rate $\alpha/2 - \epsilon$ in a random direction $\theta \in [\phi + \pi/2, \phi + 3\pi/2]$.

5.6.1 Derivation of continuum equations

To derive continuum equations for the globally biased system, we split the integral of θ into two halves: towards ϕ; and away from ϕ:

$$\begin{aligned} C_i(\mathbf{x}, t + \Delta t) = & \\ & C_i(\mathbf{x}, t) \left(1 - \alpha \Delta t + \left(\frac{\alpha}{2} + \epsilon \right) \frac{\Delta t}{\pi} \int_{\phi - \pi/2}^{\phi + \pi/2} P_i(\mathbf{x}, \theta, t) \, \mathrm{d}\theta + \left(\frac{\alpha}{2} - \epsilon \right) \frac{\Delta t}{\pi} \int_{\phi + \pi/2}^{\phi + 3\pi/2} P_i(\mathbf{x}, \theta, t) \, \mathrm{d}\theta \right) \\ & + \left(\frac{\alpha}{2} + \epsilon \right) \frac{\Delta t}{\pi} \int_{\phi - \pi/2}^{\phi + \pi/2} \left(1 - P_i(\mathbf{x} - \mathbf{d}_\theta, \theta, t) \right) C_i(\mathbf{x} - \mathbf{d}_\theta, t) \, \mathrm{d}\theta \\ & + \left(\frac{\alpha}{2} - \epsilon \right) \frac{\Delta t}{\pi} \int_{\phi + \pi/2}^{\phi + 3\pi/2} \left(1 - P_i(\mathbf{x} - \mathbf{d}_\theta, \theta, t) \right) C_i(\mathbf{x} - \mathbf{d}_\theta, t) \, \mathrm{d}\theta + \mathcal{O}\left(\Delta t^2 \right), \end{aligned} \tag{5.55}$$

and, letting $\Delta t \to 0$,

$$\begin{aligned} \frac{\partial C_i}{\partial t} = & \frac{\alpha}{2\pi} \int_{-\pi}^{\pi} \left[C_i(\mathbf{x} - \mathbf{d}_\theta, t) - C_i(\mathbf{x}, t) \right] \, \mathrm{d}\theta \\ & + \frac{\alpha}{2\pi} \int_{-\pi}^{\pi} \left[P_i(\mathbf{x}, \theta, t) C_i(\mathbf{x}, t) - P_i(\mathbf{x} - \mathbf{d}_\theta, \theta, t) C_i(\mathbf{x} - \mathbf{d}_\theta, t) \right] \, \mathrm{d}\theta \\ & + \frac{\epsilon}{\pi} \left(\int_{\phi - \pi/2}^{\phi + \pi/2} - \int_{\phi + \pi/2}^{\phi + 3\pi/2} \right) \left[C_i(\mathbf{x} - \mathbf{d}_\theta, t) - C_i(\mathbf{x}, t) \right] \, \mathrm{d}\theta \\ & + \frac{\epsilon}{\pi} \left(\int_{\phi - \pi/2}^{\phi + \pi/2} - \int_{\phi + \pi/2}^{\phi + 3\pi/2} \right) \left[P_i(\mathbf{x}, \theta, t) C_i(\mathbf{x}, t) - P_i(\mathbf{x} - \mathbf{d}_\theta, \theta, t) C_i(\mathbf{x} - \mathbf{d}_\theta, t) \right] \, \mathrm{d}\theta. \end{aligned} \tag{5.56}$$

Hence $C = \sum_i C_i$ satisfies

$$\begin{aligned} \frac{\partial C}{\partial t} = & \frac{\alpha d^2}{4} \nabla^2 C - \frac{4 \epsilon d}{\pi} \left(1 - 8 d R \frac{N - 1}{N} C \right) \left(\cos\phi \frac{\partial C}{\partial x} + \sin\phi \frac{\partial C}{\partial y} \right) \\ & + \frac{N - 1}{N} \alpha R d^2 (R\pi - d) \nabla \cdot (C \nabla C) + \mathcal{O}\left((d + R)^5 \right), \end{aligned} \tag{5.57}$$

and taking $d \to 0$ with $\hat{\alpha} = \lim_{d\to 0} \alpha d^2/4$ and $\hat{\epsilon} = \lim_{d\to 0} 4\epsilon d/\pi$ finite, we obtain

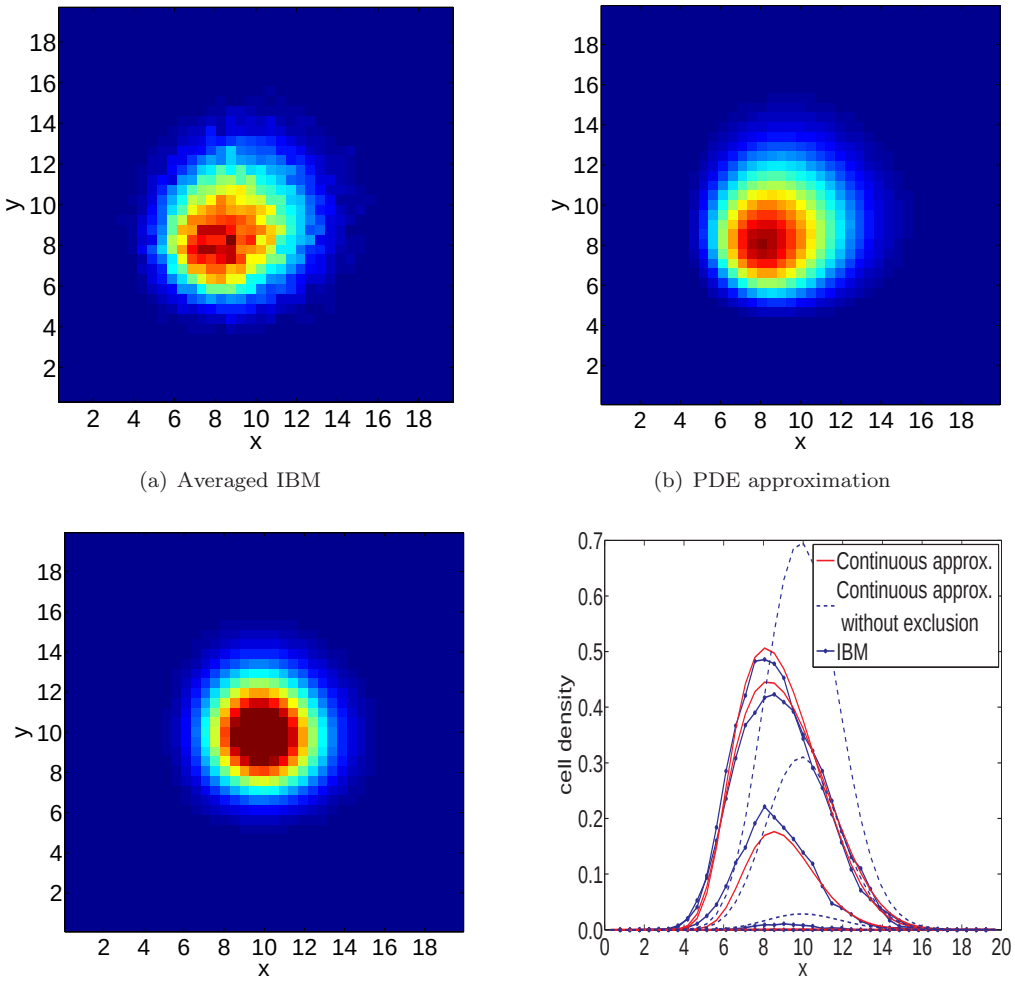
$$\frac{\partial C}{\partial t} = \hat{\alpha}\nabla \cdot \left(\left(1 + \frac{N-1}{N} 4R^2 C \right) \nabla C \right) - \hat{\epsilon} \left(\cos\phi \frac{\partial C}{\partial x} + \sin\phi \frac{\partial C}{\partial y} \right). \tag{5.58}$$

The inclusion of a global bias thus introduces a further term to our equations, corresponding to a flux in the direction of bias, as was also found in one dimension (equation (4.74)).

Simulations of individuals initialised in one group with biased movement compare well with numerical solutions to equation (5.58) (Figure 5.10). The pdf profile evolves as a two-dimensional wavepulse, moving in the direction of bias and spreading out into a less tightly packed group of individuals as time increases. As in one dimension, the addition of biased movement increases the effect of volume exclusion, since movement is severely limited at the back of the cell group. The prevention of some cell jumps also reduces the speed of movement of the group of cells so that both the speed and profile of the pdf are not accurately predicted by the continuum equation without volume exclusion (Figure 5.10). The discrepancy in speed and profile predictions without volume exclusion for biased movement (Figures 5.11(a), 5.11(c) and 5.11(e)) lead to exclusion effects of up to 2.86 times those without biased movement (Figures 5.4(a), 5.4(c) and 5.4(e)), measured as the distance between averaged simulations and the continuum equation derived without volume exclusion. Greater jump distances, d, and cell radii, R, further increase the chance of individuals at the back of the group encountering those ahead, and hence increase the overall effect of volume exclusion (Figures 5.11(c) and 5.11(e)). Similarly, larger cell populations have even slower group travelling speeds, since trailing cells are more likely to encounter those ahead and thus display greater volume exclusion effects (Figures 5.11(d) and 5.11(f)).

5.7 Discussion

In this chapter we have extended our technique from Chapter 4 into two spatial dimensions, finding the probability of volume exclusion by integrating over a suitable area close to the moving cell. To reduce this to a local PDE model, we assume that the radius of a cell is small and take the limit as the distance moved during a time step tends towards zero so that local values of the concentration and its derivatives are enough to predict cell population development over time. We find similarities between the equations derived in one and two dimensions, and the addition of distributed jump distances and domain growth are exactly analogous to one-dimensional derivations and equations.



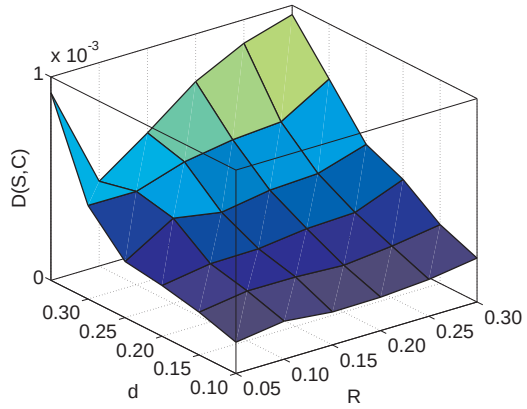
(a) Averaged IBM

(b) PDE approximation

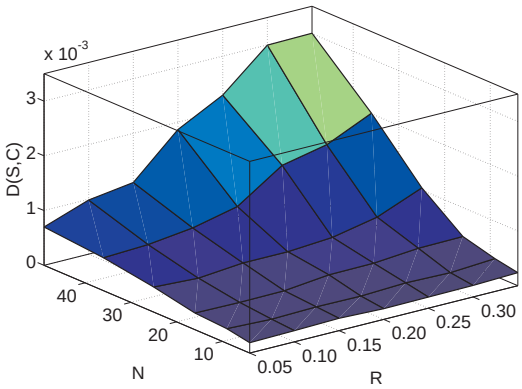
(c) PDE approximation without exclusion

(d) y-axis slices

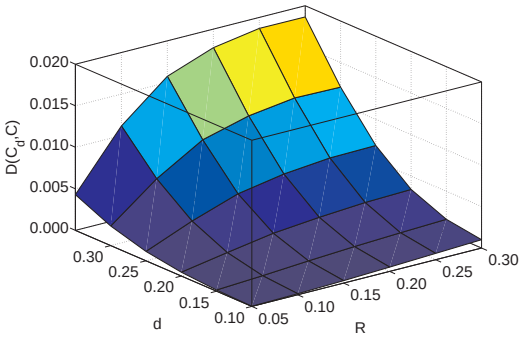
Figure 5.10: Comparison between the IBM with biased movement and the continuum equation with or without volume exclusion (equation (5.57)). Movement is biased with $\epsilon = 0.2$ towards $\phi = \pi/4$, distance moved is taken as $d = 0.3$, the radius of a cell is $R = 0.3$ and other parameter values are given in Table 5.1. To allow room to migrate in the direction of bias, the centre coordinates of the bottom left-most cell are drawn from the normal distribution, $\mathcal{N}(5 - \sqrt{2RN}/0.8, 1)$ and other cells are initialised in a grid as described, with other details about the simulations in Section 5.2.3.



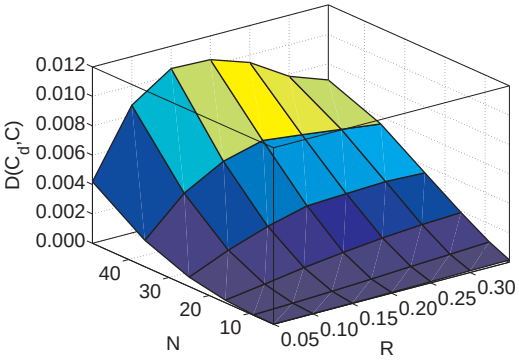
(a) Error over $d - R$ space



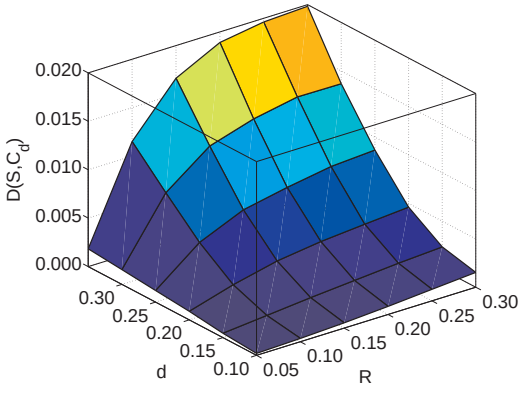
(b) Error over $N - R$ space



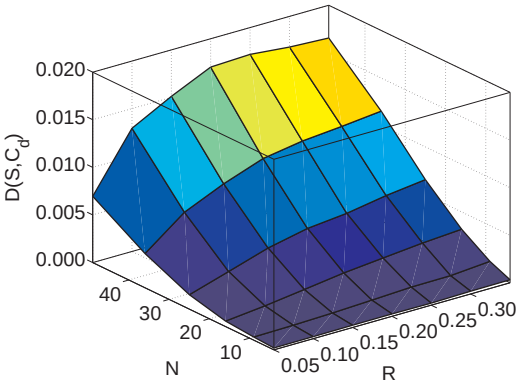
(c) Predicted exclusion effect over $d - R$ space



(d) Predicted exclusion effect over $N - R$ space



(e) Exclusion effect over $d - R$ space



(f) Exclusion effect over $N - R$ space

Figure 5.11: Exploring the effect of changing the likelihood of successful movement in the model with biased movement by changing the distance moved, d, and the radius of a cell, R ((a), (c) and (e)) and changing the occupied space in the model by changing the number of cells, N, and the radius of a cell, R ((b), (d) and (f)). We consider the effect on: (a)-(b) the error introduced by our approximations (the difference between the solution to equation (5.57) and the average of our simulations); (c)-(d) our prediction of the effect of volume exclusion (the difference between the solution to our equation and the solution to the diffusion equation); and (e)-(f) the actual effect of volume exclusion (the difference between the average of our simulations and the solution to the diffusion equation). For more details, see Sections 4.6 and 5.6. Movement is biased with $\epsilon = 0.2$ towards $\phi = \pi/4$ and other parameter values are given in Table 5.1. To allow room to migrate in the direction of bias, the centre coordinates of the bottom left-most cell are drawn from the normal distribution, $\mathcal{N}(5 - \sqrt{2RN}/0.8, 1)$ and other cells are initialised in a grid as described, with other details about the simulations in Section 5.2.3.

As in one dimension, the inclusion of a global movement bias changes both the profile of the cell density and the speed with which a group of individuals migrate across the domain. Despite the two-dimensional setting providing more unoccupied volume in which to move, biased movement results in a large number of interactions between individuals so that it is very important to consider volume exclusion when deriving continuum equations. We also studied heterogeneous populations, where we may define several subpopulations with different characteristics. This was not considered in a one-dimensional setting since the inability of cells to migrate past each other means that there would be hardly any interaction between different subpopulations. In a two-dimensional setting, the interface between the subpopulations was captured well by our derived equations and this is a promising avenue for studying heterogeneous and crowded environments.

It is expected that a hybrid chemotactic system in which individuals deplete chemoattractant close to them will result in a movement bias away from areas that have been occupied for longer periods of time (as in Chapters 2 and 3). Chemotactic systems of this type have been widely studied in a continuum setting (Keller and Segel, 1971; Hillen and Painter, 2009; Nicolau, 2008) but the stochastic nature of our IBMs may result in new and interesting behaviours. Any kind of widespread bias would be likely to increase interactions between individuals travelling in a group. The bias introduced by the inclusion of chemotaxis may therefore also result in significant effects from volume exclusion leading us to consider chemotactic systems in the next chapter.

Since biological systems that rely upon chemotaxis often rely upon relatively high concentrations of chemoattractant guiding relatively few individuals (Dallon and Othmer, 1997; Guo *et al.*, 2008), such as the system studied in Chapters 2 and 3, we will consider hybrid systems in which discrete individuals are guided by a continuum chemoattractant. Hybrid systems are studied in a variety of systems, with and without chemotaxis, and can include strong coupling between the two systems. It is important, therefore, to find methods of analysis that are successful at predicting averaged profiles in the presence of such coupling between both species. In Chapter 6 we will extend our analytical technique to chemotactic systems, particularly those in which chemotaxis is modelled as a hybrid system. In addition, we will investigate the coupling between the discrete and continuous systems and discuss how to find hybrid formulations of given discrete systems.

Chapter 6

Chemotaxis

Chapters 4 and 5 have focused on deriving continuum equations from one- and two-dimensional IBMs with volume exclusion. In this chapter we consider coupled chemotactic systems, where the motion of cells is influenced by the concentration of a chemical which is, in turn, affected by the presence of cells. We then use the techniques from Chapters 4 and 5 to consider volume exclusion in one- and two-dimensional chemotactic systems in preparation for deriving a system of equations in Chapter 7 from the chemotactic models in Chapters 2 and 3.

6.1 Model 1: on-lattice cells and chemoattractant

Since coupled chemotaxis systems are not well-studied in IBMs in either an on-lattice or off-lattice setting, we begin by considering a one-dimensional on-lattice model with discrete cells and chemical. We take a similar approach to Baker *et al.* (2010, Appendix A.3) who investigate a model consisting of a signalling molecule that is produced by cells and decays linearly. However, since we wish to later extend our method to an off-lattice setting, where they find the probability distribution of the number of cells at a lattice point, we instead consider the positions, $\mathbf{c} = (c_1, \cdots, c_N)$, of the N cells. In addition, we take $\mathbf{s} = (s_1, \cdots, s_L)$ chemical molecules initially present at each lattice point that are consumed by cells at that lattice site with a rate λ_s. Movement of the chemical is incorporated at a rate d_s to the left or right. Cell movement may depend on the distribution of chemical molecules, with a cell at lattice site i moving at a rate $T_i^-(\mathbf{s})$ and $T_i^+(\mathbf{s})$ to the left and right, respectively. Whilst this formulation allows for cell movement to depend on the chemical

concentration far from the position of the cell, in this thesis we consider only movement dependent on chemical concentrations close to the cell. In addition to the migration of cells and chemical, there is a flux, P_c, of cells into the first lattice point. Cells and chemical molecule movements that would leave the domain are prevented so that there is no flux of cells or chemical out of the domain.

6.1.1 Derivation of continuum equations

We begin by deriving the evolution of the probability, $C_k(i,t)$, that cell k will be at lattice point i at time t, using an analogous technique to that in Chapters 4 and 5. We note that the probability of cell k moving right (for example) from lattice point i at time t is given by

$$\text{Prob(cell } k \text{ moving right from site } i) = \sum_{s_1} \cdots \sum_{s_L} \text{Prob(moving right given } \mathbf{s}) \text{Prob}(\mathbf{s}), \tag{6.1}$$

$$= \sum_{\mathbf{s}} T_i^+(\mathbf{s}) P(\mathbf{s},t), \tag{6.2}$$

where $P(\mathbf{s},t)$ is the probability that $\mathbf{s} = (s_1, \cdots, s_L)$ chemical molecules are present at time t, and $\sum_{\mathbf{s}} = \sum_{s_1} \cdots \sum_{s_L}$. Hence

$$C_k(i,t+\Delta t) = \overbrace{C_k(i,t)\left(1 - \Delta t \sum_{\mathbf{s}} \left(T_i^+(\mathbf{s})P(\mathbf{s},t) + T_i^-(\mathbf{s})P(\mathbf{s},t)\right)\right)}^{\text{cell } k \text{ at } i \text{ stays at } i}$$
$$+ \underbrace{\Delta t \sum_{\mathbf{s}} T_{i-1}^+(\mathbf{s})P(\mathbf{s},t)C_k(i-1,t)}_{\text{cell } k \text{ at } i-1 \text{ moves to } i} + \underbrace{\Delta t \sum_{\mathbf{s}} T_{i+1}^-(\mathbf{s})P(\mathbf{s},t)C_k(i+1,t)}_{\text{cell } k \text{ at } i+1 \text{ moves to } i}, \tag{6.3}$$

where we make the moment closure assumption that the chemical profile is approximately uncorrelated with the position of cell k when averaged over multiple simulations. That is

$$\sum_{\mathbf{s}} \sum_{l} T_i^{\pm}(\mathbf{s}) \text{Prob}(C_k^l(i,t), \mathbf{s}(t)) \approx \langle C_k^l(i,t) \rangle \langle T_i^{\pm}(\mathbf{s}) \rangle, \tag{6.4}$$

$$= C_k(i,t) \sum_{\mathbf{s}} T_i^{\pm}(\mathbf{s}) P(\mathbf{s},t), \tag{6.5}$$

where $C_k(i,t)$ can be thought of as the average position of cell k over all simulations, so that $\text{Prob}(C_k^l(i,t), \mathbf{s}(t))$ is the probability that cell k is at position i at time t during simulation l and

that the chemical profile is given by $\mathbf{s}$ at time t. Taking $\Delta t \to 0$

$$\frac{\partial C_k}{\partial t} = \sum_{\mathbf{s}} \left[T^+_{i-1}(\mathbf{s}) P(\mathbf{s},t) C_k(i-1,t) - \left(T^+_i(\mathbf{s}) + T^-_i(\mathbf{s}) \right) P(\mathbf{s},t) C_k(i,t) \right.$$
$$\left. + T^-_{i+1}(\mathbf{s}) P(\mathbf{s},t) C_k(i+1,t) \right]. \tag{6.6}$$

The evolution of the chemical distribution, $P(\mathbf{s},t)$ may then be found as follows:

$$\frac{\partial P}{\partial t} = \sum_{i=1}^{L} [\overbrace{d_s \left((s_i+1) P(J_i^+ \mathbf{s}, t) - s_i P(\mathbf{s},t) \right)}^{s \text{ moving right}} + \overbrace{d_s \left((s_i+1) P(J_i^- \mathbf{s}, t) - s_i P(\mathbf{s},t) \right)}^{s \text{ moving left}}]$$
$$+ \sum_{i=1}^{L} \sum_{k=1}^{N} \overbrace{\lambda_s \left((s_i+1) \text{Prob}(A_i^c \mathbf{s}(t), C_k(i,t)) - s_i \text{Prob}(\mathbf{s}(t), C_k(i,t)) \right)}^{\text{consumption by cell } k \text{ at } i}, \tag{6.7}$$

where the operators

$$J_i^+ \mathbf{s} = (s_1, \cdots, s_{i-1}, s_i + 1, s_{i+1} - 1, \cdots, s_L), \tag{6.8}$$

$$J_i^- \mathbf{s} = (s_1, \cdots, s_{i-1} - 1, s_i + 1, s_{i+1}, \cdots, s_L), \tag{6.9}$$

$$A_i^c \mathbf{s} = (s_1, \cdots, s_i + 1, \cdots, s_L), \tag{6.10}$$

describe a chemical molecule at i moving right, left or being consumed, respectively and the probability that the chemical profile is $\mathbf{s}(t)$ at time t and that cell k is at position i at time t is given by $\text{Prob}(\mathbf{s}(t), C_k(i,t))$. We now derive the average chemical profile:

$$(\langle s_1 \rangle, \ldots, \langle s_L \rangle) = \sum_{s_1} \cdots \sum_{s_L} (s_1 P(\mathbf{s},t), \ldots, s_L P(\mathbf{s},t)), \tag{6.11}$$

$$\langle (s_1, \ldots, s_L) \rangle = \sum_{s_1} \cdots \sum_{s_L} (s_1, \ldots, s_L) P(\mathbf{s},t), \tag{6.12}$$

$$\langle \mathbf{s} \rangle = \sum_{\mathbf{s}} \mathbf{s} P(\mathbf{s},t), \tag{6.13}$$

by multiplying equation (6.7) by $\mathbf{s}$ and summing over all $\mathbf{s}$. Then s_j satisfies

$$\frac{\partial \langle s_j \rangle}{\partial t} = \sum_{\mathbf{s}} \sum_{i=1}^{L} \left[d_s \left((s_i+1) s_j P(J_i^+ \mathbf{s}, t) - s_i s_j P(\mathbf{s},t) \right) + d_s \left((s_i+1) s_j P(J_i^- \mathbf{s}, t) - s_i s_j P(\mathbf{s},t) \right) \right]$$
$$+ \sum_{i=1}^{L} \sum_{k=1}^{N} C_k(i,t) \lambda_s \left((s_i+1) s_j P(A_i^c \mathbf{s}, t) - s_i s_j P(\mathbf{s},t) \right), \tag{6.14}$$

where we employ the moment closure assumption

$$\sum_{s_i} s_i \text{Prob}(\mathbf{s}(t), C_k(i,t)) = \sum_{\mathbf{s}} s_i \text{Prob}\left(\mathbf{s}(t), \sum_l C_k^l(i,t)\right), \tag{6.15}$$

$$\approx \langle s_i \rangle \langle C_k^l(i,t) \rangle, \tag{6.16}$$

$$= C_k(i,t) \sum_{s_i} s_i P(\mathbf{s},t), \tag{6.17}$$

as before. For $i \neq j-1, j$, taking the first two terms on the right-hand side of equation (6.14),

$$d_s \sum_{s_i} (s_i+1) s_j P(J_i^+ \mathbf{s}, t) - d_s \sum_{s_i} s_i s_j P(\mathbf{s},t) = d_s \sum_{\hat{s}_i} \hat{s}_i s_j P(\hat{\mathbf{s}},t) - d_s \sum_{s_i} s_i s_j P(\mathbf{s},t) = 0, \tag{6.18}$$

where $\hat{s}_i = s_i + 1$ and $\hat{s}_{i+1} = s_{i+1} - 1$ in equation (6.18) so that $J_i^+ \mathbf{s} = \hat{\mathbf{s}}$. Similarly, relabelling $\hat{s}_i = s_i + 1$ and $\hat{s}_{i-1} = s_{i-1} - 1$ for $i \neq j, j+1$ ensures that the second two terms in equation (6.14) cancel and relabelling $\hat{s}_i = s_i + 1$ for $i \neq j$ cancels the final two terms.

If $i = j - 1$ then

$$d_s \sum_{s_i} (s_i+1) s_{i+1} P(J_i^+ \mathbf{s}, t) - d_s \sum_{s_i} s_i s_{i+1} P(\mathbf{s},t) = d_s \sum_{\hat{s}_i} \hat{s}_i (\hat{s}_{i+1}+1) P(\hat{\mathbf{s}},t) - d_s \sum_{s_i} s_i s_{i+1} P(\mathbf{s},t), \tag{6.19}$$

$$= d_s \sum_{s_i} s_i P(\mathbf{s},t). \tag{6.20}$$

Proceeding similarly for $i = j$, $j+1$, or $j-1$ in the other terms we obtain

$$\frac{\partial \langle s_j \rangle}{\partial t} = d_s \left(\langle s_{j-1} \rangle - 2\langle s_j \rangle + \langle s_{j+1} \rangle \right) - \lambda_s \langle s_j \rangle \sum_{k=1}^{N} C_k(j,t), \tag{6.21}$$

and, taking the continuous functions $s(x,t) = s(j\Delta x, t) = \langle s_j(t) \rangle$ and $\hat{C}_k(x,t) = \hat{C}_k(j\Delta x, t) = C_k(j,t)$ for lattice spacing $\Delta x \ll 1$ and Taylor expanding in Δx,

$$\frac{\partial s}{\partial t} = d_s (\Delta x)^2 \frac{\partial^2 s}{\partial x^2} - \lambda_s s \sum_{k=1}^{N} \hat{C}_k + \mathcal{O}(\Delta x^3). \tag{6.22}$$

Similarly to Chapters 4 and 5, to maintain cell migration as the lattice spacing $\Delta x \to 0$, the

diffusivity d_s must scale so that $D = \lim_{\Delta x \to 0, d_s \to \infty} d_s(\Delta x)^2$ is constant and

$$\frac{\partial s}{\partial t} = D\frac{\partial^2 s}{\partial x^2} - \lambda_s s \sum_{k=1}^{N} \hat{C}_k. \tag{6.23}$$

6.1.2 Types of sensing

We consider two forms of chemotactic sensing, as in Baker *et al.* (2010):

1. local sensing: $T_i^{\pm} = \alpha s_i$;
2. non-local sensing: $T_i^{\pm} = \alpha s_{i\pm 1}$.

6.1.2.1 Local sensing ($T_i^{\pm} = \alpha s_i$)

Local sensing leads to movement *down* the concentration gradient of s, since movement rates increase when cells are in areas with large numbers of chemicals. Cells therefore move frequently until they have moved to an area of low chemical. Hence the chemical is a chemorepellent (as in the excluded regions between CNCC migratory streams). Taking $T_i^{\pm} = \alpha s_i$ in equation (6.6) gives

$$\frac{\partial C_k}{\partial t} = \alpha \sum_{\mathbf{s}} \left[s_{i-1} P(\mathbf{s}, t) C_k(i-1, t) - 2 s_i P(\mathbf{s}, t) C_k(i, t) + s_{i+1} P(\mathbf{s}, t) C_k(i+1, t) \right], \tag{6.24}$$

$$= \alpha \left(\langle s_{i-1} \rangle C_k(i-1, t) - 2\alpha \langle s_i \rangle C_k(i, t) + \langle s_{i+1} \rangle C_k(i+1, t) \right], \tag{6.25}$$

where $\langle s_i \rangle = \sum_{\mathbf{s}} s_i P(\mathbf{s}, t)$, as before.

Taking $s(x, t) = s(j\Delta x, t) = \langle s_j(t) \rangle$ and $\hat{C}_k(x, t) = \hat{C}_k(j\Delta x, t) = C_k(j, t)$ for lattice spacing $\Delta x \ll 1$ as before, we Taylor expand in Δx to give

$$\frac{\partial \hat{C}_j}{\partial t} = \alpha(\Delta x)^2 \frac{\partial^2}{\partial x^2} \left(s\hat{C}_j \right) + \mathcal{O}(\Delta x^3). \tag{6.26}$$

Now we let $\Delta x \to 0$ and $\alpha \to \infty$ with $\hat{\alpha} = \lim_{\Delta x \to 0} \alpha(\Delta x)^2$ constant and note that the resultant equation is the same for all C_j. Hence if all cells have the same initial conditions then $C_j = C_1 \ \forall j$ and $C(x, t) = \sum_j \hat{C}_j(x, t)$ satisfies

$$\frac{\partial s}{\partial t} = D\frac{\partial^2 s}{\partial x^2} - \lambda_s s C, \tag{6.27}$$

$$\frac{\partial C}{\partial t} = \hat{\alpha} \frac{\partial^2}{\partial x^2} (sC). \tag{6.28}$$

At $x = L\Delta x$

$$\frac{\partial C}{\partial t} = \frac{\hat{\alpha}}{\Delta x}\frac{\partial}{\partial x}\left(s\frac{\partial C}{\partial x} + C\frac{\partial s}{\partial x}\right), \tag{6.29}$$

so that as $\Delta x \to 0$ we require

$$\frac{\partial}{\partial x}(sC)\bigg|_{x=L} = 0, \tag{6.30}$$

and no-flux boundary conditions are similarly derived for the chemorepellent at $x = 0$ and $x = L\Delta x$. At $x = 0$

$$\frac{\partial C}{\partial t} = \alpha s(\Delta x)C(\Delta x) - \alpha s(0)C(0) + P_c, \tag{6.31}$$

$$= \alpha\Delta x\left(s\frac{\partial C}{\partial x} + C\frac{\partial s}{\partial x}\right) + P_c. \tag{6.32}$$

This typifies a common problem with flux boundary conditions, in which the order of the approximation at the boundary is lower than in the domain. In this example, we take $\hat{\alpha} = \lim_{\Delta x\to 0}\alpha(\Delta x)^2$ to be finite, but the boundary condition depends upon $\alpha\Delta x$. Hence as $\Delta x \to 0$ whilst keeping $\hat{\alpha}$ finite, then $\alpha\Delta x \to \infty$. For any particular value of Δx, however, we may solve the system using equation (6.32), and this compares well with averaged simulation data (Figure 6.1(a)) for a range of parameter values (results not shown). Alternatively, if the influx of cells, P_c, is of order $\mathcal{O}(1/\Delta x)$ then taking $\Delta x \to 0$ in equation (6.32) gives the flux boundary condition

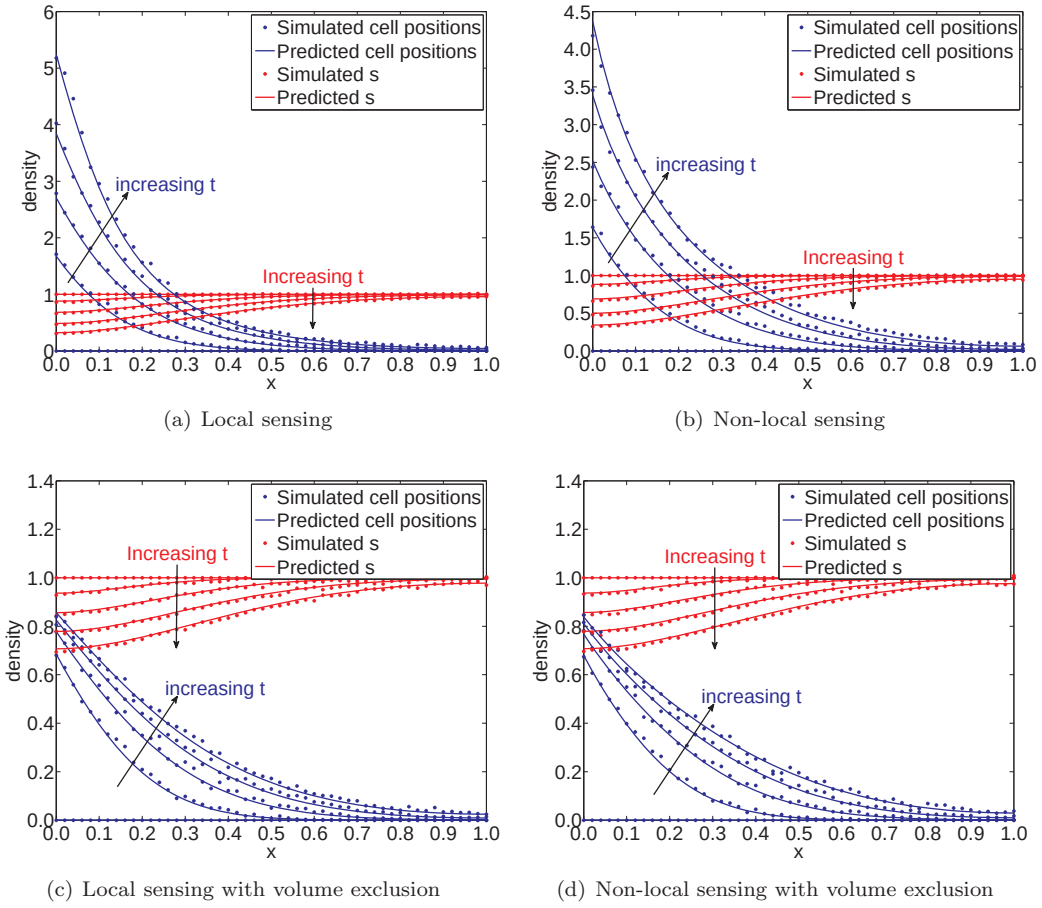
$$\hat{\alpha}\left(s\frac{\partial C}{\partial x} + C\frac{\partial s}{\partial x}\right) + \hat{P}_c = 0, \tag{6.33}$$

where $\hat{P}_c = \lim_{\Delta x\to 0} P_c\Delta x$ is finite since $P_c = \mathcal{O}(1/\Delta x)$. Numerical solutions to the system with equation (6.33) are found to be equivalent to those with equation (6.32) (result not shown).

The right-hand side of equation (6.28) can be thought of as two terms:

$$\frac{\partial C}{\partial t} = \hat{\alpha}\frac{\partial}{\partial x}\left(s\frac{\partial C}{\partial x} + C\frac{\partial s}{\partial x}\right). \tag{6.34}$$

The first is a chemical-dependent diffusion term, with higher diffusivity where there is more chemical, whilst the second is an advection term down concentration gradients of s. Thus the cells move away quickly from areas of high chemical, down the concentration gradient, to areas where the chemical is sparse, thus demonstrating chemorepellence. Since cells enter the domain at the left-hand boundary



(a) Local sensing

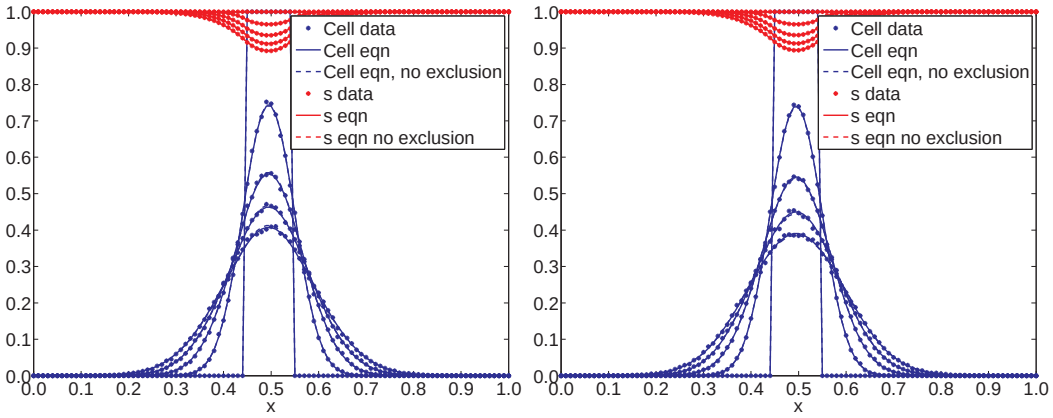
(b) Non-local sensing

(c) Local sensing with volume exclusion

(d) Non-local sensing with volume exclusion

Figure 6.1: A comparison of typical averaged simulations of the on-lattice chemotaxis models described in Section 6.1 and numerical solutions to the derived equations for (a) local sensing (equations (6.27) and (6.28)); (b) non-local sensing (equations (6.38) and (6.39)); (c) local sensing with volume exclusion (equations (6.48) and (6.49)) and (d) non-local sensing with volume exclusion (equations (6.56) and (6.57)). Each simulation is performed with $\alpha = 2.22$, $\Delta t = 1$, $\Delta x = 1$, $D_s = 1$, $\lambda_s = 0.01$ and is averaged over 1000 runs. The domain is initially empty with an influx of cells at a rate $P_c = 0.05$.

and reduce the chemorepellent in that area, thus creating an area of lower chemical concentration, the cellular density grows next to the boundary. The increasing levels of cell densities next to the domain entrance is demonstrated in both the IBM simulations and numerical solutions to equations (6.27) and (6.28) with equation (6.32) (Figure 6.1(a)).



(a) Local sensing with volume exclusion

(b) Non-local sensing with volume exclusion

Figure 6.2: A comparison of typical averaged simulations of the on-lattice chemotaxis models described in Section 6.1 and numerical solutions to the derived equations for (a) local sensing with volume exclusion (equations (6.48) and (6.49)) and (b) non-local sensing with volume exclusion (equations (6.56) and (6.57)). Each simulation is performed with $\alpha = 2.22$, $\Delta t = 1$, $\Delta x = 1$, $D_s = 1$, $\lambda_s = 0.01$ and is averaged over 1000. Ten cells are initially placed, one at each lattice site, in the middle of the domain and no flux conditions are taken at each boundary.

6.1.2.2 Non-local sensing ($T_i^{\pm} = \alpha s_{i\pm1}$)

Non-local sensing gives rise to movement up concentration gradients of s, so that s is now a chemoattractant. Substituting $T_i^{\pm} = \alpha s_{i\pm1}$ into equation (6.6):

$$\frac{\partial C_k}{\partial t} = \alpha \sum_{\mathbf{s}} \left[s_i P(\mathbf{s},t) C_k(i-1,t) - \left(s_{i+1} + s_{i-1}\right) P(\mathbf{s},t) C_k(i,t) + s_i P(\mathbf{s},t) C_k(i+1,t) \right], \tag{6.35}$$

$$= \alpha \left(\langle s_i \rangle C_k(i-1,t) - \left(\langle s_{i+1} \rangle + \langle s_{i-1} \rangle \right) C_k(i,t) + \langle s_i \rangle P(\mathbf{s},t) \right). \tag{6.36}$$

If $s(x,t) = s(j\Delta x, t) = \langle s_j(t) \rangle$ and $\hat{C}_k(x,t) = \hat{C}_k(j\Delta x, t) = C_k(j,t)$ for lattice spacing $\Delta x \ll 1$ as before, then expanding in Δx gives

$$\frac{\partial \hat{C}_k}{\partial t} = \alpha (\Delta x)^2 \frac{\partial}{\partial x} \left(s \frac{\partial \hat{C}_k}{\partial x} - C \frac{\partial s}{\partial x} \right), \tag{6.37}$$

and if $\Delta x \to 0$ and $\alpha \to \infty$ with $\hat{\alpha} = \lim_{\Delta x \to 0} \alpha (\Delta x)^2$ constant then $C(x,t) = \sum_j \hat{C}_j(x,t)$ and s satisfy

$$\frac{\partial s}{\partial t} = D_s \frac{\partial^2 s}{\partial x^2} - \lambda_s s C, \tag{6.38}$$

$$\frac{\partial C}{\partial t} = \hat{\alpha} \frac{\partial}{\partial x} \left(s \frac{\partial C}{\partial x} - C \frac{\partial s}{\partial x} \right). \tag{6.39}$$

At $x = L\Delta x$

$$\frac{\partial C}{\partial t} = \frac{\hat{\alpha}}{\Delta x}\frac{\partial}{\partial x}\left(s\frac{\partial C}{\partial x} - C\frac{\partial s}{\partial x}\right), \tag{6.40}$$

so that as $\Delta x \to 0$

$$s\frac{\partial C}{\partial x}|_{x=L} - C\frac{\partial s}{\partial x}\Big|_{x=L} = 0. \tag{6.41}$$

No flux conditions are similarly taken at both boundaries for the chemoattractant, as before. At $x = 0$ there is an influx of cells, with

$$\frac{\partial C}{\partial t} = \alpha s(0)C(\Delta x, t) - \alpha s(\Delta x)C(0, t) + P_c, \tag{6.42}$$

$$= \alpha\Delta x\left(s\frac{\partial C}{\partial x} - C\frac{\partial s}{\partial x}\right) + P_c. \tag{6.43}$$

As before we may solve for a particular value of Δx and find that this compares to averaged simulation data (Figure 6.1(b)) for a range of parameter values (results not shown). In this case the diffusion term with diffusivity s is modified by an advection term up concentration gradients of the chemical, as we would expect, since our sensing mechanism results in chemoattraction up gradients of s. Individuals entering at the left-hand boundary still reduce the levels of s as in Section 6.1.2.1, but the positive gradient in s now drives movement away from the boundary (Figure 6.1(b)), so that the cell density profile is more spread out towards the end of the domain, and the numbers of individuals at the left-hand boundary is reduced compared to local sensing (Figure 6.1(a)).

6.1.3 Excluded volume effects

This model may be extended to incorporate volume exclusion by allowing at most one cell at each lattice point. equations (6.27) and (6.38) are not affected by volume exclusion, whilst equation (6.3)

is modified in the following way:

$$\frac{\partial C_k}{\partial t} = -\sum_{\mathbf{s}} C_k(i,t)\left(T_i^+(\mathbf{s})P(\mathbf{s},t)\left(1-\sum_{j\neq k}C_j(i+1,t)\right)+T_i^-(\mathbf{s})P(\mathbf{s},t)\left(1-\sum_{j\neq k}C_j(i-1,t)\right)\right)$$
$$+\sum_{\mathbf{s}} T_{i-1}^+(\mathbf{s})P(\mathbf{s},t)C_k(i-1,t)\left(1-\sum_{j\neq k}C_j(i,t)\right)$$
$$+\sum_{\mathbf{s}} T_{i+1}^-(\mathbf{s})P(\mathbf{s},t)C_k(i+1,t)\left(1-\sum_{j\neq k}C_j(i,t)\right), \tag{6.44}$$

taking the moment closure assumption that the probability that cell k is at point i and cell j is at point $i+1$ is approximately equal to $C_k(i,t)C_j(i+1,t)$.

6.1.3.1 Local sensing

For local sensing, substituting $T_i^{\pm}(\mathbf{s}) = \alpha s_i$ into equation (6.44),

$$\frac{\partial C_k}{\partial t} = -\alpha\langle s_i\rangle C_k(i,t)\left(\left(1-\sum_{j\neq k}C_j(i+1,t)\right)+\left(1-\sum_{j\neq k}C_j(i-1,t)\right)\right)$$
$$+\alpha\langle s_{i-1}\rangle C_k(i-1,t)\left(1-\sum_{j\neq k}C_j(i,t)\right)+\alpha\langle s_{i+1}\rangle C_k(i+1,t)\left(1-\sum_{j\neq k}C_j(i,t)\right), \tag{6.45}$$

for $\langle s_i\rangle = \sum_{\mathbf{s}} s_i P(\mathbf{s},t)$. If $s(x,t) = s(j\Delta x,t) = \langle s_j(t)\rangle$ and $\hat{C}_k(x,t) = \hat{C}_k(j\Delta x,t) = C_k(j,t)$ for lattice spacing $\Delta x \ll 1$ as before, then expanding in Δx gives

$$\frac{\partial \hat{C}_k}{\partial t} = \alpha(\Delta x)^2\frac{\partial^2}{\partial x^2}\left(s\hat{C}_k\right) - \alpha(\Delta x)^2\frac{\partial^2}{\partial x^2}\left(s\hat{C}_k\right)\sum_{j\neq k}\hat{C}_j + \alpha(\Delta x)^2 s\hat{C}_k\sum_{j\neq k}\frac{\partial^2\hat{C}_j}{\partial x^2} + \mathcal{O}(\Delta x^3). \tag{6.46}$$

If all cells have the same initial conditions, then $\hat{C}_k = \hat{C}_1$, for $k = 1,\dots N$, so that

$$\frac{\partial \hat{C}_k}{\partial t} = \alpha(\Delta x)^2\frac{\partial}{\partial x}\left(\frac{\partial}{\partial x}\left(s\hat{C}_k\right) - (N-1)\hat{C}_k^2\frac{\partial s}{\partial x}\right) + \mathcal{O}(\Delta x^3). \tag{6.47}$$

Letting $\Delta x \to 0$ and $\alpha \to \infty$ with $\hat{\alpha} = \lim_{\Delta x\to 0,\alpha\to\infty}\alpha(\Delta x)^2$ and $D_x = \lim_{\Delta x\to 0, d_s\to\infty} d_s(\Delta x)^2$ constant then $C(x,t) = \sum_j \hat{C}_j(x,t)$ and s satisfy

$$\frac{\partial s}{\partial t} = D_s\frac{\partial^2 s}{\partial x^2} - \lambda_s sC, \tag{6.48}$$
$$\frac{\partial C}{\partial t} = \hat{\alpha}\frac{\partial}{\partial x}\left(\frac{\partial}{\partial x}(sC) - \frac{N-1}{N}C^2\frac{\partial s}{\partial x}\right), \tag{6.49}$$

and we take no-flux boundary conditions on s at $x = 0$, $x = L$ and C at $x = L$, and

$$\frac{\partial C}{\partial t} = \alpha s(\Delta x, t)C(\Delta x, t)\left(1 - \frac{N-1}{N}C(0,t)\right) - \alpha s(0,t)C(0,t)\left(1 - \frac{N-1}{N}C(\Delta x, t)\right) + P_c(1 - C(0,t)), \tag{6.50}$$

$$= \alpha \Delta x \left(s\frac{\partial C}{\partial x} + C\frac{\partial s}{\partial x} - \frac{N-1}{N}C^2\frac{\partial s}{\partial x}\right) + P_c(1-C), \tag{6.51}$$

at $x = 0$. Volume exclusion thus modifies the equation previously derived for local sensing (equation (6.28)), reducing the flux. Equation (6.49) can be thought of as

$$\frac{\partial C}{\partial t} = \hat{\alpha}\frac{\partial}{\partial x}\left(s\frac{\partial C}{\partial x} + \frac{\partial s}{\partial x}C\left(1 - \frac{N-1}{N}C\right)\right), \tag{6.52}$$

so that the diffusive term calculated previously in equation (6.28) is unchanged (since all modifications to diffusion due to volume exclusion are higher order terms), whilst the rate of advection down gradients of s is reduced by the presence of other individuals. Note that taking s to be homogeneous, so that there is no chemotaxis, reduces equation (6.52) to on-lattice diffusion with volume exclusion, which is simply given by the diffusion equation.

Numerical solutions to equations (6.50) and (6.51) compare well to averaged simulation data (Figure 6.1(c)) for a range of parameter values (results not shown) and display a significantly different profile to the system without volume exclusion (Figure 6.1(a)). As is expected, volume exclusion further slows the rate of cellular migration into the domain, since new cells are often prevented from entering the first lattice site by the presence of other cells. We also expect the rate of migration through the domain to increase due to the reduction in advection down the chemical gradient, however this effect is so small as to be negligible when the dominating flux boundary effects are removed (Figure 6.2(a)).

6.1.3.2 Non-local sensing

Similarly, with non-local sensing and excluded volume, substituting $T_i^{\pm}(\mathbf{s}) = \alpha s_{i\pm1}$ in equation (6.44),

$$\frac{\partial C_k}{\partial t} = -\alpha C_k(i,t)\left(\langle s_{i+1}\rangle\left(1 - \sum_{j\neq k}C_j(i+1,t)\right) + \langle s_{i-1}\rangle\left(1 - \sum_{j\neq k}C_j(i-1,t)\right)\right) + \alpha\langle s_i\rangle C_k(i-1,t)\left(1 - \sum_{j\neq k}C_j(i,t)\right) + \alpha\langle s_i\rangle C_k(i+1,t)\left(1 - \sum_{j\neq k}C_j(i,t)\right), \tag{6.53}$$

where $\langle s_i \rangle = \sum_{\mathbf{s}} s_i P(\mathbf{s}, t)$. If $s(x,t) = s(j\Delta x, t) = \langle s_j(t) \rangle$ and $\hat{C}_k(x,t) = \hat{C}_k(j\Delta x, t) = C_k(j,t)$ for lattice spacing $\Delta x \ll 1$ as before, then expanding in Δx gives

$$\frac{\partial \hat{C}_k}{\partial t} = \alpha(\Delta x)^2 \frac{\partial}{\partial x}\left(s\frac{\partial \hat{C}_k}{\partial x} - \hat{C}_k \frac{\partial s}{\partial x}\right) + \alpha(\Delta x)^2 \hat{C}_k \frac{\partial^2}{\partial x^2}\left(s\sum_{j\neq k} C_j\right) - s\frac{\partial^2 C}{\partial x^2}\sum_{j\neq k} C_j + \mathcal{O}(\Delta x^3). \tag{6.54}$$

If all cells have the same initial conditions, then $\hat{C}_k = \hat{C}_1$, for $k = 1, \ldots N$, so that

$$\frac{\partial \hat{C}_k}{\partial t} = \alpha(\Delta x)^2 \frac{\partial}{\partial x}\left(s\frac{\partial \hat{C}_k}{\partial x} - \hat{C}_k \frac{\partial s}{\partial x} + (N-1)C_k^2 \frac{\partial s}{\partial x}\right) + \mathcal{O}(\Delta x^3). \tag{6.55}$$

Letting $\Delta x \to 0$ and $\alpha \to \infty$ with $\hat{\alpha} = \lim_{\Delta x \to 0, \alpha \to \infty} \alpha(\Delta x)^2$ and $D_x = \lim_{\Delta x \to 0, d_s \to \infty} d_s(\Delta x)^2$ constant, $C(x,t) = \sum_j \hat{C}_j(x,t)$ and s satisfy

$$\frac{\partial s}{\partial t} = D_s \frac{\partial^2 s}{\partial x^2} - \lambda_s s C, \tag{6.56}$$

$$\frac{\partial C}{\partial t} = \hat{\alpha}\frac{\partial}{\partial x}\left(s\frac{\partial C}{\partial x} - C\frac{\partial s}{\partial x} + \frac{N-1}{N}C^2\frac{\partial s}{\partial x}\right), \tag{6.57}$$

with no-flux boundary conditions on s at $x = 0$, $x = L$ and on C at $x = L$, and

$$\begin{aligned}\frac{\partial C}{\partial t} &= \alpha s(0) C(\Delta x, t)\left(1 - \frac{N-1}{N}C(0,t)\right) - \alpha s(\Delta x) C(0,t)\left(1 - \frac{N-1}{N}C(\Delta x, t)\right) \\ &\quad + P_c(1 - C(0,t)), \qquad (6.58)\\ &= \alpha \Delta x \left(s\frac{\partial C}{\partial x} - C\frac{\partial s}{\partial x} + \frac{N-1}{N}C^2\frac{\partial s}{\partial x}\right) + P_c(1-C), \qquad (6.59)\end{aligned}$$

at $x = 0$. As with local sensing, we may rewrite equation (6.57) as

$$\frac{\partial C}{\partial t} = \hat{\alpha}\frac{\partial}{\partial x}\left(s\frac{\partial C}{\partial x} - C\frac{\partial s}{\partial x}\left(1 - \frac{N-1}{N}C\right)\right), \tag{6.60}$$

where the diffusive term is again unchanged by volume exclusion, but the advective flux up gradients of chemical is reduced by the presence of other individuals.

Numerical solutions to equations (6.50) and (6.51) also compare well to averaged simulation data (Figure 6.1(d)) for a range of parameter values (results not shown), with volume exclusion slowing the rate of entry into the domain as predicted (Figure 6.1(d) compared to Figure 6.1(b)). In fact, when the rates of advection in the local and non-local sensing cases are reduced by volume exclusion the

resulting cell density profiles are very similar (Figures 6.1(c) and 6.1(d)). The similarity in profiles with volume exclusion indicates that the diffusion term, which is the same in both equations (6.52) and (6.60), is dominant over the exclusion-reduced advection term down (equation (6.52)) or up (equation (6.60)) concentration gradients of the chemical. As in the local sensing case, the expected slowing of migration rate through the domain is negligible when the flux boundary effects are removed (Figure 6.2(b)).

6.2 Model 2: on-lattice cells with continuum chemoattractant

For large numbers of chemical molecules, the models in Section 6.1 may be inefficient, since higher numbers of molecules take longer to simulate. However, models can often be accurately approximated by a continuum equation, even within a single simulation. Numerical solutions to a continuum equation are typically faster to compute than the equivalent individual-based system. Since biological systems often involve many different species of chemicals and cell populations (Vigmond *et al.*, 2009; Perfahl *et al.*, 2011), the speed-up from using such hybrid models, in which some species are represented by continuum equations whilst others are modelled using an individual-based framework, may be significant. Large-scale computational models are becoming more widespread, leading to an increasing demand to analyse the effect of moving from a purely individual-based model to a hybrid model. Determining which equations to use for continuum-modelled species and how best to link between the continuum and discrete parts of models is a current and highly relevant question in the field (Franz and Erban, 2013).

6.2.1 Hybrid model

As in Chapters 2 and 3, the chemoattractant will diffuse and be consumed by the cells around their position so that

$$\frac{\partial s}{\partial t} = D_s \frac{\partial^2 s}{\partial x^2} - \lambda_s s \sum_{i=1}^{N} e^{-(x-c_i)^2/\delta}, \tag{6.61}$$

where D_s is the diffusion coefficient, the current position of cell i is given by c_i and δ gives a measure of how far away cells may consume chemoattractant. With this model each cell consumes the chemoattractant at their current lattice site at a rate, λ_s, as in the non-hybrid system. The

cells are modelled as on-lattice individuals, as in Section 6.1, and move to the left or right with rate $T_i^{\pm} = \alpha s(c_i \pm \Delta x)$.

6.2.2 Hybrid simulation algorithm

In previous simulations we have used the Gillespie simulation algorithm (Gillespie, 1977), which derives the distribution of waiting times for the next "event" and then draws the next time step in the simulation from this distribution. In this derivation, the rate of events is dependent only upon the distribution of species at the previous time point, which is assumed not to change before the next time step. In the hybrid model, however, the chemoattractant equation is continuously changing in time, so that the rates of cell movement, which depend upon the chemical concentration, also change over time. Hence the distribution of next time points cannot be derived as in the Gillespie algorithm.

Note that in the special case where there is no diffusion of chemical (*i.e.* $D_s = 0$), it is possible to modify the derivation of the Gillespie algorithm to find the modified distribution of waiting times in the following way. Let $P_\mu(\tau|\mathbf{c}, t)$ be the probability that the next μ event occurs in $(t+\tau, t+\tau+\delta\tau)$ given that the cells are currently at positions $\mathbf{c}$. If $a_\mu(t)$ is the propensity function for event μ at time t,

$$P_\mu(\tau|\mathbf{c}, t) = \text{Prob(no } \mu \text{ events in } [t, t+\tau))\text{Prob}(\mu \text{ event in } [t+\tau, t+\tau+\delta\tau)), \tag{6.62}$$

$$= G_\mu(\tau)\text{Prob}(\mu \text{ event in } [t+\tau, t+\tau+\delta\tau)), \tag{6.63}$$

where $G_\mu(\tau)$ is the probability that no μ events occur in $[t, t+\tau)$. We consider the case where $a_\mu(t) = \alpha s(c_\mu, t)$ (other forms of $T_i^{\pm}$ can be derived similarly). Since we take $D_s = 0$

$$\frac{\partial s}{\partial t} = -\lambda_s s \sum_{i=1}^{N} e^{-(x-c_i)^2/\delta}, \tag{6.64}$$

so $s(x, t+\tau) = s(x,t)\exp[-A(\mathbf{c},x)\tau]$ where $A(\mathbf{c},x) = \lambda_s \sum_j \exp[-(c_j - x)^2/\delta]$. Then

$$G_\mu(\sigma + \delta\sigma) = \text{Prob(no } \mu \text{ events in } [t, t+\sigma))\text{Prob(no } \mu \text{ events in } [t+\sigma, t+\sigma+\delta\sigma)), \tag{6.65}$$

$$= \text{Prob(no } \mu \text{ events in } [t, t+\sigma))(1 - \text{Prob}(\mu \text{ events in } [t+\sigma, t+\sigma+\delta\sigma)), \tag{6.66}$$

$$= G_\mu(\sigma)(1 - \int_\sigma^{\sigma+\delta\sigma} \alpha s(c_\mu, t+\bar{\sigma})\ \mathrm{d}\bar{\sigma}, \tag{6.67}$$

$$= G_\mu(\sigma)(1 - \int_\sigma^{\sigma+\delta\sigma} \alpha s(c_\mu, t) e^{-A(\mathbf{c},c_\mu)\bar{\sigma}}\ \mathrm{d}\bar{\sigma}, \tag{6.68}$$

but

$$\int_\sigma^{\sigma+\delta\sigma} e^{-A(\mathbf{c},c_\mu)\bar{\sigma}}\ \mathrm{d}\bar{\sigma} = \left[-\frac{e^{-A(\mathbf{c},c_\mu)\bar{\sigma}}}{A(\mathbf{c},c_\mu)}\right]_\sigma^{\sigma+\delta\sigma}, \tag{6.69}$$

$$= \frac{e^{-A(\mathbf{c},c_\mu)\sigma}}{A(\mathbf{c},c_\mu)}\left(1 - e^{-A(\mathbf{c},c_\mu)\delta\sigma}\right), \tag{6.70}$$

$$= \delta\sigma e^{-A(\mathbf{c},c_\mu)\sigma} + \mathcal{O}(\delta\sigma^2). \tag{6.71}$$

Hence

$$G_\mu(\sigma + \delta\sigma) = G_\mu(\sigma)(1 - \alpha s(c_\mu, t)\delta\sigma e^{-A(\mathbf{c},c_\mu)\sigma} + \mathcal{O}(\delta\sigma^2), \tag{6.72}$$

$$\frac{\partial G_\mu}{\partial \sigma} = -\alpha s(c_\mu, t) e^{-A(\mathbf{c},c_\mu)\sigma}, \tag{6.73}$$

letting $\delta\sigma \to 0$. Solving for $G_\mu(\sigma)$ whilst noting that $G_\mu(0) = \text{Prob(no } \mu \text{ events in } [t,t)) = 1$,

$$G_\mu(\sigma) = \exp\left(\alpha \frac{s(c_\mu,t)}{A(\mathbf{c},c_\mu)} e^{-A(\mathbf{c},c_\mu)\sigma}\right), \tag{6.74}$$

and hence

$$P_\mu(\tau|\mathbf{c},t) = \exp\left(\alpha \frac{s(c_\mu,t)}{A(\mathbf{c},c_\mu)} e^{-A(\mathbf{c},c_\mu)\tau}\right) \text{Prob}(\mu \text{ event in } [t+\tau, t+\tau+\delta\tau)), \tag{6.75}$$

$$= \alpha s(c_\mu,t) e^{-A(\mathbf{c},c_\mu)\tau} \exp\left(\alpha \frac{s(c_\mu,t)}{A(\mathbf{c},c_\mu)} e^{-A(\mathbf{c},c_\mu)\tau}\right). \tag{6.76}$$

The cumulative distribution function (CDF) is distributed uniformly (similarly to the Gillespie

algorithm) and is given by

$$F_\mu(\tau|\mathbf{c},t) = \int_0^\tau P_\mu(\tau'|\mathbf{c},t)\ \mathrm{d}\tau', \tag{6.77}$$

$$= \exp\left(\alpha\frac{s(c_\mu,t)}{A(\mathbf{c},c_\mu)}e^{-A(\mathbf{c},c_\mu)\tau}\right) - \exp\left(\alpha\frac{s(c_\mu,t)}{A(\mathbf{c},c_\mu)}\right), \tag{6.78}$$

so if $r \sim U([0,1])$

$$\tau_\mu = -\frac{1}{A(\mathbf{c},c_\mu)}\log\left(\frac{A(\mathbf{c},c_\mu)}{\alpha s(c_\mu,t)}\log\left(r + e^{\alpha\frac{s(c_\mu,t)}{A(\mathbf{c},c_\mu)}}\right)\right), \tag{6.79}$$

is distributed according to $F_\mu(\tau|\mathbf{c},t)$. Thus to simulate the model with no chemical diffusion we may draw a uniformly-distributed random number, r, for each reaction μ and use equation (6.79) to find the next time for that reaction. The smallest of these times is then used and the corresponding reaction occurs.

To simulate the model with chemical diffusion, however, we cannot use the Gillespie algorithm described here since there is no simple form of solution to equation (6.61). Instead we take a fixed time step, small enough that there is very little chance that more than one movement will occur during any time step. Then at each time step any particular movement happens with a calculated probability and we assume that at most one movement occurs. This is less efficient than the Gillespie algorithm, since during most time steps there is no change in individual positions, the only calculation that occurs is to solve for the chemical profile at that time point. However, this does allow simulation of a wider range of chemical evolution equations, including those with spatial dependencies.

6.2.3 Derivation of continuum equations

Modelling a species as a continuum whilst keeping track of the concentration at each spatial point presents an additional challenge to the models discussed in Section 6.1. The question is how to find the probability of the chemoattractant functional form in some tractable way that enables us to find the rates of cell movement. Since the cells are on-lattice, their rates of movement depend only on the chemoattractant concentration at the lattice points, and not in the entirety of the x-axis. The evolution of the chemoattractant may be continuous in space, however, we may restrict it naturally to depend only upon values at lattice sites by taking a finite difference approximation to spatial derivatives. We will then let the lattice spacing Δx tend to zero to obtain a continuous cell density (equation (6.28)), thus recovering the original continuous derivatives. Hence we consider $P(\mathbf{s},t) =$

$P(\mathbf{s}+\delta\mathbf{s},t)$ to be the probability that the concentration of s at the lattice points $\mathbf{x}=[1,\ldots,L]$ is in the region $\mathbf{s}+\delta\mathbf{s}=(s_1+\delta_1,\ldots,s_L+\delta_L)$ at time t for $|\delta s|\ll 1$. Then the probability, $C_k(i,t)$, that a cell k will be at lattice point i at time t is given by equation (6.6), as before.

If the chemoattractant evolves according to

$$\frac{\partial s}{\partial t}=g(\mathbf{c},s,x), \tag{6.80}$$

then to find the evolution of $s(\mathbf{x})=\mathbf{s}$, we take a finite difference approximation, $\hat{g}(\mathbf{c},\mathbf{s},i)$, to equation (6.80) for lattice spacing $\Delta x\ll 1$ and time step $\Delta t\ll 1$ so that

$$s_i(t+\Delta t)=s_i(t)+\Delta t\hat{g}(\mathbf{c},\mathbf{s}(t),i). \tag{6.81}$$

Hence

$$\text{Prob}(\mathbf{s}\text{ at }t+\Delta t)=\sum_{\mathbf{c}}\text{Prob}(\text{cells at }\mathbf{c}\text{ and }s(\mathbf{x},t)\text{ is such that }s(\mathbf{x},t+\Delta t)=\mathbf{s}\text{ given cells at }\mathbf{c}), \tag{6.82}$$

and the average chemoattractant profile,

$$\langle(s_1,\cdots,s_L)\rangle=\int_{s_1}\cdots\int_{s_L}(s_1,\cdots,s_L)P(\mathbf{s},t)\,\mathrm{d}s_1\cdots\,\mathrm{d}s_L, \tag{6.83}$$

$$\langle\mathbf{s}\rangle=\int_{\mathbf{s}}\mathbf{s}P(\mathbf{s},t)\,\mathrm{d}\mathbf{s}, \tag{6.84}$$

can be found by multiplying equation (6.84) by s_i and integrating over all $\mathbf{s}$:

$$\int_{\mathbf{s}}s_iP(\mathbf{s},t+\Delta t)\,\mathrm{d}\mathbf{s}=\int_{\mathbf{s}}\sum_{\mathbf{c}}\text{Prob}\left(\sum_l C_1^l(c_1,t),\ldots,\sum_l C_N^l(c_N,t),\mathbf{s}\text{ given }\mathbf{c}\right), \tag{6.85}$$

$$\approx\langle C_1^l(c_1,t)\rangle\cdots\langle C_N^l(c_N,t)\rangle\langle\mathbf{s}\text{ given }\mathbf{c}\rangle, \tag{6.86}$$

$$=\int_{\mathbf{s}}\sum_{c_1}\cdots\sum_{c_N}C_1(c_1,t)\cdots C_N(c_N,t)s_iP(\mathbf{R}(\mathbf{c},\mathbf{s})^{-1},t)\left|\det\left(\frac{\partial\mathbf{s}}{\partial\mathbf{R}(\mathbf{s},\mathbf{c})}\right)\right|\,\mathrm{d}\mathbf{s}, \tag{6.87}$$

where we define $(\mathbf{R}(\mathbf{c},\mathbf{s}))_i=s_i+\Delta t\hat{g}(\mathbf{c},\mathbf{s},i)$, and the Jacobian, $|\det(\partial\mathbf{s}/\partial\mathbf{R}(\mathbf{s},\mathbf{c}))|$, is due to the change in volume of the infinitesimal, $\delta\mathbf{s}$, when changing variables. Here, equation (6.86) follows from equation (6.85) by taking a moment closure approximation, as in Section 6.1.1.

Now we make the change of variables $\hat{\mathbf{s}} = \mathbf{R}(\mathbf{c}, \mathbf{s})^{-1}$ and note that this introduces a factor $|\det(\partial \mathbf{R}(\mathbf{s}, \mathbf{c})/\partial \mathbf{s})|$ so that

$$\int_{\mathbf{s}} s_i P(\mathbf{s}, t + \Delta t)\ \mathrm{d}\mathbf{s} = \int_{\hat{\mathbf{s}}} \sum_{\mathbf{c}} C_1(c_1, t) \cdots C_N(c_N, t)(\mathbf{R}(\mathbf{c}, \hat{\mathbf{s}}))_i P(\hat{\mathbf{s}}, t)\ \mathrm{d}\hat{\mathbf{s}}, \tag{6.88}$$

$$= \int_{\hat{\mathbf{s}}} \sum_{\mathbf{c}} C_1(c_1, t) \cdots C_N(c_N, t)(\hat{s}_i + \Delta t \hat{g}(\mathbf{c}, \hat{\mathbf{s}}, i))P(\hat{\mathbf{s}}, t)\ \mathrm{d}\hat{\mathbf{s}}. \tag{6.89}$$

Since $\sum_{c_j} C_j(c_j, t) = 1$,

$$\frac{1}{\Delta t}\left(\int_{\mathbf{s}} s_i P(\mathbf{s}, t + \Delta t)\ \mathrm{d}\mathbf{s} - \int_{\hat{\mathbf{s}}} \hat{s}_i P(\hat{\mathbf{s}}, t)\ \mathrm{d}\hat{\mathbf{s}}\right) = \int_{\hat{\mathbf{s}}} \sum_{\mathbf{c}} C_1(c_1, t) \cdots C_N(c_N, t)\hat{g}(\mathbf{c}, \mathbf{s}, i))P(\hat{\mathbf{s}}, t)\ \mathrm{d}\hat{\mathbf{s}}, \tag{6.90}$$

and as $\Delta t \to 0$

$$\frac{\partial \langle s_i \rangle}{\partial t} = \int_{\hat{\mathbf{s}}} \sum_{\mathbf{c}} C_1(c_1, t) \cdots C_N(c_N, t)\hat{g}(\mathbf{c}, \mathbf{s}, i))P(\hat{\mathbf{s}}, t)\ \mathrm{d}\hat{\mathbf{s}}. \tag{6.91}$$

For example, if

$$g(\mathbf{c}, s, x) = D_s \frac{\partial^2 s}{\partial x^2} - \lambda_s s \sum_{j=1}^{N} e^{-(x - c_j)^2/\delta}, \tag{6.92}$$

then, making a finite difference approximation using the lattice spacing $\Delta x \ll 1$

$$\hat{g}(\mathbf{c}, \mathbf{s}, i) = D_s \frac{s_{i+1} - 2s_i + s_{i-1}}{(\Delta x)^2} - \lambda_s s_i \sum_{j=1}^{N} e^{-(i - c_j)^2/\delta}, \tag{6.93}$$

and

$$\frac{\partial \langle s_i \rangle}{\partial t} = \int_{\hat{\mathbf{s}}} \sum_{\mathbf{c}} C_1(c_1, t) \cdots C_N(c_N, t)\hat{g}(\mathbf{c}, \mathbf{s}, i))P(\hat{\mathbf{s}}, t)\ \mathrm{d}\hat{\mathbf{s}}, \tag{6.94}$$

$$= D_s \frac{\langle s_{i+1} \rangle - 2\langle s_i \rangle + \langle s_{i-1} \rangle}{(\Delta x)^2} - \lambda_s \sum_{\mathbf{c}} C_1(c_1, t) \cdots C_N(c_N, t)\langle s_i \rangle \sum_{j=1}^{N} e^{-(i - c_j)^2/\delta}, \tag{6.95}$$

$$= D_s \frac{\langle s_{i+1} \rangle - 2\langle s_i \rangle + \langle s_{i-1} \rangle}{(\Delta x)^2} - \lambda_s \sum_{j=1}^{N} \prod_{k \neq j} \sum_{c_k} C_k(c_k, t)\langle s_i \rangle \sum_{c_j} C_j(c_j, t) e^{-(i - c_j)^2/\delta}. \tag{6.96}$$

For $k \neq j$ $\sum_{c_k} C_k(c_k, t) = 1$, so

$$\frac{\partial \langle s_i \rangle}{\partial t} = D_s \frac{\langle s_{i+1} \rangle - 2\langle s_i \rangle + \langle s_{i-1} \rangle}{(\Delta x)^2} - \lambda_s \sum_{j=1}^{N} \sum_{l=1}^{L} \hat{C}_j(l\Delta x, t) \langle s_i \rangle e^{-(i-l)^2/\delta}, \tag{6.97}$$

$$= D_s \frac{\langle s_{i+1} \rangle - 2\langle s_i \rangle + \langle s_{i-1} \rangle}{(\Delta x)^2} - \lambda_s \sum_{l=1}^{L} C(l\Delta x, t) \langle s_i \rangle e^{-(i-l)^2/\delta}, \tag{6.98}$$

and as $\delta \to 0$ and $\Delta x \to 0$ with $d_s = D_s/(\Delta x)^2$ constant

$$e^{-(i-l)^2/\delta} \to \begin{cases} 0 & l \neq i \\ 1 & l = i \end{cases}, \tag{6.99}$$

$$= \delta_{il}, \tag{6.100}$$

the Kronecker delta function. Hence

$$\frac{\partial \hat{s}}{\partial t} = d_s \frac{\partial^2 \hat{s}}{\partial x^2} - \lambda_s \hat{s} C, \tag{6.101}$$

for $\hat{s}(k\Delta x, t) = \langle s_k(t) \rangle$. Since $\mathbf{s} = s(\mathbf{x}, t)$, the cell evolution equations are exactly as before (see Section 6.1) and so the complete system remains unchanged from the non-hybrid case with

$$\frac{\partial \hat{s}}{\partial t} = d_s \frac{\partial^2 \hat{s}}{\partial x^2} - \lambda_s \hat{s} C, \tag{6.102}$$

$$\frac{\partial C}{\partial t} = \hat{\alpha} \frac{\partial}{\partial x} \left(\hat{s} \frac{\partial C}{\partial x} - C \frac{\partial \hat{s}}{\partial x} \right), \tag{6.103}$$

for non-local sensing and

$$\frac{\partial \hat{s}}{\partial t} = d_s \frac{\partial^2 \hat{s}}{\partial x^2} - \lambda_s \hat{s} C, \tag{6.104}$$

$$\frac{\partial C}{\partial t} = \hat{\alpha} \frac{\partial^2}{\partial x^2} \left(\hat{s} C \right), \tag{6.105}$$

for local sensing. Simulations of this system, averaged over multiple realisations to obtain a histogram of cell positions and the average chemical density at each lattice point, compare well with numerical solutions to equations (6.102) and (6.103) (Figure 6.3(b)) and equations (6.104) and (6.105)) (Figure 6.3(a)) for a range of parameter values, including changes in δ (results not shown). The addition of volume exclusion modifies equations (6.103) and (6.105) to give equations (6.49) and (6.57), respectively, derived in the non-hybrid system (Section 6.1). As in the on-lattice non-hybrid system

(Figure 6.2), the addition of volume exclusion makes very little difference to either the simulations or numerical solutions to the equations within the domain (Figures 6.3(c) and 6.3(d)).

6.3 Model 3: off-lattice cells with continuum chemoattractant

Off-lattice IBMs often give different results than the equivalent on-lattice models (Plank and Simpson, 2012). To investigate whether an off-lattice version of our hybrid model from Section 6.2 behaves similarly to the on-lattice version, and work towards a method that can be used to analyse the models from Chapters 2 and 3, we will extend Section 6.2 to include off-lattice cells. Cells in our model move a distance $d \ll 1$ at a rate $T^{\pm}(s,x) = \alpha s(x \pm d)$ left or right from a position x. The rates $T^{\pm}(s,x)$ depend on the chemoattractant concentration, s which evolves according to

$$\frac{\partial s}{\partial t} = g\left(\mathbf{c}, s, \frac{\partial^2 s}{\partial x^2}, x\right). \tag{6.106}$$

We begin with a model without volume exclusion. We take the same chemical evolution equation as our on-lattice model (Section 6.2), that is,

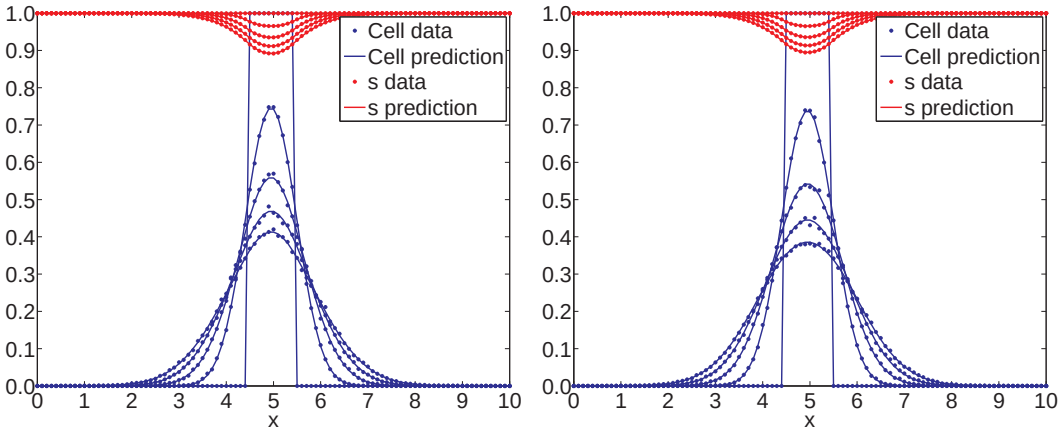
$$g\left(\mathbf{c}, s, \frac{\partial^2 s}{\partial x^2}, x\right) = D_s \frac{\partial^2 s}{\partial x^2} - \lambda_s s \sum_{j=1}^{N} e^{-(x-c_j)^2/\delta}. \tag{6.107}$$

6.3.1 Derivation of continuum equations

To analyse the off-lattice model, we consider the pdf, $P(s,x,t)$, for the chemoattractant concentration at position, x, and time, t. Then the pdf, $C_k(x,t)$, for the position x, of cell k at time t, satisfies

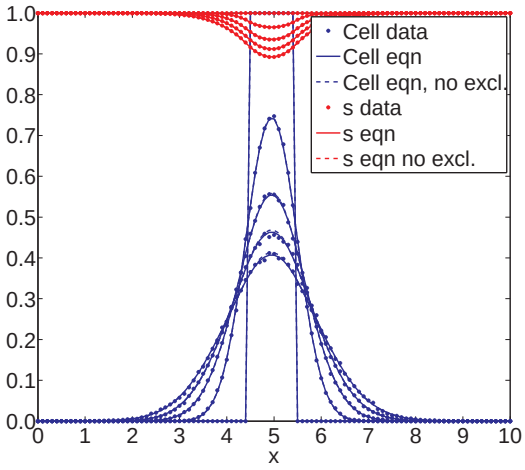
$$\begin{aligned} C_k(x,t+\Delta t) = {} & \overbrace{C_k(x,t)\left(1-\alpha\Delta t\left(\int_s sP(s,x+d,t) + sP(x,x-d,t)\ \mathrm{d}s\right)\right)}^{\text{cell } k \text{ at } x \text{ stays at } x} \\ & + \underbrace{\alpha\Delta t \int_s sP(s,x,t)\ \mathrm{d}s\ C_k(x-d,t)}_{\text{cell } k \text{ at } x-d \text{ moves to } x} + \underbrace{\alpha\Delta t \int_s sP(s,x,t)\ \mathrm{d}s\ C_k(x+d,t)}_{\text{cell } k \text{ at } x+d \text{ moves to } x}, \end{aligned} \tag{6.108}$$

$$\begin{aligned} = {} & C_k(x,t)(1-\alpha\Delta t(\langle s(x+d,t)\rangle + \langle s(x-d,t)\rangle)) + \alpha\Delta t\langle s(x,t)\rangle C_k(x-d,t) \\ & + \alpha\Delta t\langle s(x,t)\rangle C_k(x+d,t), \end{aligned} \tag{6.109}$$

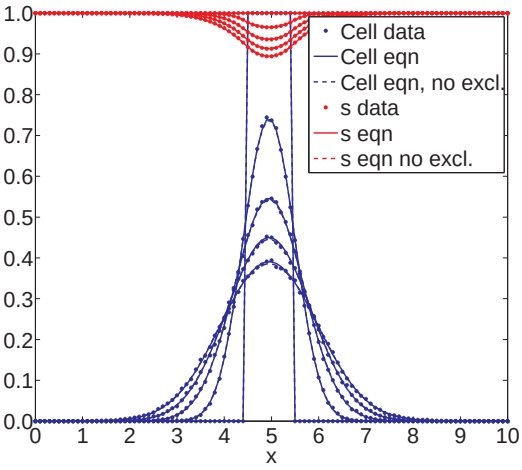


(a) Local sensing

(b) Non-local sensing



(c) Local sensing with volume exclusion



(d) Non-local sensing with volume exclusion

Figure 6.3: A comparison of typical averaged simulations of the on-lattice hybrid chemotaxis models described in Section 6.2 and numerical solutions to the derived equations for (a) local sensing (equations (6.104) and (6.105)); (b) non-local sensing (equations (6.102) and (6.103))); (c) local sensing with volume exclusion (equations (6.49) and (6.104)) and (d) non-local sensing with volume exclusion (equations (6.57) and (6.104)). Each simulation is performed with $\alpha = 0.5$, $\Delta t = 0.5$, $\Delta x = 0.1$, $D_s = 0$, $\lambda_s = 0.02$ and is averaged over $10,000$ runs. Ten cells are initially placed, one at each lattice site, in the middle of the domain and no flux conditions are taken at each boundary. Solutions are plotted at $t = 10$.

where $\langle s(x,t)\rangle = \int_s sP(s,x,t)\ \mathrm{d}s$ and we have used the moment closure approximation

$$\int_{\mathbf{s}} s\text{Prob}(\text{cell } k \text{ at } (x,t) \text{ and chemoattractant is } s \text{ at } (x,t))\ \mathrm{d}s = \int_{\mathbf{s}} \text{Prob}\left(\sum_l C_k^l(x,t), s\right)\ \mathrm{d}s, \tag{6.110}$$

$$\approx \langle C_k^l(x,t)\rangle\langle s(x,t)\rangle, \tag{6.111}$$

$$= C_k(x,t)\int_{\mathbf{s}} sP(s,x,t)\ \mathrm{d}s, \tag{6.112}$$

where $C_k^l(x,t)$ is the probability that cell, k, is at position, x, at time, t, during simulation l. Now if $\hat{s}(x,t) = \langle s(x,t)\rangle$ and we Taylor expand in terms of $d \ll 1$,

$$\frac{\partial C_k}{\partial t} = \alpha d^2 \frac{\partial}{\partial x}\left(\hat{s}\frac{\partial C_k}{\partial x} - C_k\frac{\partial \hat{s}}{\partial x}\right) + \mathcal{O}(d^3), \tag{6.113}$$

as $\Delta t \to 0$.

To find $P(s,x,t)$ we use a similar technique to Section 6.2.3, but cell movement rates now require the chemoattractant concentration at arbitrary points in the domain, since the cells are no longer only present at lattice sites. We therefore use $\partial^2 s/\partial x^2$ instead of a finite difference approximation to the second derivative. Then

$$P(s,x,t+\Delta t) = \int_{\mathbf{c}}\int_y P(g^{-1}(\mathbf{c},s,y,x,t),x,t)\text{Prob}\left(\frac{\partial^2 s}{\partial x^2} = y\right)\left|\det\left(\frac{\partial \mathbf{s}}{\partial g(\mathbf{c},s,y,x,t)}\right)\right|\ \mathrm{d}y\ \mathrm{d}\mathbf{c}, \tag{6.114}$$

and hence

$$\int_s sP(s,x,t+\Delta t)\ \mathrm{d}s$$

$$= \int_s\int_{\mathbf{c}}\int_y sP(g^{-1}(\mathbf{c},s,y,x,t),x,t)\text{Prob}\left(\frac{\partial^2 s}{\partial x^2} = y\right)\left|\det\left(\frac{\partial \mathbf{s}}{\partial g(\mathbf{c},s,y,x,t)}\right)\right|\ \mathrm{d}y\ \mathrm{d}\mathbf{c}\ \mathrm{d}s, \tag{6.115}$$

$$= \int_{\hat{s}}\int_{\mathbf{c}}\int_y g(\mathbf{c},\hat{s},y,x,t)P(\hat{s},x,t)\text{Prob}\left(\frac{\partial^2 s}{\partial x^2} = y\right)\ \mathrm{d}y\ \mathrm{d}\mathbf{c}\ \mathrm{d}\hat{s}, \tag{6.116}$$

making a change of variables as before. Then

$$\langle s(x,t+\Delta t)\rangle = D_s \int_y y \text{ Prob}\left(\frac{\partial^2 s}{\partial x^2} = y\right) \, \mathrm{d}y$$
$$- \lambda_s \int_{\hat{s}} \hat{s} P(\hat{s},x,t) \, \mathrm{d}\hat{s} \int_y \sum_{j=1}^{N} C_j(y,t) e^{-(x-y)^2/\delta} \, \mathrm{d}y, \tag{6.117}$$
$$= D_s \langle s_{xx}(x,t)\rangle - \lambda_s \langle s(x,t)\rangle \int_y C(y,t) e^{-(x-y)^2/\delta} \, \mathrm{d}y, \tag{6.118}$$

for $C = \sum_j C_j$. Assuming $\langle s_{xx}(x,t)\rangle = \partial^2\langle s\rangle/\partial x^2 = \partial^2 \hat{s}/\partial x^2$

$$\frac{\partial \hat{s}}{\partial t} = D_s \frac{\partial^2 \hat{s}}{\partial x^2} - \lambda_s \hat{s} \int_y C(y,t) e^{-(x-y)^2/\delta} \, \mathrm{d}y. \tag{6.119}$$

For $\delta \ll 1$ we use Laplace's method,

$$\int_y C(y,t) e^{-(x-y)^2/\delta} \, \mathrm{d}y$$
$$= \int_y e^{-(x-y)^2/\delta} \left(C(x,t) + C_x(x,t)(y-x) + \frac{1}{2} C_{xx}(x,t)(y-x)^2 + \mathcal{O}((y-x)^3) \right) \, \mathrm{d}y, \tag{6.120}$$

then, extending the integral to $y \in (-\infty, \infty)$ (introducing at most exponentially small errors),

$$\int_y C(y,t) e^{-(x-y)^2/\delta} \, \mathrm{d}y = C(x,t) \int_{-\infty}^{\infty} e^{-(x-y)^2/\delta} \, \mathrm{d}y + \sum_{n=1}^{\infty} C^{(n)}(x,t) \int_{-\infty}^{\infty} (x-y)^n e^{-(x-y)^2/\delta} \, \mathrm{d}y, \tag{6.121}$$
$$= \sqrt{\delta\pi} C(x,t) + \sum_{n \text{ even}} \sqrt{2\pi} \left(\frac{\delta}{2}\right)^{\frac{n+1}{2}} (n-1)\cdots(3)(1) C^{(n)}(x,t), \tag{6.122}$$
$$= \sqrt{\delta\pi} C(x,t) + \mathcal{O}\left(\delta^{\frac{3}{2}}\right). \tag{6.123}$$

Hence as $\delta \to 0$ we take $\lambda_s \to \infty$ with $\hat{\lambda}_s = \sqrt{\delta\pi}\lambda_s$ constant so that summing equation (6.113) over $k = 1, \ldots, N$ and taking $d \to 0$:

$$\frac{\partial \hat{s}}{\partial t} = D_s \frac{\partial^2 \hat{s}}{\partial x^2} - \hat{\lambda}_s \hat{s} C(x,t), \tag{6.124}$$
$$\frac{\partial C}{\partial t} = \hat{\alpha} \frac{\partial}{\partial x} \left(\hat{s} \frac{\partial C}{\partial x} - C \frac{\partial \hat{s}}{\partial x} \right), \tag{6.125}$$

where $\hat{\alpha} = \lim_{d\to 0, \alpha \to \infty} \alpha d^2$. Thus we find that moving to an off-lattice setting for the individual-based cells, using an approximation to the continuous delta function in place of the previously taken approximation to the Kronecker delta, results in an equivalent system of continuum equations (equa-

tions (6.124) and (6.125)) to that for on-lattice cells moving via non-local sensing (equations (6.38) and (6.39)). The lattice spacing, Δx, is replaced by the distance, d, moved in a single jump, since when individuals move a constant distance at each jump they remain a distance di away from their initial position, where i is an integer. Hence each individual is effectively on a lattice, defined by their initial position, with spacing d. Since we do not consider volume exclusion, it makes no difference that each individual moves on an independent lattice, as there is no direct interaction between the individuals. Thus the only change due to moving to an off-lattice setting is that the consumption term in equation (6.124) is scaled by $\sqrt{\delta\pi}$. This difference in scaling is due to the difference between the approximation to the Kronecker delta, δ_{ij}, and the approximation we take to the delta function $\delta(x-y)$. Hence to regain an equivalent averaged system to equation (6.102) we need to take

$$g\left(\mathbf{c}, s, \frac{\partial^2 s}{\partial x^2}, x\right) = D_s \frac{\partial^2 s}{\partial x^2} - \frac{\lambda_s}{\sqrt{\delta\pi}} s \sum_{j=1}^{N} e^{-(x-c_j)^2/\delta}. \tag{6.126}$$

If instead we consider local sensing, $T_x^{\pm} = \alpha s(x)$, (with equation (6.107)) then equation (6.119) remains the same, whilst

$$\begin{aligned} C_k(x, t+\Delta t) = {} & \overbrace{C_k(x,t)\left(1 - 2\alpha\Delta t \int_s sP(s,x,t)\ \mathrm{d}s\right)}^{\text{cell } k \text{ at } x \text{ stays at } x} \\ & + \underbrace{\alpha\Delta t \int_s sP(s, x-d, t)\ \mathrm{d}s C_k(x-d,t)}_{\text{cell } k \text{ at } x-d \text{ moves to } x} + \underbrace{\alpha\Delta t \int_s sP(s, x+d, t)\ \mathrm{d}s C_k(x+d,t)}_{\text{cell } k \text{ at } x+d \text{ moves to } x}, \end{aligned} \tag{6.127}$$

$$\begin{aligned} = {} & C_k(x,t)(1 - 2\alpha\Delta t\langle s(x,t)\rangle) + \alpha\Delta t\langle s(x-d,t)\rangle C_k(x-d,t) \\ & + \alpha\Delta t\langle s(x+d,t)\rangle C_k(x+d,t), \end{aligned} \tag{6.128}$$

where $\langle s(x,t)\rangle = \int_s sP(s,x,t)\ \mathrm{d}s$. Now setting $\hat{s}(x,t) = \langle s(x,t)\rangle$, and expanding in terms of d,

$$\frac{\partial C_k}{\partial t} = \alpha d^2 \frac{\partial^2}{\partial x^2}(\hat{s}C_k) + \mathcal{O}(d^3), \tag{6.129}$$

as $\Delta t \to 0$. Summing over k and taking $d \to 0$ with $\hat{\alpha} = \lim_{d\to 0, \alpha\to\infty} \alpha d^2$ held constant, gives the

full system

$$\frac{\partial \hat{s}}{\partial t} = D_s \frac{\partial^2 \hat{s}}{\partial x^2} - \hat{\lambda}_s \hat{s} C(x,t), \tag{6.130}$$

$$\frac{\partial C}{\partial t} = \hat{\alpha} \frac{\partial^2}{\partial x^2}(\hat{s} C), \tag{6.131}$$

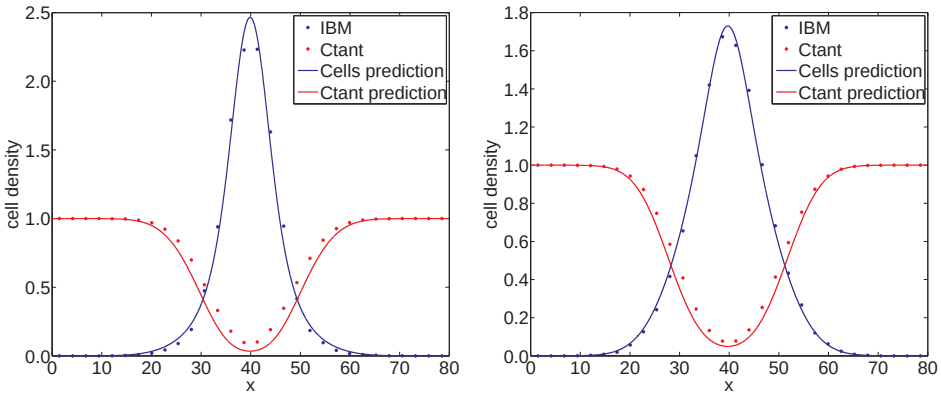
where $C(x,t) = \sum_{k=1}^{N} C_k(x,t)$, as before. Here again we find that moving to an off-lattice setting makes very little difference to the cell equation, merely replacing the lattice spacing, Δx, with the distance moved, d. Simulations of the hybrid system averaged over many repetitions compare well with numerical solutions to the systems derived, with both local (equations (6.130) and (6.131)) and non-local (equations (6.124) and (6.125)) sensing (Figures 6.4(a) and 6.4(b)). Hence we conclude that moving to an off-lattice hybrid system in this manner does not affect the cell concentration equations, only the chemical equation.

We now investigate the effect of changing the consumption parameter, δ, whilst scaling λ_s in the simulations resulting in constant $\hat{\lambda}_s$ so that the PDEs to be solved are the same for all values of δ. Unlike the on-lattice hybrid system, however, changing δ has a significant effect on the chemical density profile (Figure 6.5). Surprisingly, our predictions are less good as δ is reduced, despite taking $\delta \to 0$ in our derivation. There are two possible reasons for this. Firstly, highly localised consumption terms may be missed in the hybrid simulations if they do not have any impact on the chemical on the lattice used for numerical solutions. Secondly, less localised consumption terms reduce the dependence on the exact location of each cell. Hence the stochastic effects are reduced and thus the deterministic equations derived may be more exact.

Volume exclusion is known to affect off-lattice systems differently to on-lattice systems, for example in the case of unbiased motion, where volume exclusion does not alter the derived diffusion coefficient in the on-lattice setting. We would expect, therefore, that volume exclusion would also affect the off-lattice hybrid system differently than the on-lattice hybrid system and will consider next the effects of volume exclusion on the off-lattice hybrid system.

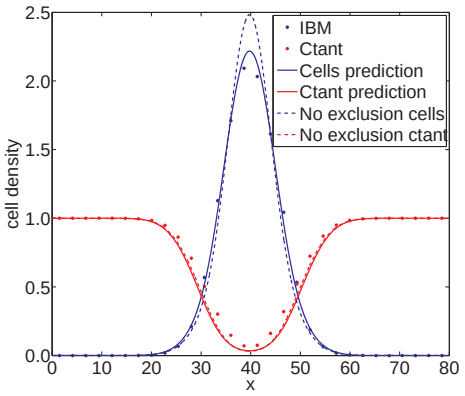
6.3.2 Excluded volume effects

The off-lattice cell models in this section may be extended to incorporate volume exclusion using similar methods to Chapter 4. In particular, the modification to cell movement due to volume exclusion does not affect the chemoattractant evolution equation (equation (6.119)) since the derivation

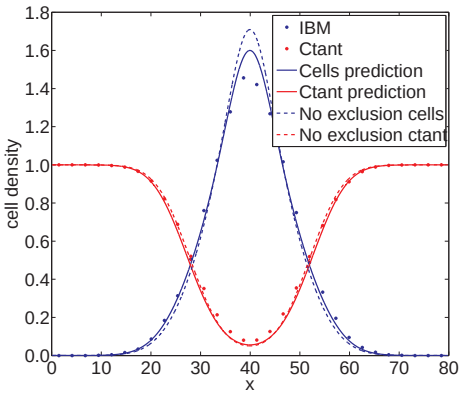


(a) Local sensing without exclusion

(b) Non-local sensing without exclusion



(c) Local sensing with exclusion



(d) Non-local sensing with exclusion

Figure 6.4: A comparison of typical averaged simulations of the off-lattice hybrid chemotaxis models described in Section 6.3 and numerical solutions to the derived equations for (a) local sensing (equations (6.130) and (6.131)); (b) non-local sensing (equations (6.124) and (6.125)); (c) local sensing with volume exclusion (equations (6.130) and (6.140)) and (d) non-local sensing with volume exclusion (equations (6.130) and (6.149)). Each simulation is performed with $\alpha = 1.25$, $\Delta t = 1$, $D_s = 0.0706$, $\lambda_s = 0.01$, $\delta = 0.01$, $d = 0.2$, $R = 0.15$ and is averaged over $1,000$ runs. Thirty cells are initially placed at regular intervals in the middle of the domain with the leftmost cell centre drawn from the normal distribution at $x \sim \mathcal{N}((B_R - B_L)/2 - NR/0.8, 10)$ and no flux conditions are taken at each boundary. Solutions are plotted at $t = 200$.

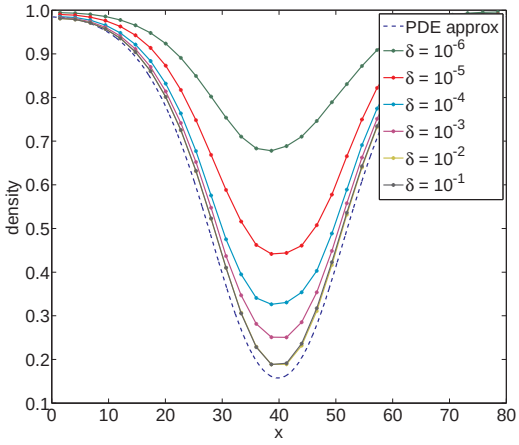
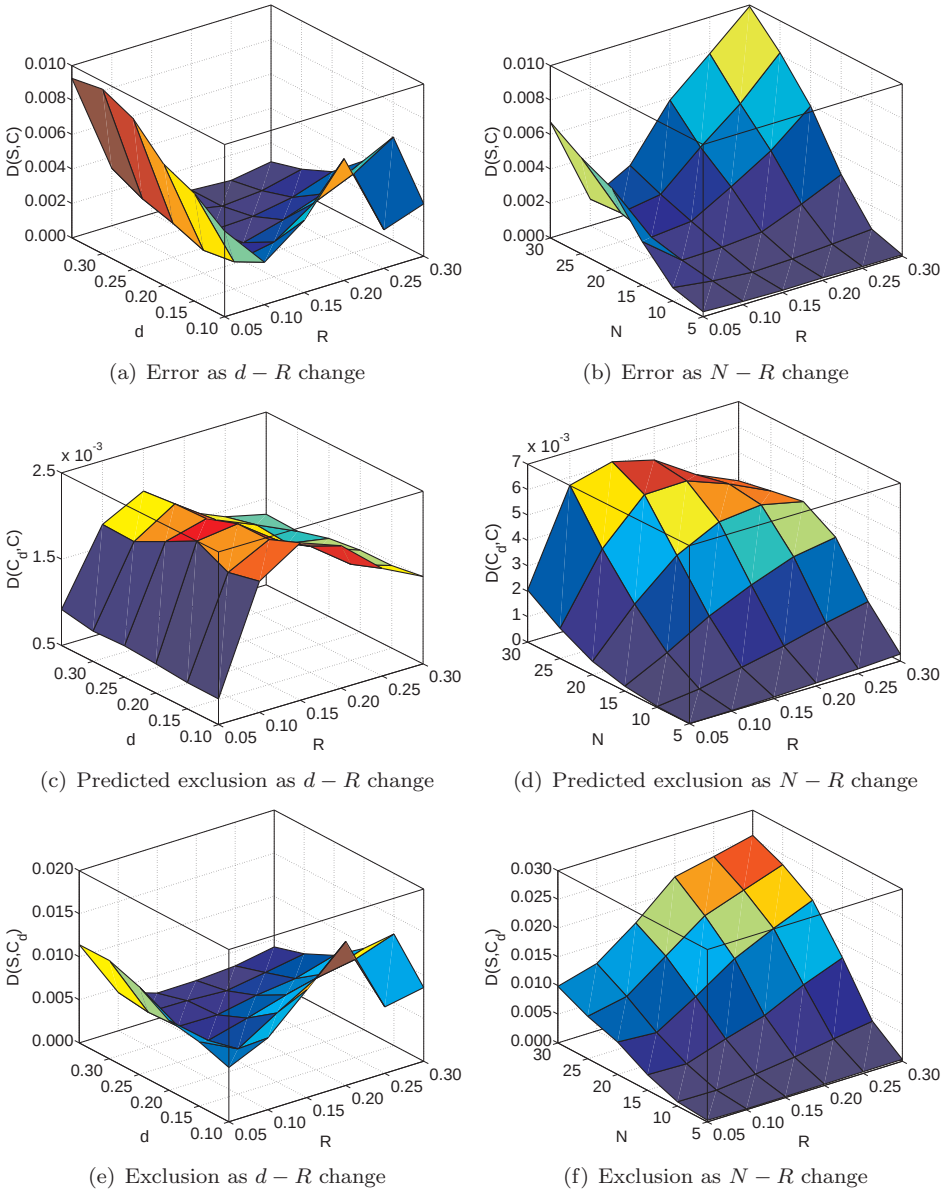


Figure 6.5: An investigation of the effects of varying δ in the chemical consumption term. Each simulation is performed with $\alpha = 0.5$, $\Delta t = 1$, $D_s = 0.0706$, $d = 0.15$, $R = 0.15$ and is averaged over $1,000$ runs. Thirty cells are initially placed at regular intervals in the middle of the domain with the left most cell centre drawn from the normal distribution at $x \sim \mathcal{N}((B_R - B_L)/2 - NR/0.8, 10)$ and no flux conditions are taken at each boundary. Solutions are plotted at $t = 200$. λ_s is taken to be $0.01/\sqrt{\delta\pi}$, so that $\hat{\lambda}_s$ in equation (6.124) remains constant. We therefore plot the numerical solution to equation (6.124), and simulations as δ changes.

of equation (6.119) does not consider cell movement. However, equation (6.108) is modified if each cell occupies a distance R from its centre. For local sensing ($T^{\pm}(s,x) = \alpha s(x)$), we have

$$
\begin{aligned}
C_k(x,t+\Delta t) = \\
&\overbrace{C_k(x,t)}^{\text{cell } k \text{ at } x}\Bigg(\overbrace{1-\alpha\Delta t\left(\int_s sP(s,x+d,t)+sP(x,x-d,t)\ \mathrm{d}s\right)}^{\text{does not try to move}} \\
&+\underbrace{\alpha\Delta t\left(\int_s sP(s,x+d,t)P_R^k(x,t)+sP(x,x-d,t)P_L^k(x,t)\ \mathrm{d}s\right)}_{\text{tries to move and fails}}\Bigg) \\
&+\underbrace{\alpha\Delta t\int_s sP(s,x,t)\ \mathrm{d}s\ C_k(x-d,t)(1-P_R^k(x-d,t))}_{\text{cell } k \text{ at } x-d \text{ moves to } x} \\
&+\underbrace{\alpha\Delta t\int_s sP(s,x,t)\ \mathrm{d}s\ C_k(x+d,t)(1-P_L^k(x+d,t))}_{\text{cell } k \text{ at } x+d \text{ moves to } x}, \qquad (6.132)\\
&= C_k(x,t)(1-\alpha\Delta t(\langle s(x+d,t)\rangle P_R^k(x,t)+\langle s(x-d,t)\rangle P_L^k(x,t)) \\
&\quad+\alpha\Delta t\left(\langle s(x+d,t)\rangle P_R^k(x,t)+\langle s(x-d,t)\rangle P_L^k(x,t)\right) \\
&\quad+\alpha\Delta t\langle s(x,t)\rangle C_k(x-d,t)(1-P_R^k(x-d,t))+\alpha\Delta t\langle s(x,t)\rangle C_k(x+d,t)(1-P_L^k(x+d,t)), \qquad (6.133)
\end{aligned}
$$



(a) Error as $d - R$ change

(b) Error as $N - R$ change

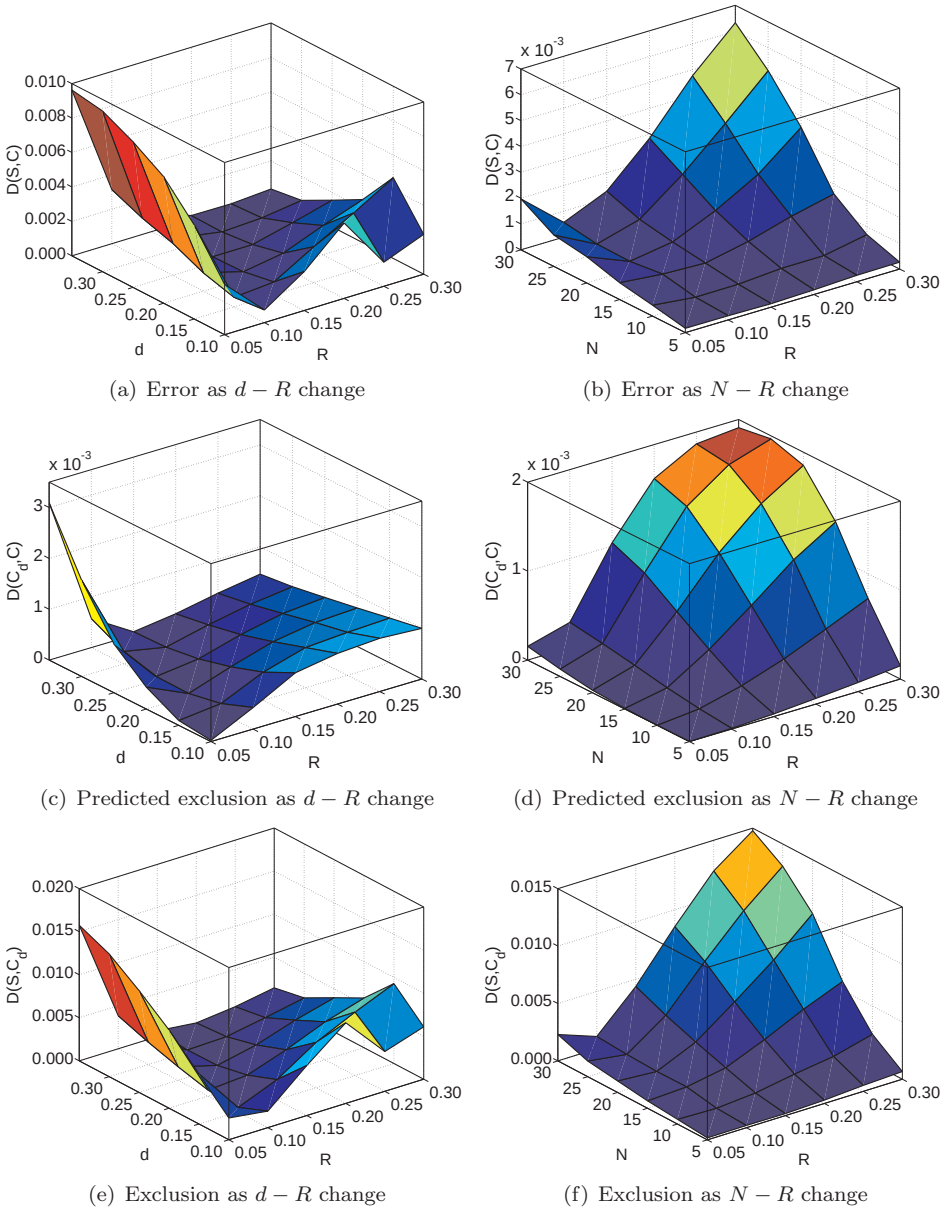
(c) Predicted exclusion as $d - R$ change

(d) Predicted exclusion as $N - R$ change

(e) Exclusion as $d - R$ change

(f) Exclusion as $N - R$ change

Figure 6.6: Exploring the effect of changing the likelihood of successful movement in Model 3 with local chemotactic sensing by changing the distance moved, d, and the radius of a cell, R ((a), (c) and (e)) and changing the occupied space in the model by changing the number of cells, N, and the radius of a cell, R ((b), (d) and (f)). We consider the effect on: (a)-(b) the error introduced by our approximations (the difference between the solution to equations (6.130) and (6.140) and the average of our simulations); (c)-(d) our prediction of the effect of volume exclusion (the difference between the solution to equations (6.130) and (6.140) and equations (6.130) and (6.131); and (e)-(f) the actual effect of volume exclusion (the difference between the average of our simulations and the solution to equations (6.130) and (6.131)). Differences are calculated using equation (4.46) and non-varying parameters are: $R = 0.05$, $d = 0.1$, $\Delta t = 1$, $\Delta x = 8/3$, $D_s = 0.07$, $\lambda_s = 0.01$, $\delta = 0.1$, whilst $\alpha = 0.05/d^2$ and simulations are averaged over $1,000$ runs. Thirty cells are initially placed at regular intervals in the middle of the domain with the left most cell centre drawn from the normal distribution at $x \sim \mathcal{N}((B_R - B_L)/2 - NR/0.8, 10)$ and no flux conditions are taken at each boundary.



(a) Error as $d - R$ change

(b) Error as $N - R$ change

(c) Predicted exclusion as $d - R$ change

(d) Predicted exclusion as $N - R$ change

(e) Exclusion as $d - R$ change

(f) Exclusion as $N - R$ change

Figure 6.7: Exploring the effect of changing the likelihood of successful movement in Model 3 with non-local chemotactic sensing by changing the distance moved, d, and the radius of a cell, R ((a), (c) and (e)) and changing the occupied space in the model by changing the number of cells, N, and the radius of a cell, R ((b), (d) and (f)). We consider the effect on: (a)-(b) the error introduced by our approximations (the difference between the solution to equations (6.124) and (6.149) and the average of our simulations); (c)-(d) our prediction of the effect of volume exclusion (the difference between the solution to equations (6.124) and (6.149) and equations (6.124) and (6.125); and (e)-(f) the actual effect of volume exclusion (the difference between the average of our simulations and the solution to equations (6.124) and (6.125)). Differences are calculated using equation (4.46) and non-varying parameters are: $R = 0.05$, $\Delta t = 1$, $\Delta x = 8/3$, $D_s = 0.07$, $\lambda_s = 0.01$, $\delta = 0.1$, whilst $\alpha = 0.05/d^2$ and simulations are averaged over $1,000$ runs. Thirty cells are initially placed at regular intervals in the middle of the domain with the left most cell centre drawn from the normal distribution at $x \sim \mathcal{N}((B_R - B_L)/2 - NR/0.8, 10)$ and no flux conditions are taken at each boundary.

where $P_L^k(x,t)$ and $P_R^k(x,t)$ are the probabilities of a cell, other than cell k, being present in the regions $(x-d,x]$ and $[x,x+d)$, respectively, given that cell k is at x at time t. The moment closure assumption has been taken as before (see equation (6.111)). We have

$$P_R^k(x,t) = \sum_{i\neq k}\int_{2R}^{2R+d} C_i(x+\bar{x},t)\,\mathrm{d}\bar{x}, \tag{6.134}$$

$$= \sum_{i\neq k}\left(dC_i + \frac{d}{2}(4R+d)\frac{\partial C_i}{\partial x} + \frac{d}{6}(12R^2+6Rd+d^3)\frac{\partial^2 C_i}{\partial x^2} + \mathcal{O}((2R+d)^4)\right), \tag{6.135}$$

$$P_L^k(x,t) = \sum_{i\neq k}\int_{-2R-d}^{-2R} C_i(x+\bar{x},t)\,\mathrm{d}\bar{x}, \tag{6.136}$$

$$= \sum_{i\neq k}\left(dC_i - \frac{d}{2}(4R+d)\frac{\partial C_i}{\partial x} + \frac{d}{6}(12R^2+6Rd+d^3)\frac{\partial^2 C_i}{\partial x^2} + \mathcal{O}((2R+d)^4)\right), \tag{6.137}$$

assuming $2R+d$ is small compared to the length scale on which C changes so that we can Taylor expand $P_L^k(x,t)$ and $P_R^k(x,t)$. Substituting equations (6.135) and (6.137) into equation (6.133) and Taylor expanding in d:

$$\frac{\partial C_k}{\partial t} = \alpha d^2\frac{\partial^2}{\partial x^2}\left(\hat{s}C_k\right) + 4\alpha R d^2\frac{\partial}{\partial x}\left(\hat{s}C_k\sum_{i\neq k}\frac{\partial C_i}{\partial x}\right) - \alpha d^3\frac{\partial}{\partial x}\left(\frac{\partial}{\partial x}\left(\hat{s}C_k\right)\sum_{i\neq k}C_i\right)$$
$$- 2\alpha R^2 d^3\frac{\partial^2}{\partial x^2}\left(\hat{s}C_k\right)\sum_{i\neq k}C_i + \mathcal{O}\left((R+d)^4\right), \tag{6.138}$$

as $\Delta t \to 0$. As before, if the initial condition for all cells is the same, then $C_i = C_j\ \forall i,j$ so $C = \sum_{k=1}^N C_k = NC_k$ and

$$\frac{\partial C}{\partial t} = \alpha d^2\frac{\partial}{\partial x}\left(\left(1-\frac{N-1}{N}dC\right)\frac{\partial}{\partial x}\left(sC\right) + 4R\frac{N-1}{N}sC\frac{\partial C}{\partial x}\right). \tag{6.139}$$

As $d\to 0$ with $\alpha\to\infty$ so that $\hat{\alpha} = \alpha d^2$ remains constant,

$$\frac{\partial C}{\partial t} = \hat{\alpha}\frac{\partial}{\partial x}\left(\frac{\partial}{\partial x}(\hat{s}C) + 4R\frac{N-1}{N}\hat{s}C\frac{\partial C}{\partial x}\right). \tag{6.140}$$

Incorporating excluded volume into the off-lattice model results in a modified form (equation (6.140)) of the previously derived equation (equation (6.131)), with an additional radius-dependent term. Although without volume exclusion the on- and off-lattice models resulted in the same equations with Δx replaced by d (equations (6.28) and (6.131)), when excluded volume is included, the form of the equations is different. This is due to modelling differences, since in the on-lattice model volume exclusion allows at most one individual at each lattice site. Hence the lattice spacing is the same as

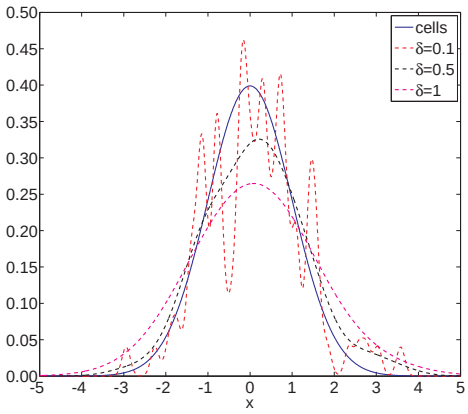


Figure 6.8: Different values of δ result in different consumption profiles from a population of 100 normally distributed cells.

the radius of an individual and letting $\Delta x \to 0$ also lets the radius tend to zero, affecting the volume exclusion. Thus there cannot be any dependence on the cell radius in the final equation on-lattice as there is in the off-lattice equation, where the radius of an individual is independent of the distance moved during a jump.

In addition, however, the form of the equation is changed in an off-lattice formulation. In the on-lattice model, volume exclusion reduces advection down the concentration gradient of other cells (equation (6.49)), whereas off-lattice, equation (6.140) can be thought of as

$$\frac{\partial C}{\partial t} = \hat{\alpha}\frac{\partial}{\partial x}\left(\hat{s}\frac{\partial C}{\partial x}\left(1 + 4R\frac{N-1}{N}C\right) + \frac{\partial \hat{s}}{\partial x}C\right), \tag{6.141}$$

so that the advection term is unchanged, whilst the diffusion term has increased diffusivity due to the presence of other individuals. Thus volume exclusion still interferes with the efficiency of the advective flux, since the smoothing effects of diffusion can work against chemotactic directed motion, without directly influencing the advection term. Indeed, numerical solutions to equations (6.119) and (6.140), validated by averaged repetitions of the hybrid model (Figure 6.4(c)), are significantly different when compared to the equations without volume exclusion (equations (6.130) and (6.131)). Specifically, with volume exclusion the initial population spreads more rapidly, as we would expect from the increased diffusivity term in equation (6.141).

It is known that using consumption terms of the form $\exp(-(x-c_i)^2/\delta)$ can result in wider chemical consumption than the underlying cell density for large δ, whilst for small δ the consumption density oscillates widely (Figure 6.8; Franz and Erban, 2013). When including volume exclusion, however, we may include chemical consumption as a constant rate of consumption over the body of each cell.

Hence a cell at position x consumes chemoattractant at a rate λ in the segment $[x-R, x+R]$. Then equation (6.119) becomes

$$\frac{\partial \hat{s}}{\partial t} = D_s \frac{\partial^2 \hat{s}}{\partial x^2} - \lambda_s \hat{s} \int_{-R}^{R} C(x+y,t)\ \mathrm{d}y, \tag{6.142}$$

$$= D_s \frac{\partial^2 \hat{s}}{\partial x^2} - \lambda_s \hat{s} \int_{-R}^{R} \left(C(x,t) + yC_x(x,t) + \frac{y^2}{2} C_{xx}(x,t) + \mathcal{O}(y^3) \right)\ \mathrm{d}y, \tag{6.143}$$

$$= D_s \frac{\partial^2 \hat{s}}{\partial x^2} - \lambda_s \hat{s} 2RC + \mathcal{O}(R^2), \tag{6.144}$$

$$= D_s \frac{\partial^2 \hat{s}}{\partial x^2} - \hat{\lambda}_s \hat{s}, \tag{6.145}$$

for $\hat{\lambda}_s = 2R\lambda_s$ finite as $R \to 0$. Numerical solutions to equation (6.145) compare well to averaged simulations as R varies (Figure 6.9). As for the δ consumption term, the derived equations are less accurate for smaller values of R, but the reduction in accuracy is much less marked than in the δ case. To enable comparison with the earlier models without volume exclusion, we will continue to use the δ consumption term, whilst being careful to maintain the same value of δ compared to the lattice spacing.

We may measure the error in our approximations as the model parameters change by using equation (5.20) to find the difference between numerical solutions to equation (6.140) and a histogram derived from average positions of cells over multiple realisations of the model. For local sensing, our approximation is less accurate for lower values of R or d (Figure 6.6(a)). This is particularly surprising, since we would expect the Taylor expansions in our derivation to be better approximations for small values of d and R. However, for local sensing, small movement distances relative to the radius of a cell result in underprediction of population spreading whereas larger movement distances result in overprediction of population spreading. Since local sensing leads to the aggregation of individuals in areas of low chemical, we expect our moment closure approximation to be less good since the positions of different individuals are correlated. Cell aggregation is exacerbated by a smaller cell radius, R, leading to slower diffusion, and higher cell movement distances, d, which is reflected in a higher multiplier, $\hat{\alpha}$, for the advection term and may explain the relatively high error for small R and larger d.

The actual and predicted effects of volume exclusion can be measured similarly to Section 5.2.5 by finding the difference between numerical solutions to equation (6.131) and averaged simulations of the model or numerical solutions to equation (6.140), respectively. The error in our approximations affects the shape of the surface of the predicted effect of volume exclusion as R and d change

(Figure 6.6(c)) compared to the shape of the surface of the actual effect of volume exclusion (Figure 6.6(e)). We note, however, that for all but three of the parameter combinations simulated, the error in our approximations was lower than the error obtained by using the equations without volume exclusion. Hence including volume exclusion increases the accuracy of our predictions compared to naïvely ignoring the space occupied by cells. We expect from equation (6.140) that increasing the radius of a cell will have an impact on volume exclusion and, whilst the effect is less marked than is predicted in Figure 6.6(c), this is still generally found to be the case when comparing against averaged simulations of the model (Figure 6.6(e)).

When varying the number of individuals in the population and the radius of a cell, we expect more numerous populations and larger individuals will be more affected by incorporating volume exclusion. We make this prediction both because we would expect individuals to encounter each other more often and because, in equation (6.140), greater values of R and larger populations leading to a higher total C when integrated over the domain increase the volume exclusion term $4R(N-1)/NsC\partial C/\partial x$ (Figure 6.6(d)). Whilst the higher population numbers and larger cells also increase the error in our approximations (Figure 6.6(b)), our prediction of greater volume exclusion effects is also validated in general by averaged simulation data (Figure 6.6(f)).

With non-local sensing ($T^{\pm}(s,x) = \alpha s(x \pm d)$), equation (6.133) becomes

$$\begin{aligned} C_k(x,t+\Delta t) = \\ & C_k(x,t)(1-\alpha\Delta t\langle s(x,t)\rangle(P_R^k(x,t)+P_L^k(x,t)) + \alpha\Delta t\langle s(x,t)\rangle\left(P_R^k(x,t)+P_L^k(x,t)\right)) \\ & + \alpha\Delta t\langle s(x-d,t)\rangle C_k(x-d,t)(1-P_R^k(x-d,t)) + \alpha\Delta t\langle s(x+d,t)\rangle C_k(x+d,t)(1-P_L^k(x+d,t)), \end{aligned} \tag{6.146}$$

with $P_L^k(x,t)$ and $P_R^k(x,t)$ given by equations (6.134) and (6.136), as before. Hence, expanding in d and R,

$$\frac{\partial C_k}{\partial t} = \alpha d^2 \frac{\partial}{\partial x}\left(\hat{s}\frac{\partial C_k}{\partial x} - C_k\frac{\partial \hat{s}}{\partial x}\right) + 4\alpha R d^2 \frac{\partial}{\partial x}\left(\hat{s}C_k\sum_{i\neq k}\frac{\partial C_i}{\partial x}\right) + \mathcal{O}\left(d^3\right), \tag{6.147}$$

where $\hat{s}(x,t) = \langle s(x,t)\rangle$ as before and if the initial conditions are the same for all cells then $C = \sum_{k=1}^{N} C_k = NC_k$ and

$$\frac{\partial C}{\partial t} = \alpha d^2 \frac{\partial}{\partial x}\left(\left(1-\frac{N-1}{N}dC\right)\left(s\frac{\partial C}{\partial x} - C\frac{\partial s}{\partial x}\right) + 4R\frac{N-1}{N}sC\frac{\partial C}{\partial x}\right). \tag{6.148}$$

Then, as $d \to 0$ with $\hat{\alpha} = \lim_{d\to 0, \alpha \to \infty} \alpha d^2$ constant

$$\frac{\partial C}{\partial t} = \hat{\alpha} \frac{\partial}{\partial x} \left(\hat{s} \frac{\partial C}{\partial x} - C \frac{\partial \hat{s}}{\partial x} \right) + 4\hat{\alpha} R \frac{N-1}{N} \frac{\partial}{\partial x} \left(\hat{s} C \frac{\partial C}{\partial x} \right). \tag{6.149}$$

As with local sensing, volume exclusion introduces the same additional term to the original equation (equation (6.125)), corresponding to the same increase in diffusivity when considered as

$$\frac{\partial C}{\partial t} = \hat{\alpha} \frac{\partial}{\partial x} \left(\hat{s} \frac{\partial C}{\partial x} \left(1 + 4R \frac{N-1}{N} C \right) - C \frac{\partial \hat{s}}{\partial x} \right). \tag{6.150}$$

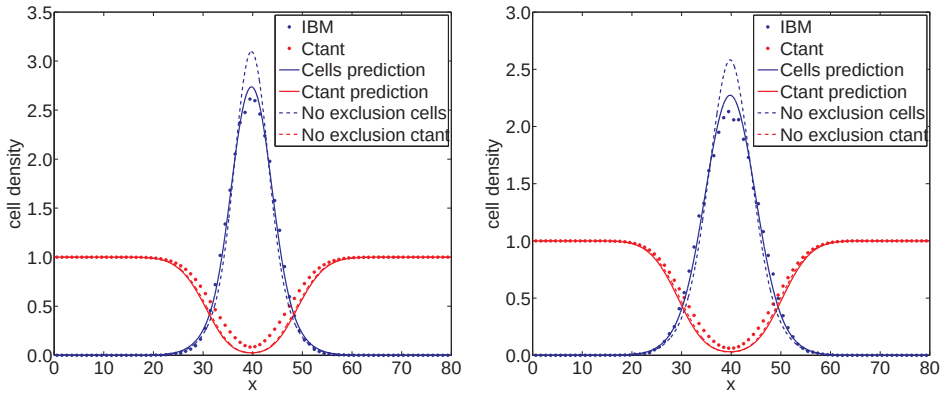
Due to the increased diffusivity, numerical solutions to equations (6.119) and (6.149), validated by averaged repetitions of the hybrid model, also display a faster spread of the initial population compared to equations (6.124) and (6.125), without volume exclusion (Figure 6.4(d)).

Hence volume exclusion in an off-lattice setting increases diffusivity, compared to reducing the advective flux in the on-lattice regime. In both on- and off-lattice formulations, volume exclusion modifies the equations in the equivalent ways for local and non-local sensing.

We find that the errors in our approximations and the predicted and actual volume exclusion effects as the number and size of cells change is very similar to those with local sensing (Figures 6.7(b), 6.7(d) and 6.7(f)), so that it does not seem to matter whether the system is chemoattractive or chemorepellent. Conversely, when changing R and d, whilst the volume-exclusion-effects surface has not changed significantly (Figure 6.7(e)); our predictions are more exact (Figure 6.7(a)) and thus our predicted-volume-exclusion effects surface (Figure 6.7(c)) is more similar to Figure 6.7(e). The slight reduction in errors in the non-local sensing model compared to the local sensing case may be explained by the change from chemotaxis down gradients in the chemical, leading to cell aggregation, to chemotaxis up chemical gradients thus spreading individuals away from the initial population of cells. We therefore expect errors due to our moment closure approximation to be reduced compared to predictions for the model with local sensing since the positions of individuals are less correlated in the model with non-local sensing.

6.4 Model 4: two spatial dimensions

To be able to compare these results with those in Chapters 2 and 3, we must extend our methods to two spatial dimensions. We achieve this by combining the coupled systems techniques used in this



(a) $R = 0.1$ (b) $R = 0.15$

(c) $R = 0.2$ (d) $R = 0.25$

Figure 6.9: An investigation of the effects of varying R when cells consume chemoattractant at a constant rate over their cell body. Each simulation is performed with $\alpha = 0.5$, $\Delta t = 1$, $D_s = 0.0706$, $d = 0.15$, $R = 0.15$ and is averaged over $1,000$ runs. Thirty cells are initially placed at regular intervals in the middle of the domain with the left most cell centre drawn from the normal distribution at $x \sim \mathcal{N}((B_R - B_L)/2 - NR/0.8, 10)$ and no flux conditions are taken at each boundary. Solutions are plotted at $t = 200$. λ_s is taken to be $0.01/\sqrt{\delta\pi}$, so that $\hat{\lambda}_s$ in equation (6.124) remains constant. We therefore plot the numerical solution to equation (6.124), and simulations as δ changes.

chapter, and the volume exclusion methods used in Chapter 5.

We use a similar model formulation to Section 6.3 with a continuum chemoattractant, $s(\mathbf{x},t)$, and individual-based cells that move a distance d in a direction θ at a rate $T(s,\theta,\mathbf{x},t)$ away from $\mathbf{x}=(x,y)$. Cells consume the chemoattractant, s, which diffuses with diffusivity D_s, so that the evolution of the chemoattractant is given by

$$\frac{\partial s}{\partial t} = g(\mathbf{c}, s, \nabla^2 s, \mathbf{x}), \tag{6.151}$$

$$= D_s \nabla^2 s - \lambda_s s \sum_{j=1}^{N} e^{-(|\mathbf{x}-\mathbf{c}_j|^2)/\delta}, \tag{6.152}$$

where $|\mathbf{x}-\mathbf{c}_j|^2$ is the square of the Euclidean distance between $\mathbf{x}$ and the position, $\mathbf{c}_j=(x_j,y_j)$, of cell j. This is analogous to equation (6.107) since the exponential term in equation (6.152) is a continuous approximation to a delta function, scaled so that cells consume chemoattractant with rate λ_s at their cell centre.

As in Section 6.3.1, we find the pdf $P(s,\mathbf{x},t)$ for the value of the chemoattractant, $s(\mathbf{x},t)$, at position, $\mathbf{x}=(x,y)$, and time, t. Then if $G(\mathbf{c},s,\nabla^2 s,\mathbf{x}) = s+\Delta t g(\mathbf{c},s,\nabla^2 s,\mathbf{x})$ so that $s(x,t+\Delta t) = G(\mathbf{c},s,\nabla^2 s,\mathbf{x})$ and

$$P(s,\mathbf{x},t+\Delta t) = \int_{\bar{\mathbf{c}}}\int_{f} \text{Prob(the chemoattractant was } G^{-1}(\mathbf{c},s,f,\mathbf{x}) \text{ and } \nabla^2 s = f)\ \mathrm{d}f\ \mathrm{d}\bar{\mathbf{c}}, \tag{6.153}$$

where $\int_{\bar{\mathbf{c}}} = \int_{\mathbf{c}_1}\cdots\int_{\mathbf{c}_N}$ and $\mathrm{d}\bar{\mathbf{c}} = \mathrm{d}\mathbf{c}_1\cdots\mathrm{d}\mathbf{c}_N$. Then to find the average chemoattractant profile, we multiply equation (6.153) by s and integrate over all s:

$$\int_s sP(s,\mathbf{x},t+\Delta t)\ \mathrm{d}s$$
$$= \int_s\int_{\bar{\mathbf{c}}}\int_f sP(G^{-1}(\mathbf{c},s,f,\mathbf{x}),\mathbf{x},t)\text{Prob}\left(\nabla^2 s = f\right)\left|\det\left(\frac{\partial s}{\partial G(\mathbf{c},s,f,\mathbf{x})}\right)\right|\ \mathrm{d}f\ \mathrm{d}\bar{\mathbf{c}}\ \mathrm{d}s, \tag{6.154}$$

taking the moment closure approximation

$$\int_{\bar{\mathbf{c}}}\int_f s\text{Prob(the chemoattractant was } G^{-1}(\mathbf{c},s,f,\mathbf{x}) \text{ and } \nabla^2 s = f)\ \mathrm{d}f\ \mathrm{d}\bar{\mathbf{c}}$$
$$\approx \int_s\int_{\bar{\mathbf{c}}}\int_f sP(G^{-1}(\mathbf{c},s,f,\mathbf{x}),\mathbf{x},t)\text{Prob}\left(\nabla^2 s = f\right)\left|\det\left(\frac{\partial s}{\partial G(\mathbf{c},s,f,\mathbf{x})}\right)\right|\ \mathrm{d}f\ \mathrm{d}\bar{\mathbf{c}}\ \mathrm{d}s. \tag{6.155}$$

Then

$$\int_s sP(s,\mathbf{x},t+\Delta t)\ \mathrm{d}s \int_{\hat{s}}\int_{\bar{\mathbf{c}}}\int_f G(\mathbf{c},s,f,\mathbf{x})P(\hat{s},\mathbf{x},t)\mathrm{Prob}\left(\nabla^2 s=f\right)\ \mathrm{d}f\ \mathrm{d}\bar{\mathbf{c}}\ \mathrm{d}\hat{s}, \tag{6.156}$$

by making the change of variables $\hat{s}=G^{-1}(\mathbf{c},s,f,\mathbf{x})$. Hence

$$\frac{\partial\langle s\rangle}{\partial t} = D_s\int_f f\mathrm{Prob}\left(\nabla^2 s=f\right)\ \mathrm{d}f - \lambda_s\int_{\hat{s}}\hat{s}P(\hat{s},\mathbf{x},t)\ \mathrm{d}\hat{s}\int_{\hat{\mathbf{x}}}\sum_{j=1}^{N}C_j(\hat{\mathbf{x}},t)e^{-|\mathbf{x}-\hat{\mathbf{x}}|^2/\delta}\ \mathrm{d}\hat{\mathbf{x}}, \tag{6.157}$$

$$= D_s\langle\nabla^2 s\rangle - \lambda_s\langle s\rangle\int_{\hat{\mathbf{x}}}C(\hat{\mathbf{x}},t)e^{-|\mathbf{x}-\hat{\mathbf{x}}|^2/\delta}\ \mathrm{d}\hat{\mathbf{x}}, \tag{6.158}$$

for $C=\int_j C_j$, and since $\langle\nabla^2 s\rangle=\nabla^2\langle s\rangle=\nabla^2\hat{s}$ for $\hat{s}(\mathbf{x},t)=\langle s(\mathbf{x},t)\rangle$,

$$\frac{\partial\hat{s}}{\partial t} = D_s\nabla^2\hat{s} - \lambda_s\hat{s}\int_{\hat{\mathbf{x}}}C(\hat{\mathbf{x}},t)e^{-|\mathbf{x}-\hat{\mathbf{x}}|^2/\delta}\ \mathrm{d}\hat{\mathbf{x}}. \tag{6.159}$$

For $\delta\ll 1$ we may use Laplace's method iterated over the two nested integrals, first Taylor expanding $C(\hat{\mathbf{x}},t)$:

$$\int_{\hat{\mathbf{x}}}C(\hat{\mathbf{x}},t)e^{-|\mathbf{x}-\hat{\mathbf{x}}|^2/\delta}\ \mathrm{d}\hat{\mathbf{x}} = \sum_{n_x}\sum_{n_y}\frac{\partial^{(n_x+n_y)}C}{\partial x^{n_x}\partial y^{n_y}}\int_{\hat{x}}\frac{(x-\hat{x})^{n_x}}{i!}e^{-(x-\hat{x})^2/\delta}\ \mathrm{d}\hat{x}\int_{\hat{y}}\frac{(y-\hat{y})^{n_y}}{j!}e^{-(y-\hat{y})^2/\delta}\ \mathrm{d}\hat{y}, \tag{6.160}$$

and then extending the integral to $\mathbb{R}^2$ (introducing at most exponentially small errors),

$$\int_{\hat{\mathbf{x}}}C(\hat{\mathbf{x}},t)e^{-|\mathbf{x}-\hat{\mathbf{x}}|^2/\delta}\ \mathrm{d}\hat{\mathbf{x}}$$

$$= \sum_{n_x}\sum_{n_y}\frac{\partial^{(n_x+n_y)}C}{\partial x^{n_x}\partial y^{n_y}}\int_{-\infty}^{\infty}\frac{(x-\hat{x})^{n_x}}{i!}e^{-(x-\hat{x})^2/\delta}\ \mathrm{d}\hat{x}\int_{\infty}^{\infty}\frac{(y-\hat{y})^{n_y}}{j!}e^{-(y-\hat{y})^2/\delta}\ \mathrm{d}\hat{y}, \tag{6.161}$$

$$= \sum_{n_x,n_y\ even}\frac{\partial^{(n_x+n_y)}C}{\partial x^{n_x}\partial y^{n_y}}2\pi\left(\frac{\delta}{2}\right)^{\frac{n_x+1}{2}}\left(\frac{\delta}{2}\right)^{\frac{n_y+1}{2}}n_x!!n_y!!, \tag{6.162}$$

$$= \delta\pi C(\mathbf{x},t)+\mathcal{O}\left(\delta^{\frac{3}{2}}\right), \tag{6.163}$$

where $n!!$ is the product of every odd number from n to 1. Hence as $\delta\to 0$ with $\hat{\lambda}_s=\delta\pi\lambda_s$ constant

$$\frac{\partial\hat{s}}{\partial t} = D_s\nabla^2\hat{s} - \hat{\lambda}_s\hat{s}C(\mathbf{x},t), \tag{6.164}$$

so that $\hat{s}$ is governed by a similar equation to the one-dimensional case (equation (6.119)), with a

scaling factor of $\delta\pi$ rather than $\sqrt{\delta\pi}$, as was found in the one-dimensional model. It seems, therefore, that extending our characterisation of consumption as a continuous approximation to the dirac delta function to higher dimensions has been successful at reproducing the classical consumption term $-f(\mathbf{x})s$, where $f(\mathbf{x})$ is the rate of consumption at position $\mathbf{x}$.

The pdf for the position of cell k at time t satisfies

$$\begin{aligned} &C_k(\mathbf{x}, t+\Delta t) \\ &\quad = \text{Prob(cell } k \text{ at } \mathbf{x} \text{ stays at } \mathbf{x}) + \int_{-\pi}^{\pi} \text{Prob(cell } k \text{ moves in direction } \theta \text{ from } \mathbf{x} - \mathbf{d}_\theta)\ \mathrm{d}\theta, \quad (6.165) \\ &\quad = \overbrace{C_k(\mathbf{x},t)}^{\text{cell } k \text{ at } \mathbf{x}} \Bigg(\overbrace{1 - \Delta t \int_{-\pi}^{\pi} \int_s T(s,\theta,\mathbf{x},t) P(s,\mathbf{x},t)\ \mathrm{d}\theta\ \mathrm{d}s}^{\text{stays at } \mathbf{x}} \Bigg) \\ &\qquad + \overbrace{\Delta t \int_{-\pi}^{\pi} \int_s T(s,\theta,\mathbf{x}-\mathbf{d}_\theta,t) C_k(\mathbf{x}-\mathbf{d}_\theta,t) P(s,\mathbf{x},t)\ \mathrm{d}\theta\ \mathrm{d}s}^{\text{cell } k \text{ at } \mathbf{x}-\mathbf{d}_\theta \text{ moves to } \mathbf{x}}, \quad (6.166) \end{aligned}$$

where $\mathbf{d}_\theta = d(\cos\theta, \sin\theta)$ and we have made the moment closure approximation

$$\begin{aligned} \sum_l \int_s \text{Prob}(C_k^l(\mathbf{x},t) \text{ and } T(s,\theta,\mathbf{x},t))\ \mathrm{d}s &\approx \langle C_k^l(\mathbf{x},t)\rangle \langle T(s,\theta,\mathbf{x},t)\rangle, \quad (6.167) \\ &= C_k(\mathbf{x},t) \int_s T(s,\theta,\mathbf{x},t) P(s,\mathbf{x},t)\ \mathrm{d}s. \quad (6.168) \end{aligned}$$

As $\Delta t \to 0$ in equation (6.166)

$$\frac{\partial C_k}{\partial t} = \int_{-\pi}^{\pi} \int_s \left[T(s,\theta,\mathbf{x}-\mathbf{d}_\theta,t) C_k(\mathbf{x}-\mathbf{d}_\theta,t) - T(s,\theta,\mathbf{x},t) C_k(\mathbf{x},t) \right] P(s,\mathbf{x},t)\ \mathrm{d}s\ \mathrm{d}\theta. \quad (6.169)$$

To progress further with equation (6.169) we consider, as before, different forms for the rate of movement, $T(s,\theta,\mathbf{x},t)$, due to cell sensing.

6.4.1 Local sensing

For local sensing of s we take $T(s,\theta,\mathbf{x},t) = \alpha s(\mathbf{x},t)$ so that cells move down concentration gradients of s. Then, if d is small compared to the length scale over which C_k changes we may Taylor expand equation (6.169) in d so that

$$\frac{\partial C_k}{\partial t} = \frac{\alpha d^2 \pi}{2} \nabla^2 (s C_k) + \mathcal{O}(d^3), \quad (6.170)$$

and as $d \to 0$ and $\alpha \to \infty$ with $\hat{\alpha} = \alpha d^2 \pi/2$ constant, summing equation (6.170) over k, $C = \sum_{k=1}^{N} C_k$ satisfies

$$\frac{\partial C}{\partial t} = \hat{\alpha} \nabla^2 (sC) . \tag{6.171}$$

We find, therefore, that using an equivalent model of local chemotactic sensing results in an equivalent cell concentration equation to that derived in one dimension (equation (6.131)). We would expect this to be the case, since moving into two spatial dimensions does not significantly impact on the mechanism of migration, and only changes the multiplier $\hat{\alpha}$ from αd^2 in one dimension to $\alpha d^2 \pi/2$ in two dimensions. Thus in two dimensions, the diffusivity is changed. Indeed, even if we take a constant y-coordinate, the equations do not reduce to those derived in Section 6.3, since taking a slice through constant y does not reduce the model to the one-dimensional version. The model with constant y-coordinate differs to the one-dimensional formulation in the distance an individual can move during a single jump. Since in two dimensions individuals move a constant distance with a particular angle of movement, if that angle is not 0 or π, then when it is projected down onto a constant y-coordinate, it can be viewed as a smaller jump along the slice considered. Hence the two-dimensional system cannot be readily reduced to the initial one-dimensional model. Numerical solutions to equations (6.164) and (6.171) compare well to averaged simulations of the model (Figure 6.10(a)) for a range of parameter values (results not shown).

6.4.2 Non-local sensing

To simulate a chemoattractant we use non-local sensing, with $T(s, \theta, \mathbf{x}, t) = \alpha s(\mathbf{x} + \mathbf{d}_\theta, t)$ where $\mathbf{d}_\theta = d(\cos\theta, \sin\theta)$. Then, Taylor expanding in terms of d, equation (6.169) becomes

$$\frac{\partial C_k}{\partial t} = \frac{\alpha d^2 \pi}{2} \left(s \nabla^2 C_k - C_k \nabla^2 s \right) + \mathcal{O}(d^3), \tag{6.172}$$

and summing equation (6.172) over k, and letting $d \to 0$ and $\alpha \to \infty$ with $\hat{\alpha} = \alpha d^2 \pi/2$ constant,

$$\frac{\partial C}{\partial t} = \hat{\alpha} \left(s \nabla^2 C - C \nabla^2 s \right) , \tag{6.173}$$

where $C = \sum_{k=1}^{N} C_k$. As with the local sensing case, the two-dimensional non-local sensing model has a very similar equation (equation (6.173)) to the one-dimensional equivalent (equation (6.125)), with the one-dimensional derivatives replaced by higher-dimensional equivalents. In addition, as with

the local sensing case the scaling αd^2 is reduced to $\alpha d^2 \pi/2$ in the two-dimensional model due to the increase in possible directions of movement. Numerical solutions to equations (6.164) and (6.173) compare well to averaged simulations of the model (Figure 6.10(c)) for a range of parameter values (result not shown).

6.4.3 Volume exclusion

We may consider volume exclusion using similar techniques to Chapter 5, by excluding all cell movements which would result in cells overlapping either during or at the end of movement. We may then integrate over the area the cell moves through to find the probability of any given movement being prevented. With chemotactic movement,

$$C_k(\mathbf{x}, t+\Delta t) = \int_{-\pi}^{\pi} \int_s \{\underbrace{C_k(\mathbf{x},t)}_{\text{cell } k \text{ at } (\mathbf{x},t)} \Big[\overbrace{1-\alpha\Delta t T(s,\theta,\mathbf{x},t)}^{\text{no move attempted}} + \underbrace{\alpha\Delta t T(s,\theta,\mathbf{x},t) P_k(\mathbf{x},\theta,t)}_{\text{attempted move fails}}\Big] + \overbrace{\alpha\Delta t T(s,\theta,\mathbf{x}-\mathbf{d}_\theta,t) C_k(\mathbf{x}-\mathbf{d}_\theta,t)}^{\text{cell } k \text{ at } \mathbf{x}-\mathbf{d}_\theta \text{ tries to move}} \underbrace{[1-P_k(\mathbf{x}-\mathbf{d}_\theta,\theta,t)]}_{\text{attempted move succeeds}} P(s,\mathbf{x},t)\ \mathrm{d}s\ \mathrm{d}\theta, \quad (6.174)$$

where $P_k(\mathbf{x},\theta,t)$ is the probability that another cell prevents the movement of cell k from position $\mathbf{x}$ in direction θ and $P(s,\mathbf{x},t)$ is the probability of having chemoattractant s at position $\mathbf{x}$ at time t. $P_k(\mathbf{x},\theta,t)$ is given by equation (5.5), so that

$$P_k(\mathbf{x},\theta,t) = \sum_{j\neq k} \int_0^d \int_{-\pi/2}^{\pi/2} 2R\cos\phi\ C_j(x + r\cos\theta + 2R\cos(\theta+\phi), y + r\sin\theta + 2R\sin(\theta+\phi), t)\ \mathrm{d}\phi\ \mathrm{d}r. \quad (6.175)$$

6.4.3.1 Local sensing

For local sensing we substitute $T(s,\theta,\mathbf{x},t) = \alpha s(\mathbf{x},t)$ into equation (6.174) to give

$$\frac{\partial C_k}{\partial t} = \alpha \int_{-\pi}^{\pi} [s(\mathbf{x}-\mathbf{d}_\theta,t) C_k(\mathbf{x}-\mathbf{d}_\theta,t) - s(\mathbf{x},t) C_k(\mathbf{x},t)]\ \mathrm{d}\theta - \alpha \int_{-\pi}^{\pi} [P_k(\mathbf{x}-\mathbf{d}_\theta,\theta,t) s(\mathbf{x}-\mathbf{d}_\theta,t) C_k(\mathbf{x}-\mathbf{d}_\theta,t) - P_k(\mathbf{x},\theta,t) s(\mathbf{x},t) C_k(\mathbf{x},t)]\ \mathrm{d}\theta, \quad (6.176)$$

as $\Delta t \to 0$. Substituting equation (6.175) into equation (6.176) and Taylor expanding in R and r, since $r \in [0, d]$ gives

$$\frac{\partial C_k}{\partial t} = \frac{\alpha d^2 \pi}{2} \nabla \cdot \left(\nabla \left(s C_k \right) \left(1 - 4Rd \sum_{j \neq k} C_j \right) + 4\pi R^2 s C_k \sum_{j \neq k} \nabla \cdot C_j \right) + \mathcal{O}(d^3), \tag{6.177}$$

and as $d \to 0$ with $\hat{\alpha} = \alpha d^2 \pi / 2$ constant then, summing equation (6.177) over k,

$$\frac{\partial C}{\partial t} = \hat{\alpha} \left(\nabla^2 \left(sC \right) + \frac{N-1}{N} 4\pi R^2 \nabla \cdot \left(sC \nabla \cdot C \right) \right), \tag{6.178}$$

where $C = \sum_{k=1}^{N} C_k$. Thus volume exclusion results in modifications to equation (6.171) to give equations (6.177) and (6.178), which are equivalent to those in one dimension (equation (6.131) compared to equations (6.139) and (6.140)). As in one dimension this can be regarded as an increase in the diffusion term due to a higher proportion of successful movements towards less populated areas. The surfaces found by varying pairs of parameters and measuring the error between the numerical solution to equation (6.178) and averaged simulations (Figures 6.11(a) and 6.11(b)), the predicted volume exclusion effect (Figures 6.11(c) and 6.11(d)), and the actual volume exclusion effect (Figures 6.11(e) and 6.11(f)) are changed, however, compared to the one-dimensional system (see Figure 6.6). In particular, the error in both the $R-d$ (Figure 6.11(a)) and $R-N$ (Figure 6.11(b)) surfaces depends less sensitively on the parameter values and the $R - d$ error is much smaller compared to the volume exclusion found from data (Figure 6.11(e)). The reduction in errors may be due to the smaller correlations between cell positions in two dimensions compared to the one-dimensional case, where individuals cannot move past each other and therefore are constrained to their initial ordering. Higher correlations reduce the accuracy of our derived equations in one dimension, since our derivation assumes approximate independence of cell positions. The reduction in errors thus improves the predicted exclusion surfaces (Figures 6.11(c) and 6.11(d)) so that they are more qualitatively similar to those found from averaged simulation data (Figures 6.11(e) and 6.11(f)). In addition, the error no longer peaks at the highest values of R and N, instead it is greatest where there are a large number of small cells. This may be due to higher correlations from initialising a large number of cells very close together thus leading to higher errors for a large group of small cells. However, initialising a greater number of cells close together also increases the interactions between the individuals, so that there are comparatively higher volume exclusion effects (Figures 6.11(d) and 6.11(f)).

6.4.3.2 Non-local sensing

For non-local sensing, $T(s,\theta,\mathbf{x},t) = \alpha s(\mathbf{x}+\mathbf{d}_\theta,t)$ and equation (6.174) becomes

$$\frac{\partial C_k}{\partial t} = \alpha \int_{-\pi}^{\pi} [s(\mathbf{x},t)C_k(\mathbf{x}-\mathbf{d}_\theta,t) - s(\mathbf{x}+\mathbf{d}_\theta,t)C_k(\mathbf{x},t)]\ \mathrm{d}\theta \\ - \alpha \int_{-\pi}^{\pi} [P_k(\mathbf{x}-\mathbf{d}_\theta,\theta,t)s(\mathbf{x},t)C_k(\mathbf{x}-\mathbf{d}_\theta,t) - P_k(\mathbf{x},\theta,t)s(\mathbf{x}+\mathbf{d}_\theta,t)C_k(\mathbf{x},t)]\ \mathrm{d}\theta, \quad (6.179)$$

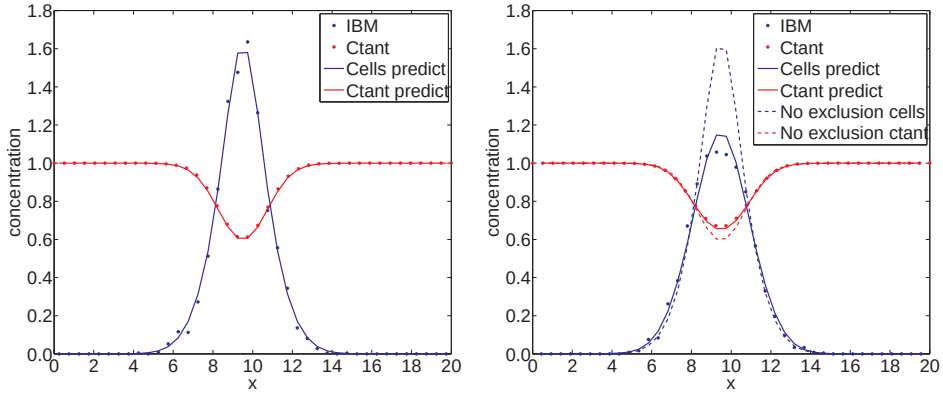
as $\Delta t \to 0$. Substituting equation (6.175) into equation (6.176) and Taylor expanding in R and r, where $r \in [0,d]$ gives

$$\frac{\partial C_k}{\partial t} = \frac{\alpha d^2 \pi}{2} \nabla \cdot \left((s\nabla C_k - C_k \nabla s) \left(1 - 4Rd \sum_{j \neq k} C_j \right) + 4\pi R^2 s C_k \sum_{j \neq k} \nabla \cdot C_j \right) + \mathcal{O}(d^3), \quad (6.180)$$

and as $d \to 0$ with $\hat{\alpha} = \alpha d^2 \pi / 2$ constant, summing equation (6.180) over k,

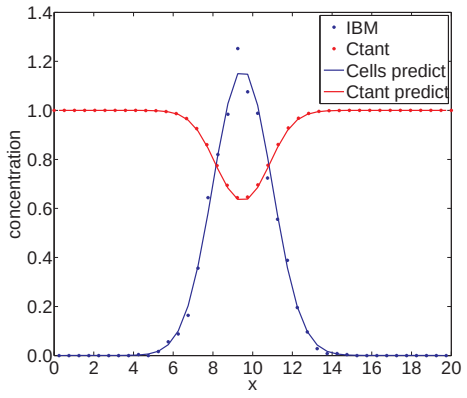
$$\frac{\partial C}{\partial t} = \hat{\alpha} \left(s\nabla^2 C_k - C_k \nabla^2 s + \frac{N-1}{N} 4\pi R^2 \nabla \cdot (sC\nabla \cdot C) \right), \quad (6.181)$$

where $C = \sum_{k=1}^{N} C_k$. Thus the terms due to volume exclusion in equations (6.180) and (6.181) are comparable to those in equations (6.177) and (6.178) and are again equivalent to those seen in one dimension (equations (6.148) and (6.149)). As with local sensing, the error surfaces (Figures 6.12(a) and 6.12(b)) are much reduced from the one-dimensional case (Figures 6.7(a) and 6.7(b)), due to decreased correlations between individuals. For non-local chemotactic sensing, the volume exclusion effects as $R-N$ (Figure 6.12(f)) and $R-d$ (Figure 6.12(e)) vary are similar to their one-dimensional counterparts (Figures 6.7(e) and 6.7(f), increasing with R and N, but also increasing as R decreases and is smaller than d. The non-monotonic profile of exclusion effects as R changes may be explainable as in the diffusion-only case (see Chapter 5) by the playoff between migration enhanced by volume exclusion (due to a higher proportion of successful moves away from populated areas) and migration hindered by volume exclusion (due to a lower number of successful moves). Thus there is a region in parameter space where the two effects cancel so that the model with volume exclusion is very similar to the model without exclusion.

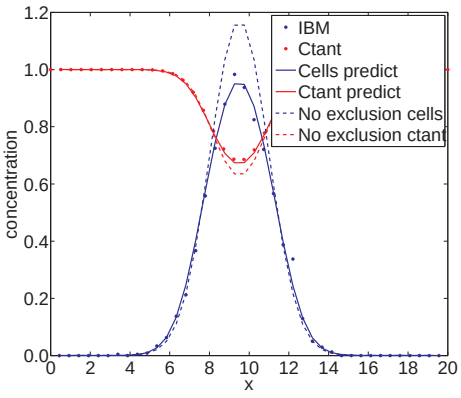


(a) Local sensing

(b) Local sensing with volume exclusion

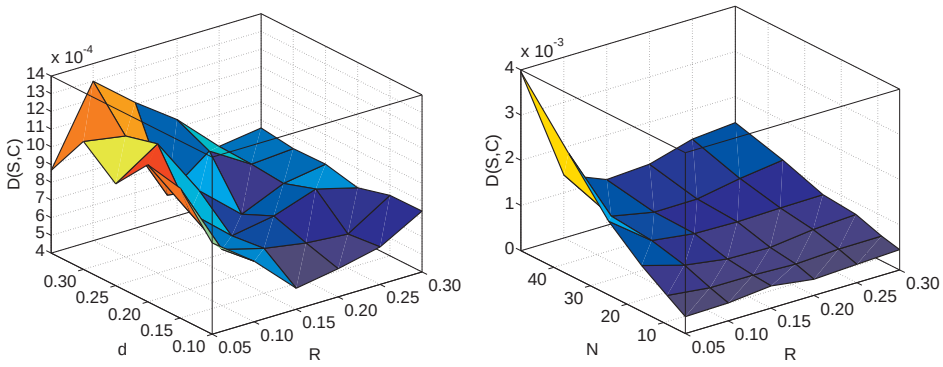


(c) Nonlocal sensing



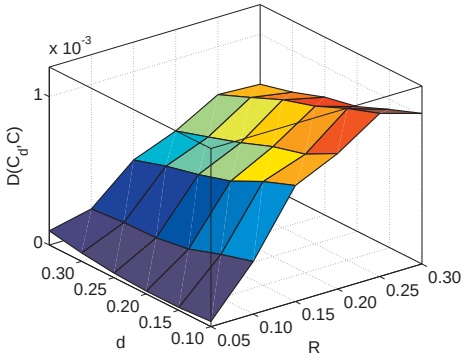
(d) Nonlocal sensing with volume exclusion

Figure 6.10: A comparison of typical averaged simulations of the two-dimensional off-lattice chemotaxis models described in Section 6.4 and numerical solutions to the derived equations for (a) local sensing (equations (6.164) and (6.171)); (b) non-local sensing (equations (6.164) and (6.173)); (c) local sensing with volume exclusion (equations (6.164) and (6.178)) and (d) non-local sensing with volume exclusion (equations (6.164) and (6.181)). Each simulation is performed with $R = 0.05$, $d = 0.1$, $\Delta t = 1$, $\Delta x = \Delta y = 0.5$, $D_s = 0.01$, $\lambda_s = 0.01$, $\delta = 0.1$, whilst $\alpha = 5$ and simulations are averaged over $1,000$ runs. Sixteen cells are initially placed in a regular grid in the middle of the domain with the left most, bottom-most cell centre (x, y) drawn from the normal distribution $x, y \sim \mathcal{N}(10 - \sqrt{2RN}/0.8, 1)$ and no flux conditions are taken at each boundary.

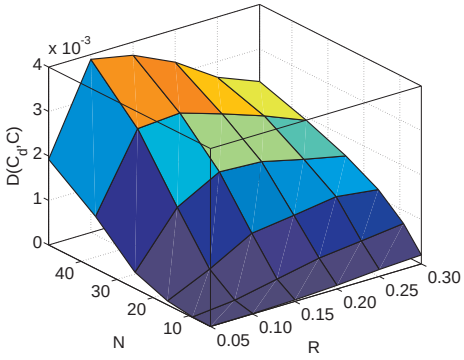


(a) Error as $d - R$ change

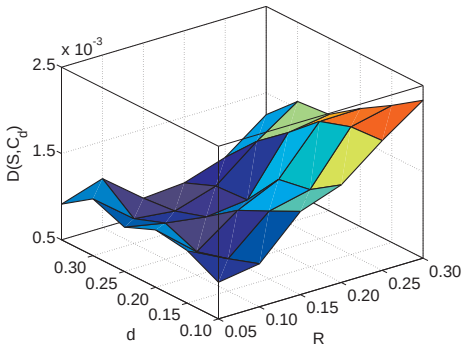
(b) Error as $N - R$ change



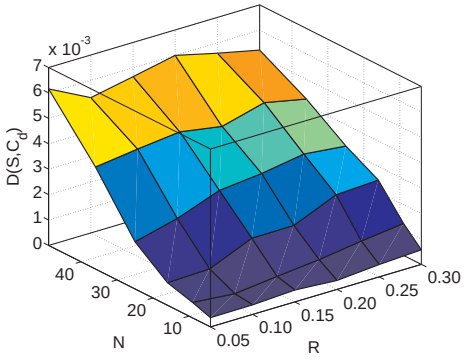
(c) Predicted volume exclusion as $d - R$ change



(d) Predicted volume exclusion as $N - R$ change

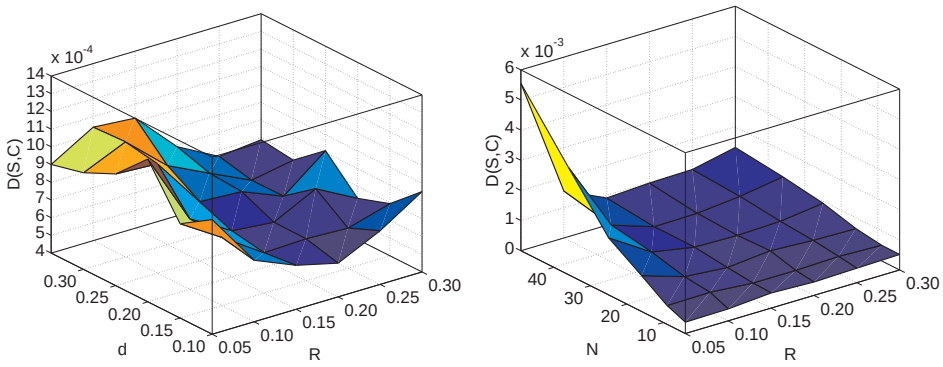


(e) Volume exclusion as $d - R$ change



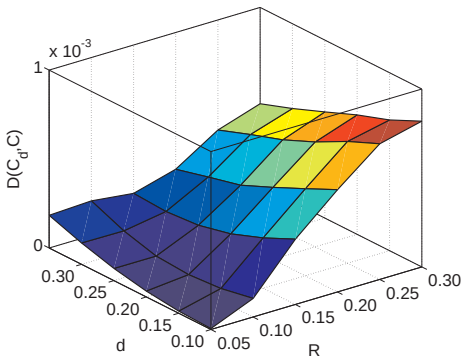
(f) Volume exclusion as $N - R$ change

Figure 6.11: Exploring the effect of changing the likelihood of successful movement in Model 4 with local chemotactic sensing by changing the distance moved, d, and the radius of a cell, R ((a), (c) and (e)) and changing the occupied space in the model by changing the number of cells, N, and the radius of a cell, R ((b), (d) and (f)). We consider the effect on: (a)-(b) the error introduced by our approximations (the difference between the solution to equations (6.164) and (6.178) and the average of our simulations); (c)-(d) our prediction of the effect of volume exclusion (the difference between the solution to equations (6.164) and (6.178) and equations (6.164) and (6.171); and (e)-(f) the actual effect of volume exclusion (the difference between the average of our simulations and the solution to equations (6.164) and (6.171). Differences are calculated using equation (5.20) and non-varying parameters are: $R = 0.05$, $d = 0.1$, $\Delta t = 1$, $\Delta x = \Delta y = 0.5$, $D_s = 0.01$, $\lambda_s = 0.01$, $\delta = 0.1$, whilst $\alpha = 0.05/d^2$ and simulations are averaged over $1,000$ runs. Sixteen cells are initially placed in a regular grid in the middle of the domain with the left most, bottom-most cell centre (x, y) drawn from the normal distribution $x, y \sim \mathcal{N}(10 - \sqrt{2RN}/0.8, 1)$ and no flux conditions are taken at each boundary.

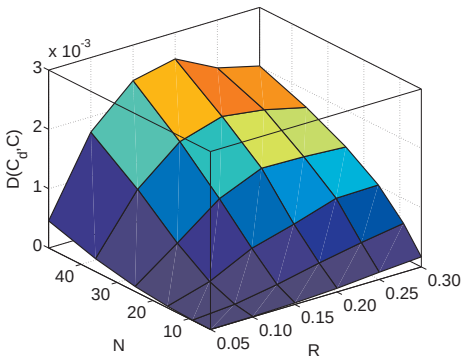


(a) Error as $d - R$ change

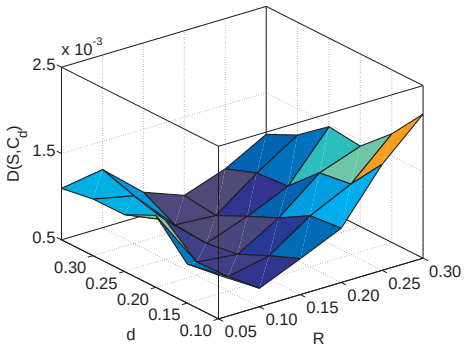
(b) Error as $N - R$ change



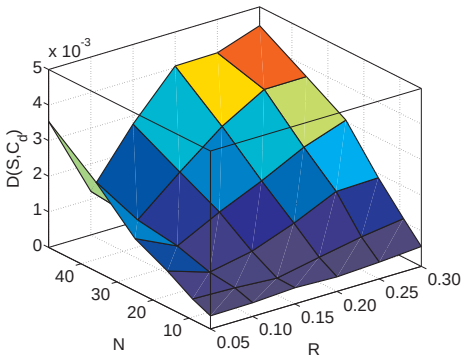
(c) Predicted volume exclusion as $d - R$ change



(d) Predicted volume exclusion as $N - R$ change



(e) Volume exclusion as $d - R$ change



(f) Volume exclusion as $N - R$ change

Figure 6.12: Exploring the effect of changing the likelihood of successful movement in Model 4 with non-local chemotactic sensing by changing the distance moved, d, and the radius of a cell, R ((a), (c) and (e)) and changing the occupied space in the model by changing the number of cells, N, and the radius of a cell, R ((b), (d) and (f)). We consider the effect on: (a)-(b) the error introduced by our approximations (the difference between the solution to equations (6.164) and (6.181) and the average of our simulations); (c)-(d) our prediction of the effect of volume exclusion (the difference between the solution to equations (6.164) and (6.181) and equations (6.164) and (6.173); and (e)-(f) the actual effect of volume exclusion (the difference between the average of our simulations and the solution to equations (6.164) and (6.173). Differences are calculated using equation (5.20) and non-varying parameters are: $R = 0.05$, $d = 0.1$, $\Delta t = 1$, $\Delta x = \Delta y = 0.5$, $D_s = 0.01$, $\lambda_s = 0.01$, $\delta = 0.1$, whilst $\alpha = 0.05/d^2$ and simulations are averaged over $1,000$ runs. Sixteen cells are initially placed in a regular grid in the middle of the domain with the left most, bottom-most cell centre (x, y) drawn from the normal distribution $x, y \sim \mathcal{N}(10 - \sqrt{2RN}/0.8, 1)$ and no flux conditions are taken at each boundary.

6.5 Conclusions

In this chapter we have discussed the problems of analysing two-species hybrid models, in which one species is modelled as a continuum whilst the other is modelled discretely. We have focussed on chemotactic systems such as the biological system investigated in Chapters 2 and 3. We began with on-lattice chemotactic systems with local or non-local sensing of the chemical. Local sensing resulted in a chemorepellent system and the derived continuum system of equations included a diffusion term together with an advection term down concentration gradients of the chemorepellent. In contrast, a chemoattractant system results from non-local sensing and gives a similar system, with the advection term leading to cell movement up concentration gradients of the chemoattractant. The addition of volume exclusion did not significantly change the cell concentration profiles within the domain, but did greatly increase cell concentrations next to the boundary when flux boundary conditions are included. In these on-lattice models volume exclusion reduced the advection term proportional to the cell concentration. Extending these models to a hybrid model with a continuum chemoattractant resulted in exactly the same equations, which also compared well with averaged simulations of the model.

We next extended our analysis to off-lattice hybrid models and found that the continuum equations derived were equivalent to those on-lattice, with a scaled chemical consumption term due to our choice of coupling between the continuum chemical and the individual-based cells. Using the techniques from Chapter 4, we incorporated volume exclusion into the model, resulting in a different modification to our derived continuum equations than was found in the on-lattice models. In particular, not only were the multipliers of the volume exclusion terms different to those on-lattice, but also the form of the equations were different. In contrast to the reduced advection on-lattice, volume exclusion in our off-lattice hybrid system modified the diffusion term in a similar way to the diffusive system studied in Chapter 4.

Finally, using the analysis technique discussed in Chapter 5 we considered two spatial dimensions and derived a continuum system from off-lattice hybrid chemotactic models. The systems derived in this way were found to be straightforward two-dimensional generalisations of the systems in one dimension, and volume exclusion modified the diffusion term as in Chapter 5.

Recall that our goal is to derive continuum equations for the chemotactic models in Chapters 2 and 3 to enable us to analyse the effects of changes of parameters on the distribution and speed of migration of the cell population. In the final chapter we will put together the techniques from

Chapters 4 to 6 to gain further insight into the models from Chapters 2 and 3.

Chapter 7

Discussion and conclusions

We began in Chapters 2 and 3 by considering the migration of CNCCs, building a hybrid model of cell migration with chemotaxis and using it to investigate the biological system and its response to experimental perturbations. However, the time taken for each simulation of the model prohibited a detailed investigation into the effects of varying parameter values. Our desire to investigate such systems further motivated us to build systematic techniques to derive PDE representations of individual-based and hybrid systems in one and two dimensions.

In this chapter we summarise the aims and results of this study and discuss how the methods derived in Chapters 4 to 6 may be used in the future to analyse models such as those in Chapters 2 and 3. Finally we consider more broadly, future avenues of research into the areas discussed here.

7.1 Summary and conclusions

In this thesis we have investigated the migration of CNCCs in the developing chick embryo using mathematical modelling and analytical examination of the resulting IBMs. Specifically, our aim was to determine possible mechanisms for cell migration consistent with experimental evidence, to make predictions for experimental outcomes using these mechanisms and test our predictions against the results of these experiments (Chapter 3). The resultant model motivated the development of systematic techniques for predicting the results of one-dimensional individual-based simulations in large numbers and extending these to two dimensions and chemotactic systems.

7.1.1 Main results

The main results from this thesis are summarised here. We began in Chapter 1 by summarising the research to date in: the experimental and biological fields of CNCC migration; the mathematical modelling of the system and related migratory systems; and the derivation of continuum equations from IBMs. In Chapter 2 we built an IBM of the system and demonstrated that a solely chemotactic system, in which cells create a gradient by consuming chemoattractant and then move up that gradient, results in later emerging cells being trapped next to the domain entrance. Thus we extended this model to include two subpopulations of cells with later emerging cells gaining directional information from earlier cells and postulate that this may be a possible mechanism for the migration of the cell stream seen *in vivo*. We did not speculate on the precise anatomy of the migratory stream, but instead present this as an alternative to the previously hypothesised chemotactic system.

In Chapter 3 we used the developed model to predict experimental outcomes. In particular, using a chemoattractant gradient dynamically produced by cell consumption, our model predicted that:

- the cell population spread along the migratory route when the neural tube is removed partway through migration, preventing later cells from entering the domain;
- cells transplanted from the leading portion of a donor migratory stream to the neural tube of a host embryo fail to migrate out of the transplanted region;
- cells transplanted from the trailing portion of a donor embryo ahead of a host migratory stream also fail to migrate out of the transplanted region.

The first two of these predictions were subsequently observed experimentally and, prior to collaborative discussions, the failure of migration in the front-to-back transplant had been regarded as a failure of the experiment techniques, rather than evidence for our modelling hypotheses. The final prediction was found to be incorrect and motivated further extensions to our model. Specifically, allowing cells to change between subpopulations and modelling the transplanted region as smaller than the migratory domain width enabled a successful simulation of the back-to-front transplant experiment.

To analyse the hybrid models developed in Chapters 2 and 3 more closely under varying parameter regimes required a measure of the general outcome of the model over many simulations. This motivated us to develop systematic techniques for deriving PDE descriptions from IBMs which may also then be compared with the more standard PDE representations. In addition, this allows for a

more explicit connection between biological and experimental results and the parameterisation of the mathematical model.

In Chapter 4 we began with one-dimensional off-lattice models with volume exclusion and derived a modified diffusion equation with diffusivity dependent on the cell concentration. Next we extended our models to include other mechanisms such as domain growth and biased movement. Our derived equations with volume exclusion were consistently more accurate than the diffusion equation at predicting the cell profile averaged over multiple simulations. We concluded, therefore, that it is important to consider the space occupied by individuals, particularly in cases involving biased movement.

Since the models studied in Chapters 2 and 3 were two-dimensional, in Chapter 5 we extended our techniques from Chapter 4 into two dimensions. To enable analysis of the system with volume exclusion, we extended the area excluded in a cell movement to all regions moved through during the jump, rather than just the final destination. We then used Taylor expansions in the distance moved in a jump and the radius of a cell to derive a continuum equation and demonstrated that the derived PDE is of a similar form to those found in Chapter 4. The multiplier of the extra term, however, was found to be the area, $\pi(2R)^2$, excluded by an individual in two dimensions, rather than the distance, $2R$, excluded by an individual in one dimension. As before, the derived equations with volume exclusion were consistently better at predicting the results of averaged simulations than the equivalent equations without volume exclusion, and we also used our technique to derive equations for models with additional mechanisms.

Finally, in Chapter 6 we considered models of chemotaxis and derived systems of equations for on- and off-lattice hybrid models with local and non-local chemotactic sensing. The derived equations consisted of a diffusion and an advection term, up or down gradients of the chemical, dependent on the choice of local or non-local chemotactic sensing. We showed that hybrid chemotactic systems result in similar continuum equations, but that the chemical consumption term must be scaled accordingly. The effect of volume exclusion on the derived equation was found to differ for off-lattice models, reducing the advection term, compared to the modification of the diffusivity found in on-lattice models. Next we extended our analysis to two-dimensional models and demonstrated that the equations derived were higher-dimensional equivalents to the one-dimensional systems studied.

7.2 Future work

Future work may be divided into three main categories. Firstly, work is required to directly apply the analytical techniques from Chapters 4 to 6 to the models used in Chapters 2 and 3. Secondly, further investigation, both experimental and through modelling, into the biological system. Finally, we consider other physical and biological systems that may be examined using similar techniques to those in this work.

7.2.1 Analysing the models from Chapters 2 and 3

In Chapter 6 we derived continuum equations from a two-dimensional hybrid chemotaxis model with volume exclusion. There are several additional mechanisms that we must consider to compare this model with that constructed in Chapter 2. In particular, the model in Chapter 2 includes domain growth parallel to the x-axis and a flux boundary condition at $x = 0$. These are easily incorporated into our derivation in a similar way to Section 5.4 and Section 5.2.3.1 in the following sections.

However, the exact form of chemotaxis in our derivation differs from that used in Chapter 2. In Chapters 2 and 3, chemotaxis is represented by attempted movement by each cell at each time step in a random direction. Movement is successful if, and only if, the chemoattractant is favourable in the desired direction and there are no other cells or boundaries blocking the movement. In contrast, in Chapter 6 we modify the rate of movement directly so that movement is only attempted at a rate determined with respect to the chemoattractant concentration. In addition, whilst the chemoattractant sensing is non-local in our derivations, it is carried out at a point a distance d away from the individual. Hence the sensing and movement distances are linked, and both tend to zero during our derivation so that in the limit sensing can be approximated from the value and derivatives of the chemical concentrations using Taylor expansions.

Future work, beyond the scope of this thesis, should consider the incorporation of truly non-local sensing into our derivation. In addition, the inclusion of subpopulations of cells, and the mechanisms governing trailing cells, is beyond the scope of this thesis. We therefore derive continuum equations using the form of chemotaxis used in Chapter 6, with only one cell population, and use these equations to explore the behaviour of the model under different parameter regimes. This methodology can then be used in the future with equations derived from the chemotactic mechanisms used in Chapters 2 and 3.

7.2.1.1 Domain growth

It is straightforward to include domain growth into our derivation from Section 6.4.3.2 in a similar way to Section 5.4 since equation (6.174) may be written as

$$C_k(\mathbf{x}, t + \Delta t) = C_k(\mathbf{x}, t) + \Delta t f(s, C_k, \mathbf{x}, \mathbf{d}_\theta), \tag{7.1}$$

where $f(s, C_k, \mathbf{x}, \mathbf{d}_\theta)$ is determined by the type of cell movement in the model. Hence, if the domain grows uniformly, parallel to the x-axis, so that the domain at time t is given by $[0, L(t)] \times [0, L_y]$,

$$\delta x \delta y C_k(\mathbf{x}, t + \Delta t) = \delta x \delta y \left(1 - \frac{\Delta t L'}{L}\right) \left(C_k(\mathbf{x} \cdot \mathbf{L}, t) + \Delta t f(s, C_k, \mathbf{x} \cdot \mathbf{L}, \mathbf{d}_\theta \cdot \mathbf{L})\right), \tag{7.2}$$

$$C_k(\mathbf{x}, t + \Delta t) = C_k(\mathbf{x}, t) + \Delta t f(s, C_k, \mathbf{x}, \mathbf{d}_\theta) - \frac{\Delta t L'}{L} C_k - \frac{\Delta t L' x}{L} \frac{\partial C_k}{\partial x} + \mathcal{O}(\Delta t^2), \tag{7.3}$$

where $\mathbf{L} = (1/(1 + \Delta t L'/L), 1) \approx (1 - \Delta t L'/L + \mathcal{O}(\Delta t^2), 1)$. Then taking $\Delta t \to 0$

$$\frac{\partial C_k}{\partial t} = f(s, C_k, \mathbf{x}, \mathbf{d}_\theta) - \frac{\partial}{\partial x}\left(\frac{L' x}{L} C_k(\mathbf{x}, t)\right). \tag{7.4}$$

Thus the effect of domain growth is an additional term $\partial/\partial x \left(L' x / L C_k(\mathbf{x}, t)\right)$ to equation (6.181). If we rescale $x' = x/L(t)$, so that the domain is stationary, with unit length, under the new coordinates then, with non-local sensing, dropping the dashes,

$$\frac{\partial C}{\partial t} + \frac{L'}{L} C = \hat{\alpha}\left(s \nabla_L^2 C - C \nabla_L^2 s + \frac{N-1}{N} 4\pi R^2 \nabla_L \cdot (s C \nabla_L \cdot C)\right), \tag{7.5}$$

$$\nabla_L = \left(\frac{1}{L}\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right), \tag{7.6}$$

$$\nabla_L^2 = \left(\frac{1}{L^2}\frac{\partial^2}{\partial x^2}, \frac{\partial^2}{\partial y^2}\right). \tag{7.7}$$

Similarly, for the chemoattractant equation, using equation (6.157),

$$\begin{aligned} \delta x \delta y \langle s(x, t + \Delta t)\rangle = \delta x \delta y \left(1 - \frac{\Delta t L'}{L}\right) \Bigg(& \langle s(\mathbf{x} \cdot \mathbf{L}, t\rangle \\ & + D_s \Delta t \int_{s_n} s_n \text{Prob}\left(\nabla^2 s = s_n\right) \, \mathrm{d}s_n \\ & - \lambda_s \Delta t \int_{\hat{s}} \hat{s} P(\hat{s}, \mathbf{x}, t) \, \mathrm{d}\hat{s} \int_{\hat{\mathbf{x}}} \sum_{j=1}^{N} C_j(\hat{\mathbf{x}}, t) e^{-|\mathbf{x} \cdot \mathbf{L} - \hat{\mathbf{x}}|^2/\delta} \, \mathrm{d}\hat{\mathbf{x}} \Bigg), \end{aligned} \tag{7.8}$$

Hence

$$\frac{\partial\langle s\rangle}{\partial t} = D_s\langle\nabla^2 s\rangle - \lambda_s\langle s\rangle\int_{\hat{\mathbf{x}}} C(\hat{\mathbf{x}},t)e^{-|\mathbf{x}-\hat{\mathbf{x}}|^2/\delta}\,\mathrm{d}\hat{\mathbf{x}} - \frac{\partial}{\partial x}\left(\frac{\Delta t L' x}{L}\langle s\rangle\right), \tag{7.9}$$

and, since $\langle\nabla^2 s\rangle = \int_{\mathbf{x}}(\nabla^2 s) = \nabla^2(\int_{\mathbf{x}} s) = \nabla^2\langle s\rangle = \nabla^2\hat{s}$ for $\hat{s}(\mathbf{x},t) = \langle s(\mathbf{x},t)\rangle$, rescaling $x' = x/L(t)$ and dropping dashes,

$$\frac{\partial\hat{s}}{\partial t} + \frac{L'}{L}\hat{s} = D_s\nabla_L^2\hat{s} - \lambda_s\hat{s}\int_{\hat{\mathbf{x}}} C(\hat{\mathbf{x}},t)e^{-|\mathbf{x}-\hat{\mathbf{x}}|^2/\delta}\,\mathrm{d}\hat{\mathbf{x}}, \tag{7.10}$$

where ∇_L is as defined in equations (7.6) and (7.7). Finally, Taylor expanding in terms of δ as in Section 6.4 gives

$$\frac{\partial\hat{s}}{\partial t} + \frac{L'}{L}\hat{s} = D_s\nabla_L^2\hat{s} - \hat{\lambda}_s\hat{s}C(\mathbf{x},t). \tag{7.11}$$

7.2.1.2 Flux boundary condition

We consider an influx of cells at $x = R$, with new individuals inserted at uniformly distributed y-positions with rate P_c whilst cancelling insertions that would result in overlapping cells. To derive an appropriate boundary condition for equation (7.5) we proceed as in Section 5.2.3.1 to consider the change in the expected number of cells in the domain at time t

$$\begin{aligned}
\frac{\partial}{\partial t}\int_{\mathcal{D}} C\,\mathrm{d}\mathbf{x} &= \int_{\mathcal{D}}\frac{\partial C}{\partial t}\,\mathrm{d}\mathbf{x}, &(7.12)\\
&= \int_{\mathcal{D}}\hat{\alpha}\nabla_L\cdot\left(s\nabla_L C - C\nabla_L s + \frac{N-1}{N}4\pi R^2\left(sC\nabla_L\cdot C\right)\right) - \frac{L'}{L}C\,\mathrm{d}\mathbf{x}, &(7.13)\\
&= \hat{\alpha}\left(s\nabla_L C - C\nabla_L s + \frac{N-1}{N}4\pi R^2\left(sC\nabla_L\cdot C\right)\right)\Big|_{\partial\mathcal{D}} - \frac{L'}{L}\int_{\mathcal{D}} C\,\mathrm{d}\mathbf{x}, &(7.14)\\
&= -L_y\hat{\alpha}\frac{1}{L}\left(s\frac{\partial C}{\partial x} - C\frac{\partial s}{\partial x} + \frac{N-1}{N}4\pi R^2\left(sC\frac{\partial C}{\partial x}\right)\right)\Big|_{x=R} - \frac{L'}{L}\int_{\mathcal{D}} C\,\mathrm{d}\mathbf{x}, &(7.15)
\end{aligned}$$

where $\mathcal{D} = [0,1]\times[0,L_y]$ is the domain with scaled x-coordinate, as described in Section 7.2.1.1. However, this may also be expressed as

$$\begin{aligned}
\frac{\partial}{\partial t}\int_{\mathcal{D}} C\,\mathrm{d}\mathbf{x} &= P_c\left(1 - \int_0^{2R}\int_{-\pi/2}^{\pi/2} rC(\mathbf{x}+\mathbf{r}_\theta^r)\,\mathrm{d}r\,\mathrm{d}\theta\right) - \frac{L'}{L}\int_{\mathcal{D}} C\,\mathrm{d}\mathbf{x}, &(7.16)\\
&= P_c\left(1 - 2\pi R^2 C\right)\Big|_{x=R} - \frac{L'}{L}\int_{\mathcal{D}} C\,\mathrm{d}\mathbf{x}, &(7.17)
\end{aligned}$$

where we Taylor expand in terms of R and r as before, and the final term is due to dilution from the expansion of the domain. Hence

$$P_c\left(1-2\pi R^2 C\right)\Big|_{x=R} = -L_y\hat{\alpha}\frac{1}{L}\left(s\frac{\partial C}{\partial x} - C\frac{\partial s}{\partial x} + \frac{N-1}{N}4\pi R^2\left(sC\frac{\partial C}{\partial x}\right)\right)\Big|_{x=R}. \tag{7.18}$$

If $L(t) = L_x$ for all time, t, then without rescaling x this becomes

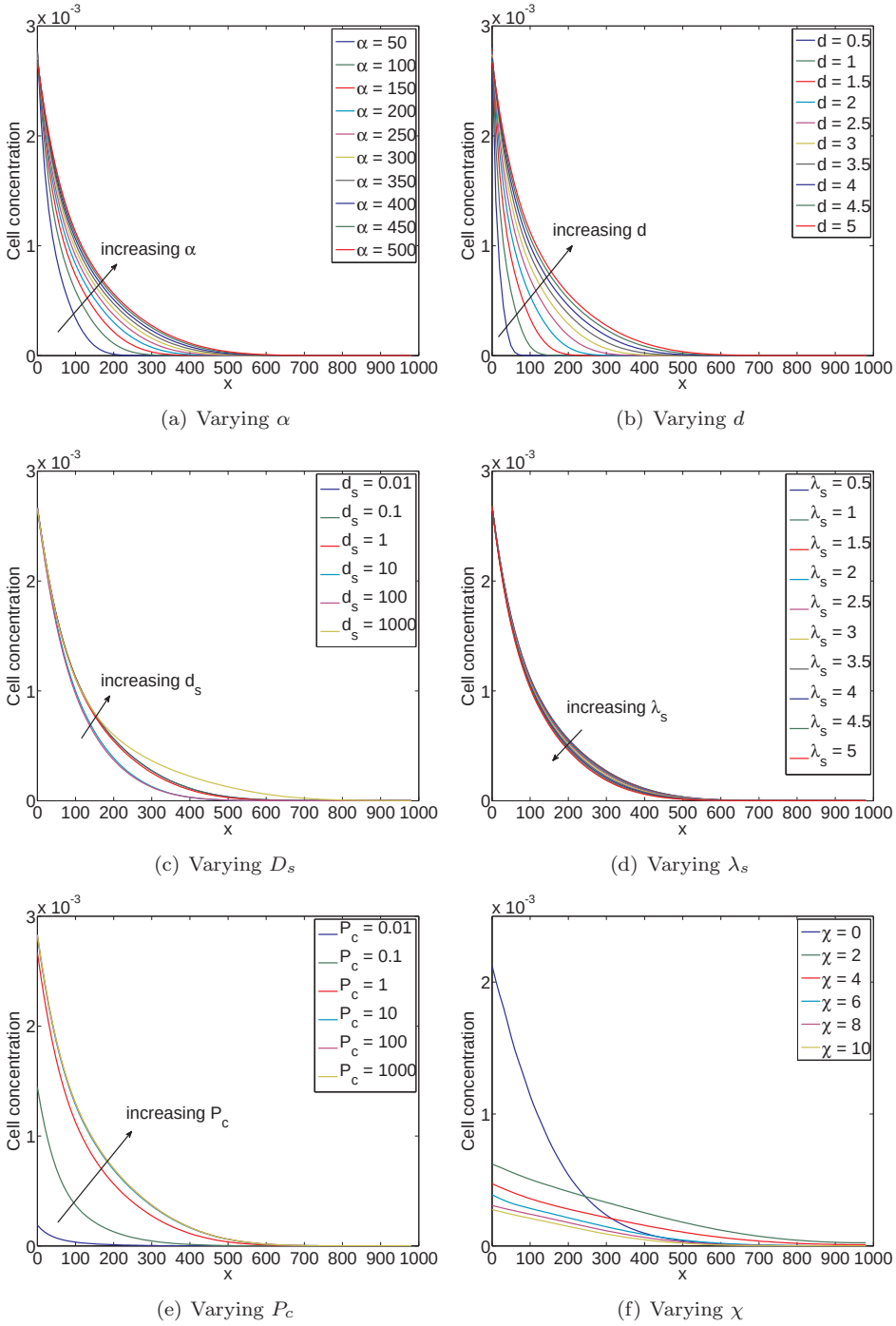
$$P_c\left(1-2\pi R^2 C\right)\Big|_{x=R} = -L_y L_x\hat{\alpha}\left(s\frac{\partial C}{\partial x} - C\frac{\partial s}{\partial x} + \frac{N-1}{N}4\pi R^2\left(sC\frac{\partial C}{\partial x}\right)\right)\Big|_{x=R}. \tag{7.19}$$

7.2.1.3 Parameters

We wish to use the same parameters set as in Table 2.1 and thus take $L_y = 120$, $R = 7.5$ and $L(t)$ given by equation (2.1). It is not clear, however, how best to use the experimentally found cell movement speed (45μm/hr) to find α and d. Whilst λ_s and D_s are specified in the models in Chapters 2 and 3, since there is little experimental data for the values taken, we will regard α, D_s, λ_s and P_c as free parameters to explore. We have assumed that δ is small in our derivations, however, when considering domain growth and rescaling the x-coordinate, we require $\delta/L \ll 1$ so that $\delta = 30\mu$m, as used in Chapters 2 and 3 still satisfies this requirement.

7.2.1.4 Results

We solve equation (7.5) with equation (7.18) over a range of parameter values with (Figure 7.1) and without (Figure 7.2) volume exclusion. In Chapters 2 and 3 the profile of the cell population changes as the y-coordinate varies, with the boundaries of the domain affecting the movement of individuals, since sensing next to the boundary is over a smaller area than in the middle of the domain, resulting in a lower concentrations being sensed next to the boundaries (see Figure 2.7). In contrast, the localised approximation to chemotaxis in our derivation results in numerical solutions, $C(x, y, t)$, that are invariant with respect to y. We therefore only show the profile for a particular y-coordinate as x changes. Since our experimental data are given for $t \leq 24$ hours, we plot the cell concentration profile at $t = 24$ hours. We compare to the distance migrated by cells *in vivo* ($\sim 600\mu$m, see Figure 3.3) and in the model described in Section 2.5, which includes only chemotactic 'leader' cells, ($\sim 500\mu$m, see Figure 2.10). We find that taking $\alpha = 500$/hr with $d = 5\mu$m does reproduce the approximate migration distances seen in Figures 2.10 and 3.3 when volume exclusion is taken into account (Figures 7.1(a) and 7.1(b)) and, furthermore, this distance is not increased further



(a) Varying α

(b) Varying d

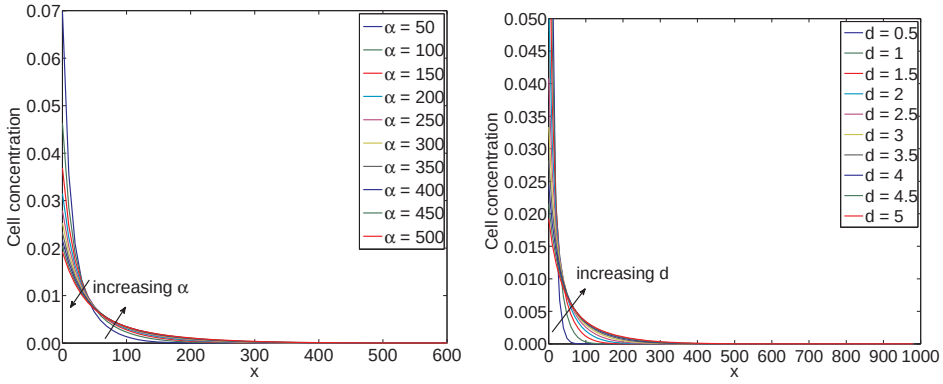
(c) Varying D_s

(d) Varying λ_s

(e) Varying P_c

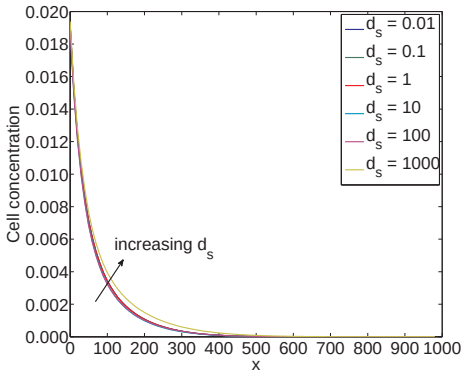
(f) Varying χ

Figure 7.1: A comparison of cell concentration profiles after 24 hours with varying parameter combinations, found by solving equation (7.5) numerically with volume exclusion, with equation (7.18) at $x = 0$ and no flux boundary conditions at the other boundaries. Solutions are given as a slice through constant $y = 60\mu$m. Other parameter values are: $\alpha = 500\mu$m/hr, $d = 5\mu$m, $D_s = 0.01\mu\text{m}^2$/hr, $\lambda_s = 0.01$/hr, $P_c = 1$/hr, $\chi = 0$/hr and $R = 7.5\mu$m.

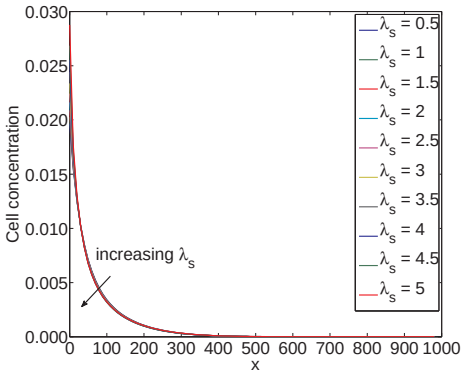


(a) Varying α

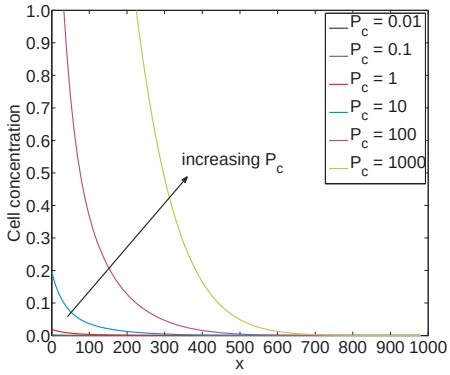
(b) Varying d



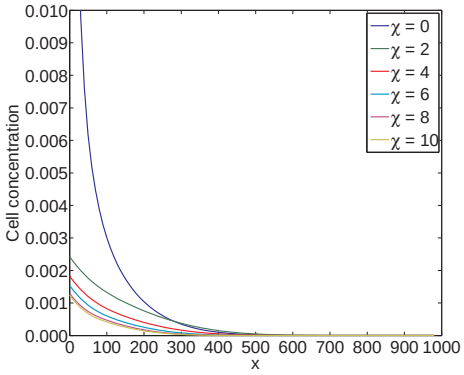
(c) Varying D_s



(d) Varying λ_s

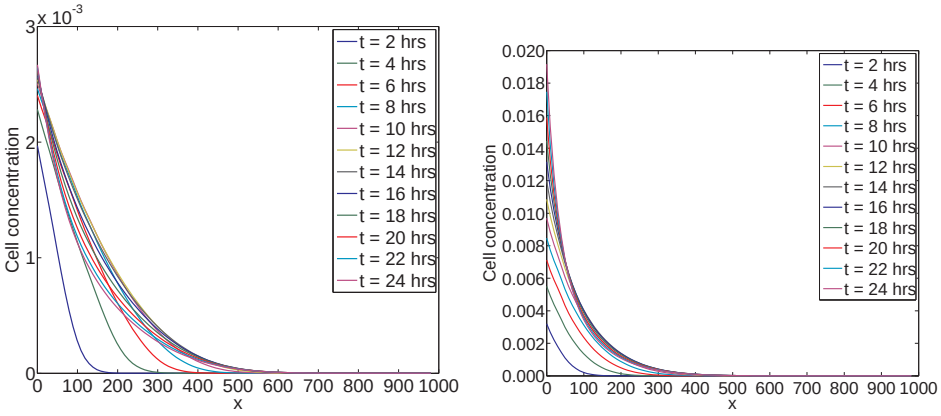


(e) Varying P_c



(f) Varying χ

Figure 7.2: A comparison of cell concentration profiles after 24 hours with varying parameter combinations, found by solving equation (7.5) numerically without volume exclusion, with equation (7.18) at $x = 0$ and no flux boundary conditions at the other boundaries. Solutions are given as a slice through constant $y = 60\mu$m. Other parameter values are: $\alpha = 500\mu$m/hr, $d = 5\mu$m, $D_s = 0.01\mu\text{m}^2$/hr, $\lambda_s = 0.01$/hr, $\chi = 0$/hr and $P_c = 1$/hr.



(a) With volume exclusion

(b) Without volume exclusion

Figure 7.3: A comparison of cell concentration profiles at different time points with varying parameter combinations, found by solving equation (7.5) numerically with and without volume exclusion, with equation (7.18) at $x = 0$ and no flux boundary conditions at the other boundaries. Solutions are given as a slice through constant $y = 60\mu$m. Other parameter values are: $\alpha = 500\mu$m/hr, $d = 5\mu$m, $D_s = 0.01\mu\text{m}^2$/hr, $\lambda_s = 0.01$/hr, $\chi = 0$/hr and $P_c = 1$/hr.

by taking larger values of α or d (result not shown). The growth of the domain may account for this surprising maximum movement distance since the migration profile is essentially fixed at the gradient attained when the chemoattractant is reduced to zero due to the rapid growth-driven dilution. When the chemoattractant is fully depleted there is no further migration in the domain and the cell concentration is fixed. If chemoattractant production is included, however, whilst the chemoattractant is not fully depleted, the growth of the domain effectively restricts further movement since the distances moved during a time step is no longer significant compared to the length of the domain (result not shown). Hence the steady state profile is essentially dependent on the form of the domain growth, and not the diffusivity or rate of consumption of the chemoattractant (Figures 7.1(c) and 7.1(d)). The final cell concentration profile is therefore not very sensitive to the parameter values chosen, with the exception of the domain growth parameters, which have been well characterised experimentally (Figure 2.2).

It is interesting to note that nearly all cell concentration profiles, over the parameter combinations tried, rise substantially next to the left-hand boundary where there is an influx of cells. The only parameter value that affects the concentration profiles after 24 hours is when a chemoattractant production term, $\chi s(1 - s)$, is included. At higher values of χ, the cell profile is less steep at the boundary. However, the profile still rises towards $x = 0$, where the flux boundary condition inputs new cells. Even when the flux rate, P_c, is reduced (Figure 7.1(e)) or the rate of movement, α (Figure 7.1(a)), (or, equivalently, the distance moved, d; Figure 7.1(b)) is increased, there are no

travelling wave-like solutions such as those predicted by Landman *et al.* (2003) in the enteric neural crest, in which a fixed wavefront profile travelled with a given speed through the domain. The difference in the two predictions is due to the lack of extensive proliferation of cells during migration in the cranial neural crest. In comparison, the enteric neural crest invade primarily due to proliferation, leading to Fisher-type invasive waves. Indeed, in our model, the concentration profiles become steeper at later time points, particularly in the case with volume exclusion (Figure 7.3(a)), as the chemotactic flux in the positive x-direction exceeds the influx of cells at the $x = 0$ boundary. In the model with volume exclusion (Figure 7.3(a)), the concentration profiles do not change significantly after $t = 24$ hours, as the concentration at $x = 0$ is already at the maximum allowed by volume exclusion and since the chemoattractant has been consumed to the point where there is no longer a gradient in the domain, there is no further migration towards the $x = L(t)$ boundary. When volume exclusion is not incorporated into the model (Figure 7.3(b)), there is no maximum concentration to limit the value of C at the $x = 0$ boundary and hence the density continues to increase. However the chemoattractant is still consumed and hence the additional cell mass that is input at the boundary does not migrate from there and instead increases the boundary values much higher than within the domain (Figure 7.3(b)). Increasing P_c without volume exclusion does give greater penetration into the domain, however the profile is the same qualitative shape as with smaller values of P_c so that massive, biologically-unrealistic, cell concentrations are required at the boundary to allow the density to reach 600μm in 24 hours (Figure 7.2(e)). Changing the other parameters does not significantly affect the cell concentration profile at $t = 24$ hours (Figures 7.1(d) and 7.2(a) to 7.2(c)). Hence the continuous equations derived from a similar model to that in Section 2.5 also display an inability to successfully migrate to the end of the domain and colonise the migratory route.

To enable direct comparison of the models considered here with those used in Chapters 2 and 3, future work will concentration on different ways of modelling chemotactic movement and the addition of subpopulations of cells that migrate by the 'following' mechanism described in Section 2.6.0.1. Thus to progress further with the analysis of the models described in Chapters 2 and 3 in the future it will be necessary to include a second cell population (as in Chapter 5), and investigate ways of allowing individuals to gain directional information from others. Finally, to compare with Chapters 2 and 3, the chemotactic movement must be modelled as a constant rate of movement where the success of movement depends solely on whether the chemoattractant gradient is favourable. In addition we will consider a nonlocal sensing mechanism in which the distance at which the chemoattractant is sensed is different to the distance moved at each time step. Thus we may take the limit as the

movement distance tends to zero, whilst keeping the sensing distance non-zero. We might expect that such a system would result in an integro-differential equation where non-local effects are explicitly included in the resulting equation. Maintaining a non-zero sensing distance is also likely to lead to non-homogeneous solutions over y-coordinates.

7.2.2 Further investigation into the biological system

Future work to further interrogate the biological system could focus on multiple different aspects of the migratory system. It is clear that any further work will require close collaboration with experimental biologists to provide further data and expertise, and enable feedback between modelling predictions and experimental outcomes. In particular, it is unclear at present whether there are more subpopulations than the two indicated by morphological analysis and gene profiling. Further work could therefore focus on establishing more precisely the makeup of the stream population and the mechanisms by which the subpopulations migrate. For example, cell adhesion profiles and cell-cell contact dynamics give insight into cell movement patterns and methods of migration. Mathematical modelling could focus on including and testing new potential mechanisms, and including different proportions of the subpopulations to test whether there is an optimum makeup for the migratory stream. In addition, it is possible that leading CNCCs may secrete substances that break down the ECM, thus making it easier for later, trailing, cells to pass through the same routes. Thus when trailing cells follow behind leading cells, this may not imply informational exchange between the cells but rather that a 'pathway' has been established that increases the likelihood of other cells migrating the same way. This could be tested by establishing whether the migratory routes of later cells change during migration or are fixed en-route, whilst taking the growth of the domain into account.

7.2.3 Future applications

We have presented a highly multi-disciplinary, collaborative exploration of a biological system using mathematical modelling techniques and used this as motivation to develop techniques for deriving continuum equations to analyse individual-based and hybrid models. These techniques are applicable to a wide variety of IBMs in which mechanisms may be approximated as local to the individual. For example, the multiple population models in Chapter 5 may be useful for studying diffusion and directed migration in crowded environments. It is known that in many biological and physical sys-

tems individuals or particles diffuse in an 'anomalous' way. That is, the mean squared displacement of the individuals does not scale linearly with time as would be predicted by the diffusion equation. There has been increasing interest in this effect, since it indicates that there may be a problem with the many mathematical models of biological systems that assume that particles and individuals move by diffusion. Our analysis of multi-population models may be used to investigate this phenomenon, since we may include populations of cells of different sizes and migratory speeds and investigate the dependence of the evolution of the mean squared displacement on these different parameters.

7.3 Concluding remarks

This thesis has investigated the migration of CNCCs through a close collaboration between mathematical modellers and experimental biologists. Mathematical models have given greater insights into *in vivo* experiments, some of which would have otherwise been regarded as failed experiments instead of evidence for the hypothesised mechanisms. Greater insight into migratory models has been gained through the development of techniques to analyse off-lattice individual-based and hybrid models with volume exclusion.

Appendix A

Numerical methods

A.1 The NAG library solver d03pc

Throughout this work we have solved systems of PDEs in one spatial dimension using the NAG algorithm d03pc from the NAG toolbox for Matlab (Numerical Algorithms Group, documentation). This routine solves a system of linear or nonlinear parabolic PDEs of the form

$$\sum_{j=1}^{\mathbf{npde}} P_{i,j}\frac{\partial U_j}{\partial t} + Q_i = x^{-m}\frac{\partial}{\partial x}\left(x^m R_i\right), \qquad i = 1, 2, \ldots, \mathbf{npde}, \qquad a \leq x \leq b, \qquad t \geq t_0, \tag{A.1}$$

where $\mathbf{U} = (U_1, \ldots, U_{\mathbf{npde}})$ and $\mathbf{U}_x = (U_{1x}, \ldots, U_{\mathbf{npde}_x})$ are the vector of solutions and their partial derivatives, respectively, $P_{i,j}$, Q_i and R_i depend on $x, t, \mathbf{U}$ and $\mathbf{U}_x$ and $m = 0$ for Cartesian coordinates. Boundary conditions are of the form

$$\beta_i(x,t)R_i(x,t,\mathbf{U},\mathbf{U}_x) = \gamma_i(x,t,\mathbf{U},\mathbf{U}_x) \qquad i = 1, 2, \ldots, \mathbf{npde}, \tag{A.2}$$

where $x = a$ or $x = b$. The algorithm uses finite differences and the method of lines to reduce the problem to a system of ODEs which are then solved using a backwards differentiation formula (BDF) method.

The solution is given at time $t = t_{out}$ and on the spatial mesh $x_i = idx$, where dx is the mesh step size. To generate solutions at multiple time points, $t_k = kdt$, the routine is run multiple times with

$t_0 = (k-1)dt$ and $t_{out} = kdt$. Then the solution at time t_k and point x_i is approximated by

$$\mathbf{u}_i^k \approx \mathbf{U}(x_i, t_k). \tag{A.3}$$

A.1.1 Method of lines

The method of lines is used to reduce a parabolic system of PDEs into a larger system of ODEs by discretising in space. For simple Cartesian problems a modified three-point method is used as described here. The problem is assumed to have the form

$$D(x, t, \mathbf{U}, \mathbf{U}_x)U_t = \frac{\partial}{\partial x}(G(x, \mathbf{U})\mathbf{U}_x) + f(x, t, \mathbf{U}, \mathbf{U}_x), \tag{A.4}$$

so that $P = D$, $Q = f$ and $R = G(x, \mathbf{U})\mathbf{U}_x$. The problem is first reformulated into a system of two first-order PDEs:

$$\mathbf{U}_x = H(x, \mathbf{U})\mathbf{V} := \frac{1}{G(x, \mathbf{U})}\mathbf{V}, \tag{A.5}$$

$$\mathbf{V}_x = Q(x, \mathbf{U}, \mathbf{U}_x, \mathbf{U}_t) := D(x, t, \mathbf{U}, \mathbf{U}_x)\mathbf{U}_t - f(x, \mathbf{U}, \mathbf{U}_x). \tag{A.6}$$

Considering a typical mesh subinterval $[\alpha, \beta]$, then integrating equation (A.6) for $\mathbf{V}$ gives

$$V(x) = V(\alpha) + \int_\alpha^x Q(\cdots)\ \mathrm{d}y, \tag{A.7}$$

and substituting this into equation (A.5) and integrating reveals

$$U(\beta) = U(\alpha) + \int_\alpha^\beta \left(V(\alpha) + \int_\alpha^x Q(\cdots)\ \mathrm{d}y \right) H(\cdots)\ \mathrm{d}x. \tag{A.8}$$

Using β instead of α in the derivation gives

$$U(\alpha) = U(\beta) - \int_\alpha^\beta \left(V(\beta) - \int_x^\beta Q(\cdots)\ \mathrm{d}y \right) H(\cdots)\ \mathrm{d}x. \tag{A.9}$$

Then, approximating H, D, and f by their values at a point $\xi = (\alpha + \beta)/2$,

$$U(\beta) = U(\alpha) + \int_\alpha^\beta H_0 \left(V(\alpha) + \int_\alpha^x (D_0 U_t(\alpha) - f_0) \, \mathrm{d}y \right) \, \mathrm{d}x + \sigma + H_0 \tau_\alpha, \tag{A.10}$$

$$U(\alpha) = U(\beta) - \int_\alpha^\beta H_0 \left(V(\beta) - \int_x^\beta (D_0 U_t(\beta) - f_0) \, \mathrm{d}y \right) \, \mathrm{d}x - \sigma + H_0 \tau_\beta, \tag{A.11}$$

where σ, τ_α and τ_β are the errors due to the approximation of H, $\int_\alpha^x Q \, \mathrm{d}y$ and $\int_x^\beta Q \, \mathrm{d}y$, respectively. Here H_0, D_0 and f_0 are the values of H, D and f, respectively, at ξ and $U(\xi)$ is determined by linear interpolation between α and β. Performing the integrations in equations (A.10) and (A.11), and using $U_x(\xi) = (U(\beta) - U(\alpha)/(\beta - \alpha)$:

$$G_0 U_x(\xi) = V(\alpha) + (D_0 U_t(\alpha) - f_0)(\xi - \alpha) + \text{error}, \tag{A.12}$$

$$G_0 U_x(\xi) = V(\beta) - (D_0 U_t(\alpha) - f_0)(\beta - \xi) + \text{error}. \tag{A.13}$$

Now if $[\alpha, \beta] = [x_{j-1}, x_j]$ and $g_{j-1/2} = G_0 U_x(\xi)$ then, changing notation in the obvious way,

$$g_{j-1/2} = v_{j-1} + (D_{j-1/2} \dot{u}_{j-1} - f_{j-1/2})(x_{j-1/2} - x_{j-1}), \tag{A.14}$$

$$-g_{j-1/2} = -v_j + (D_{j-1/2} \dot{u}_j - f_{j-1/2})(x_j - x_{j-1/2}), \tag{A.15}$$

and adding equation (A.14) at $[x_j, x_{j+1}]$ to equation (A.15) at $[x_{j-1}, x_j]$ then in the interior of the domain

$$g_{j+1/2} - g_{j-1/2} = (D_{j+1/2} \dot{u}_j - f_{j+1/2})(x_{j+1/2} - x_j) + (D_{j-1/2} \dot{u}_j - f_{j-1/2})(x_j - x_{j-1/2}). \tag{A.16}$$

The boundary conditions are determined by solving the relevant equation from equations (A.14) and (A.15) for v_j and substituting that for g in the boundary condition. Thus a system of ODEs is derived which must now be solved using a BDF method.

A.1.2 Backwards differentiation formula method

Equation (A.16) is of the form

$$\frac{\partial u}{\partial t} = \phi(u(t), t), \tag{A.17}$$

which is solved using a BDF method. An m-step BDF method finds an approximation to u at a time step, $t = t_k$, using a function of the approximations found at previous time steps, $t = t_{k-1}, \ldots, t_{k-m}$. The first six m-step methods are given by

$$\begin{aligned}
\text{BDF1:} \quad & u_{n+1} - u_n = h\phi(u_{n+1}, t_{n+1}), \\
\text{BDF2:} \quad & u_{n+2} - \tfrac{4}{3}u_{n+1} + \tfrac{1}{3}u_n = \tfrac{2}{3}h\phi(u_{n+2}, t_{n+2}), \\
\text{BDF3:} \quad & u_{n+3} - \tfrac{18}{11}u_{n+2} + \tfrac{9}{11}u_{n+1} - \tfrac{2}{11}u_n = \tfrac{6}{11}h\phi(u_{n+3}, t_{n+3}), \\
\text{BDF4:} \quad & u_{n+4} - \tfrac{48}{25}u_{n+3} + \tfrac{36}{25}u_{n+2} - \tfrac{16}{25}u_{n+1} + \tfrac{3}{25}u_n = \tfrac{12}{25}h\phi(u_{n+4}, t_{n+4}), \\
\text{BDF5:} \quad & u_{n+5} - \tfrac{300}{137}u_{n+4} + \tfrac{300}{137}u_{n+3} - \tfrac{200}{137}u_{n+2} + \tfrac{75}{137}u_{n+1} - \tfrac{12}{137}u_n = \tfrac{60}{137}h\phi(u_{n+5}, t_{n+5}), \\
\text{BDF6:} \quad & u_{n+6} - \tfrac{360}{147}u_{n+5} + \tfrac{450}{147}u_{n+4} - \tfrac{400}{147}u_{n+3} + \tfrac{225}{147}u_{n+2} \\
& \qquad - \tfrac{72}{147}u_{n+1} + \tfrac{10}{147}u_n = \tfrac{60}{147}h\phi(u_{n+6}, t_{n+6}).
\end{aligned} \tag{A.18}$$

Methods for which $m > 6$ are not given here as they are not zero-stable, that is, a perturbation in the initial conditions of size ϵ results in changes to the numerical solution that are greater than $K\epsilon$ for any constant K not dependent on the step size, h.

A.2 The NAG library solver d03ra

Systems of PDEs in two spatial dimensions in this thesis are solved numerically using the NAG toolbox for Matlab routine d03ra (Numerical Algorithms Group, documentation), which integrates PDEs on a rectangular domain. As in d03pc (Appendix A.1), the method of lines is used to reduce the system to ODEs which are then solved using a BDF method.

In two dimensions, d03ra reduces the PDE system

$$F_j(t, x, y, u, u_t, u_x, u_y, u_{xx}, u_{xy}, u_{yy}) = 0, \qquad j = 1, 2, \ldots, \mathbf{npde}, \tag{A.19}$$

by the method of lines to a system of ODEs using a standard centred second-order finite difference scheme:

$$\left[\frac{\partial u}{\partial x}\right]_{x_j} = \frac{u_{j+1} - u_j}{x_{j+1} - x_j}. \tag{A.20}$$

Here x and y in the domain $x_{min} \le x \le x_{max}$, $y_{min} \le y \le y_{max}$ and $t_0 \le t \le t_{out}$. The system is solved on a user-supplied grid, with local uniform refinement where required. Grid refinement is

controlled by requiring that

$$\max_i(\mu_i^s) < 0.9, \tag{A.21}$$

(or 1.0, depending on the grid level at the previous step) at each time step where

$$\mu_i^s = \max_{j=1,\mathbf{npde}} \left(\frac{w_j^s}{u_j^{max}\sigma} \left(\left| (\Delta x)^2 \frac{\partial^2}{\partial x^2} u_j(x_i, y_i, t) \right| + \left| (\Delta y)^2 \frac{\partial^2}{\partial y^2} u_j(x_i, y_i, t) \right| \right) \right), \tag{A.22}$$

where Δx and Δy are the grid spacing parameters in the x- and y-directions, respectively, and σ is the user-supplied tolerance and w_j^s is a weighting factor. u_j^{max} is the approximate maximum absolute value of the jth component. If the condition in equation (A.21) is not satisfied, then a new level of grid refinement is created, with all cells surrounding points at which $\mu_i^s > 0.25$ quartered in size.

The resulting system of ODEs is solved using a second-order two-step, implicit BDF method such as $BDF2$ in equation (A.18).

The method also includes variable time step size, controlled by restricting the size of

$$\mu_l^t = \sqrt{\frac{1}{N} \sum_{j=1}^{\mathbf{npde}} w_j^t \sum_{i=1}^{\mathbf{ngpts}(l)} \left(\frac{\Delta t}{\alpha_{ij}} u_t(x_i, y_i, t) \right)^2}, \tag{A.23}$$

where $\mathbf{ngpts}(l)$ is the total number of points on grid level l, $N = \mathbf{ngpts}(l) \times \mathbf{npde}$, Δt is the current time step, u_t is the first-order finite difference approximation to the time derivative of u, w_j^t is a weighting factor and

$$\alpha_{ij} = \tau \left(\frac{u_j^{max}}{100} + |u(x_i, y_i, t)| \right), \tag{A.24}$$

for the user-specified tolerance, τ. It is required that $\max_l(\mu_l^t) \leq 1.0$ at each time step, and the step is retried at all levels if the condition fails.

Bibliography

Alt, W., Lauffenburger, D., 1987. Transient behavior of a chemotaxis system modelling certain types of tissue inflammation. *J. Math. Biol.* (1985), 691–722.

Baker, R. E., Simpson, M. J., 2010. Correcting mean-field approximations for birth-death-movement processes. *Phys. Rev. E* 82 (4), 041905.

Baker, R. E., Yates, C. A., Erban, R., 2010. From microscopic to macroscopic descriptions of cell migration on growing domains. *Bull. Math. Biol.* 72 (3), 719–762.

Barkai, E., Silbey, R., 2010. Diffusion of tagged particle in an exclusion process. *Phys. Rev. E* 81 (4), 041129.

Boldajipour, B., Mahabaleshwar, H., Kardash, E., Reichman-Fried, M., Blaser, H., Minina, S., Wilson, D., Xu, Q., Raz, E., 2008. Control of chemokine-guided cell migration by ligand sequestration. *Cell* 132 (3), 463–73.

Bruna, M., Chapman, S. J., 2012. Excluded-volume effects in the diffusion of hard spheres. *Phys. Rev. E* 85 (1), 011103.

Callaghan, T., Khain, E., Sander, L., Ziff, R., 2006. A stochastic model for wound healing. *J. Stat. Phys.* 122, 909–924.

Carmona-Fontaine, C., Matthews, H., Kuriyama, S., Moreno, M., Dunn, G., Parsons, M., Stern, C., Mayor, R., 2008. Contact inhibition of locomotion in vivo controls neural crest directional migration. *Nature* 456 (7224), 957–961.

Chen, R. R., Silva, E. a., Yuen, W. W., Brock, A. a., Fischbach, C., Lin, A. S., Guldberg, R. E., Mooney, D. J., 2007. Integrated approach to designing growth factor delivery systems. *FASEB J.* 21 (14), 3896–903.

Codling, E. A., 2003. Biased random walks in biology. Ph.D. thesis, University of Leeds. http://129.11.36.6/applied/phd/ecodlingthesis.pdf

Codling, E. A., Plank, M. J., Benhamou, S., 2008. Random walk models in biology. *J. R. Soc. Interface* 5 (25), 813–34.

Dallon, J. C., Othmer, H. G., 1997. A discrete cell model with adaptive signalling for aggregation of Dictyostelium discoideum. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 352, 391–417.

de Melker, A. A., Desban, N., Duband, J.-L., 2004. Cellular localization and signaling activity of beta-catenin in migrating neural crest cells. *Dev. Dyn.* 230 (4), 708–26.

Deroulers, C., Aubert, M., Badoual, M., Grammaticos, B., 2009. Modeling tumor cell migration: From microscopic to macroscopic models. *Phys. Rev. E* 79 (3), 031917.

Dormann, D., Weijer, C. J., 2006. Chemotactic cell movement during Dictyostelium development and gastrulation. *Curr. Opin. Genet. Dev.* 16 (4), 367–73.

Dyson, L., Maini, P., Baker, R., 2012. Macroscopic limits of individual-based models for motile cell populations with volume exclusion. *Phys. Rev. E* 86 (3), 031903.

Eberhart, J. K., He, X., Swartz, M. E., Yan, Y., Song, H., Boling, T. C., Kunerth, A. K., Walker, M. B., Kimmel, C. B., Postlethwait, J. H., 2008. MicroRNA Mirn140 modulates Pdgf signaling during palatogenesis. *Nat. Genet.* 40, 290–298.

Erban, R., Othmer, H. G., 2004. From individual to collective behavior in bacterial chemotaxis. *SIAM J. Appl. Math.* 65 (2), 361–391.

Fernando, A. E., Landman, K. A., Simpson, M. J., 2010. Nonlinear diffusion and exclusion processes with contact interactions. *Phys. Rev. E* 81 (1), 011903.

Flache, A., Hegselmann, R., 2001. Do irregular grids make a difference? Relaxing the spatial regularity assumption in cellular models of social dynamics. *JASSS* 4 (4).

Franz, B., Erban, R., 2013. Hybrid modelling of individual movement and collective behaviour. *Dispersal, Individual Movement and Spatial Ecology.*

Gejji, R., Lushnikov, P., Alber, M., 2012. Macroscopic model of self-propelled bacteria swarming with regular reversals. *Phys. Rev. E*, 1–21.

Gilbert, S. F., 2006. Developmental Biology. Sinauer Associates.

Gillespie, C. S., 2009. Moment-closure approximations for mass-action models. *IET Syst. Biol.* 3 (1), 52–58.

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

Guo, Z., Sloot, P. M. A., Tay, J. C., 2008. A hybrid agent-based approach for modeling microbiological systems. *J. Theor. Biol.* 255 (2), 163–75.

Helbing, D., Farkas, I., Vicsek, T., 2000. Simulating dynamical features of escape panic. *Nature* 407 (6803), 487–90.

Hillen, T., Painter, K., 2009. A user's guide to PDE models for chemotaxis. *J. Math. Biol.* 58 (1-2), 183–217.

Johnston, M. C., Bronsky, P. T., 1995. Prenatal craniofacial-development - new insights on normal and abnormal mechanisms. *Crit. Rev. Oral Biol. Med.* 6, 25–79.

Keller, E. F., Segel, L. a., 1971. Model for chemotaxis. *J. Theor. Biol.* 30 (2), 225–34.

Kubota, Y., Ito, K., 2000. Chemotactic migration of mesencephalic neural crest cells in the mouse. *Dev. Dyn.* 217, 170–179.

Kulesa, P. M., Bailey, C. M., Kasemeier-Kulesa, J. C., McLennan, R., 2010. Cranial neural crest migration: New rules for an old road. *Dev. Biol.* 344, 543–554.

Kulesa, P. M., Ellies, D. L., Trainor, P. A., 2004. Comparative analysis of neural crest cell death, migration, and function during vertebrate embryogenesis. *Dev. Dyn.* 229, 14–29.

Kulesa, P. M., Lu, C. C., Fraser, S. E., 2005. Time-lapse analysis reveals a series of events by which cranial neural crest cells reroute around physical barriers. *Brain Behav. Evol.* 66, 255–265.

Kulesa, P. M., Teddy, J. M., Stark, D. A., Smith, S. E., McLennan, R., 2008. Neural crest invasion is a spatially-ordered progression into the head with higher cell proliferation at the migratory front as revealed by the photoactivatable protein, KikGR. *Dev. Biol.* 316, 275–287.

Landman, K. A., Pettet, G. J., Newgreen, D. F., 2003. Mathematical models of cell colonization of uniformly growing domains. *Bull. Math. Biol.* 65, 235–262.

Landman, K. A., Simpson, M. J., Newgreen, D. F., 2007. Mathematical and experimental insights into the development of the enteric nervous system and Hirschsprung's Disease. *Develop. Growth Differ.* 49, 277–286.

Latta, E. J., Golding, J. P., 2012. Regulation of PP2A activity by Mid1 controls cranial neural crest speed and gangliogenesis. *Mech. Dev.* 128 (11-12), 560–76.

Lee, J. M., Hillen, T., Lewis, M. A., 2008. Continuous traveling waves for prey-taxis. *Bull. Math. Biol.* 70 (3), 654–76.

Luster, A., 2001. Chemotaxis: role in immune response. *eLS*.

McLennan, R., Dyson, L., Prather, K. W., Morrison, J. A., Baker, R. E., Maini, P., Kulesa, P. M., 2012. Multiscale mechanisms of cell migration during development: theory and experiment. *Development* 139 (16), 2935–2944.

McLennan, R., Kulesa, P. M., 2007. In vivo analysis reveals a critical role for neuropilin-1 in cranial neural crest cell migration in chick. *Dev. Biol.* 301, 227–239.

McLennan, R., Kulesa, P. M., 2010. Neuropilin-1 interacts with the second branchial arch microenvironment to mediate chick neural crest cell dynamics. *Dev. Dyn.* 239 (6), 1664–1673.

McLennan, R., Teddy, J. M., Kasemeier-Kulesa, J. C., Romine, M. H., Kulesa, P. M., 2010. Vascular endothelial growth factor (VEGF) regulates cranial neural crest migration in vivo. *Dev. Biol.* 339 (1), 114–25.

Mellott, D. O., Burke, R. D., 2008. Divergent roles for Eph and ephrin in avian cranial neural crest. *BMC Dev. Biol.* 8, 56.

Monsonego-Ornan, E., Kosonovsky, J., Bar, A., Roth, L., Fraggi-Rankis, V., Simsa, S., Kohl, A., Sela-Donenfeld, D., 2012. Matrix metalloproteinase 9/gelatinase B is required for neural crest cell migration. *Dev. Biol.* 364 (2), 162–77.

Nicolau, D. V., 2008. Spatial modelling of chemotaxis and its evolution. Ph.D. thesis, University of Oxford.

Nie, S., Kee, Y., Bronner-Fraser, M., 2009. Myosin-X is critical for migratory ability of Xenopus cranial neural crest cells. *Dev. Biol.* 335 (1), 132–42.

Noden, D. M., 1993. Spatial integration among cells forming the cranial peripheral nervous-system. *J. Neurobiol.* 24, 248–261.

Numerical Algorithms Group. d03pc - NAG Toolbox for Matlab documentation. http://www.nag.co.uk/numeric/MB/manual64_23_1/pdf/D03/d03pc.pdf

Olesnicky Killian, E. C., Birkholz, D. A., Artinger, K. B., 2009. A role for chemokine signaling in neural crest cell migration and craniofacial development. *Dev. Biol.* 333 (1), 161–72.

Painter, K., Sherratt, J., 2003. Modelling the movement of interacting cell populations. *J. Theor. Biol.* 225 (3), 327–339.

Painter, K. J., Maini, P. K., Othmer, H. G., 2000. A chemotactic model for the advance and retreat of the primitive streak in avian development. *Bull. Math. Biol.* 62 (3), 501–525.

Perfahl, H., Byrne, H. M., Chen, T., Estrella, V., Alarcón, T., Lapin, A., Gatenby, R. a., Gillies, R. J., Lloyd, M. C., Maini, P. K., Reuss, M., Owen, M. R., 2011. Multiscale modelling of vascular tumour growth in 3D: the roles of domain size and boundary conditions. *PloS One* 6 (4), e14790.

Plank, M. J., Simpson, M. J., 2012. Models of collective cell behaviour with crowding effects: comparing lattice-based and lattice-free approaches. *J. R. Soc. Interface* 9 (76), 2983–96.

Pratt, R. M., Larsen, M. A., Johnston, M. C., 1975. Migration of Cranial Neural Crest Cells in a Cell-Free Matrix. *Dev. Biol.* 44, 298–305.

Puré, E., Cuff, C., 2001. A crucial role for CD44 in inflammation. *Trends Mol. Med.* 7 (5), 213–21.

Reynolds, O., Brightmore, A. W., 1903. Papers on mechanical and physical subjects: the sub-mechanics of the universe. Collected work, Volume III. Cambridge University Press.

Rizzoti, K., Lovell-Badge, R., 2007. SOX3 activity during pharyngeal segmentation is required for craniofacial morphogenesis. *Development* 134, 3437–3448.

Roberts, I. S., Brenchley, P. E., 2000. Mast cells: the forgotten cells of renal fibrosis. *J. Clin. Pathol.* 53 (11), 858–62.

Roussos, E., Condeelis, J., Patsialou, A., 2011. Chemotaxis in cancer. *Nat. Rev. Cancer* 11 (8), 573–87.

Rupp, P. A., Kulesa, P. M., 2007. A role for RhoA in the two-phase migratory pattern of post-otic neural crest cells. *Dev. Biol.* 311 (1), 159–71.

Simpson, M., Baker, R., 2011. Corrected mean-field models for spatially dependent advection-diffusion-reaction phenomena. *Phys. Rev. E* 83 (5), 051922.

Simpson, M. J., Baker, R. E., McCue, S. W., 2011. Models of collective cell spreading with variable cell aspect ratio: A motivation for degenerate diffusion models. *Phys. Rev. E* 83 (2), 021901.

Simpson, M. J., Landman, K. A., Hughes, B., 2009a. Pathlines in exclusion processes. *Phys. Rev. E* 79 (3), 1–5.

Simpson, M. J., Landman, K. A., Hughes, B. D., 2009b. Distinguishing between directed and undirected cell motility within an invading cell population. *Bull. Math. Biol.* 71, 781–799.

Simpson, M. J., Landman, K. A., Hughes, B. D., 2009c. Multi-species simple exclusion processes. *Physica A* 388 (4), 399–406.

Simpson, M. J., Landman, K. A., Hughes, B. D., 2010a. Cell invasion with proliferation mechanisms motivated by time-lapse data. *Physica A* 389 (18), 3779–3790.

Simpson, M. J., Landman, K. A., Hughes, B. D., Fernando, A. E., 2010b. A model for mesoscale patterns in motile populations. *Physica A* 389 (7), 1412–1424.

Simpson, M. J., Landman, K. A., Hughes, B. D., Newgreen, D. F., 2006a. Looking inside an invasion wave of cells using continuum models: Proliferation is the key. *J. Theor. Biol.* 243, 343–360.

Simpson, M. J., Landman, K. A., Newgreen, D. F., 2006b. Chemotactic and diffusive migration on a nonuniformly growing domain: numerical algorithm development and applications. *J. Comput. Appl. Math.* 192, 282–300.

Simpson, M. J., Merrifield, A., Landman, K. A., Hughes, B. D., 2007a. Simulating invasion with cellular automata: connecting cell-scale and population-scale properties. *Phys. Rev. E* 76, 21918.

Simpson, M. J., Zhang, D. C., Mariani, M., Landman, K. A., Newgreen, D. F., 2007b. Cell proliferation drives neural crest cell invasion of the intestine. *Dev. Biol.* 302 (2), 553–568.

Singer, A., 2004. Maximum entropy formulation of the Kirkwood superposition approximation. *J. Chem. Phys.* 121 (8), 3657–66.

Singer, A., Schuss, Z., 2008. Partially reflected diffusion. *SIAM J. Appl. Math.* 68 (3), 844–868.

Spicher, A., Michel, O., 2007. Declarative modeling of a neurulation-like process. *BioSystems* 87 (2-3), 281–8.

Tallquist, M. D., Soriano, P., 2003. Cell autonomous requirement for PDGFR alpha in populations of cranial and cardiac neural crest cells. *Development* 130, 507–518.

Tanimoto, J., Hagishima, A., Tanaka, Y., 2010. Study of bottleneck effect at an emergency evacuation exit using cellular automata model, mean field approximation analysis, and game theory. *Physica A* 389 (24), 5611–5618.

Teddy, J. M., Kulesa, P. M., 2004. In vivo evidence for short- and long-range cell communication in cranial neural crest cells. *Development* 131, 6141–6151.

Theveneau, E., Mayor, R., 2011. Collective cell migration of the cephalic neural crest: the art of integrating information. *Genesis* 49 (4), 164–76.

Theveneau, E., Mayor, R., 2012. Neural crest migration: interplay between chemorepellents, chemoattractants, contact inhibition, epithelial-mesenchymal transition, and collective cell migration. *WIREs Dev. Biol.* 1, 435–445.

Trainor, P. A., 2005. Specification of neural crest cell formation and migration in mouse embryos. *Semin. Cell Dev. Biol.* 16 (6), 683–93.

Trewenack, A. J., Landman, K. A., 2009. A traveling wave model for invasion by precursor and differentiated cells. *Bull. Math. Biol.* 71, 291–317.

Vigmond, E., Vadakkumpadan, F., Gurev, V., Arevalo, H., Deo, M., Plank, G., Trayanova, N., 2009. Towards predictive modelling of the electrophysiology of the heart. *Exp. Physiol.* 94 (5), 563–77.

Weng, W., Chen, T., Yuan, H., Fan, W., 2006. Cellular automaton simulation of pedestrian counter flow with different walk velocities. *Phys. Rev. E* 74 (3), 036102.

Wynn, M. L., Kulesa, P. M., Schnell, S., 2012. Computational modelling of cell chain migration reveals mechanisms that sustain follow-the-leader behaviour. *J. R. Soc. Interface* 9 (72), 1576–88.

Yates, C. A., Baker, R. E., Erban, R., Maini, P. K., 2012. Going from microscopic to macroscopic on nonuniform growing domains. *Phys. Rev. E* 86 (2), 24–29.

Zhang, D. C., Brinas, I. M., Binder, B. J., Landman, K. A., Newgreen, D. F., 2010. Neural crest regionalisation for enteric nervous system formation: Implications for Hirschsprung's disease and stem cell therapy. *Dev. Biol.* 339 (2), 280–294.